BEHAVIORAL AND NEURAL MECHANISMS OF PATCH FORAGING

Gary A. Kane

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY
THE DEPARTMENT OF PSYCHOLOGY

Adviser: Jonathan D. Cohen

June 2018
Abstract

Serial stay-or-search decisions, in which one must choose to stick with a current resource or to search for a new, potentially better one, are ubiquitous across many domains. These decisions have been studied extensively in patch-foraging paradigms, in which animals, including humans, decide to stick with a depleting reward within a "patch," or to leave the current "patch" to search for a new one yielding greater rewards, but that comes at the cost of time and effort spent traveling to the new "patch." A variety of animals, ranging from invertebrates to birds to mammals, generally follow predictions of the optimal foraging theory, particularly the Marginal Value Theorem (MVT; Charnov, 1976), but animals have the tendency to overharvest, or stick with the depleting reward longer than is predicted by MVT. Common biases in intertemporal choice, such as decreasing marginal utility for larger rewards or discounting of future reward, have been hypothesized as the cause of overharvesting, but there have been few direct tests of whether these biases influence foraging behavior. From a neural perspective, multiple brain regions that contribute to stay-or-search decisions have been identified, including the anterior cingulate cortex (ACC), which is hypothesized to monitor the difficulty in deciding to stay vs. search, and the locus coeruleus (LC), which is hypothesized to regulate the decision process. This thesis extends knowledge of the behavioral and neural mechanisms of stay-or-search decisions. In Chapter 2, I describe a novel, operant chamber based patch foraging for rats. Similar to other animals, rats follow qualitative predictions of MVT: they stay longer in patches that yield greater rewards, and longer in all patches when the cost of traveling to a new patch is greater, but overharvest, staying patches longer than is predicted by MVT. In Chapter 3, I thoroughly characterized rat foraging behavior across a series of experiments: these revealed that in the context of foraging, rats exhibit time preferences similar to delay discounting paradigms, and that suboptimal decision making in foraging and delay discounting tasks is best explained by hyperbolic discounting. In Chapter 4, I begin to examine the neural mechanism of foraging decisions, investigating the role of the LC. Stimulation of LC-NE neurons impaired rats ability to focus on salient information in the task and impaired their ability to perform the task in general. In Chapter 5, I examine the function of the anterior cingulate. Rat ACC neurons increased in activity as rewards in a patch depleted and animals were more likely to leave patches. However, manipulation of ACC activity revealed
that ACC was neither necessary for adaptive foraging decisions nor sufficient to drive decisions to leave patches. Altogether, these studies contribute to our knowledge of how animals make foraging decisions from a behavioral and neural perspective.
Acknowledgments

First and foremost, I would like to thank my three advisors: Jon Cohen, Gary Aston-Jones, and Liz Gould. I am so grateful for the guidance and support you all gave me throughout different stages of my graduate school career: Liz, for teaching me how to ask questions, critically analyze ideas, and develop them into a research project; Jon, for teaching me to be thorough and to leave no stone unturned before coming to conclusions; and Gary, for imparting your wisdom on the LC, and for teaching me how to diagnose and troubleshoot any issues that arise in my research. I couldn’t have asked for a better experience in graduate school.

I am extremely grateful to the members of my committees. Nathaniel Daw, thank you for teaching me so much about statistics and computational modeling. Your support in all of my projects has really enriched my education at Princeton. Ilana Witten, thank you for your thoughtful feedback and your encouragement in this research. Jordan Taylor, thank you for taking part in my thesis defense and for your advice in searching for postdoc positions.

I am deeply indebted to the many current and former postdocs that have taught me so much about neuroscience, from the day-to-day activities to how to handle the stresses of graduate school. Elena Vazey, you were the first to take me under your wing, and from you, I learned all the basic skills I’ve used in my research over the past 4 years. Amitai Shenhav, Bob Wilson, Aaron Bornstein: whenever I’ve struggled with a project, you all have been the first to provide advice, and on so many occasions, have eased any anxieties I’ve had about my research. Thank you for your patience and dedication to the work that has gone into this thesis. Mike Shvartsman, when analyzing data, your voice will always be in the back of my head encouraging me to not cut any corners and to think critically about the statistical methods that I use. Morgan James, no matter how many projects your juggling, you always have a few minutes to hear my latest news (good or bad), and you do so with a smile on your face. Thank you so much for your guidance both in science and in life.

Keisha Craig, Carol Agans, Beth Porter, Tina McCoy, David Carter, Michele Giovannetti, and the psychology department/PNI administrative staff, thank you for kindly teaching me to use the copier on three separate occasions, for working on frustrating grant paperwork right up to the deadline, and for everything else you do to make our lives easier! To the LAR staff – Brian
Ludwig, Susie Chow, Laura Conour, Juan Mercado, and Paul Gudz – your dedication to caring for research animals has made PNI a really special place to do neuroscience research. It’s been a delight to work with you all.

To all of my lab mates in the last six years, your friendship has made me look forward to coming into the lab day in and day out. Adam Brockett, my first confidant and friend in Princeton, I’ll never forget the subway lunches, GC dinners and the many days of spring training in preparation for the upcoming Defense Mechanisms seasons. Maya Opendak, Matt Runkle, Liz LaMarca, Elise Cope, Sahana Murthy, Brandy Briones, and Patrick Monari, I’ll always remember the happy hours, the One Direction stickers, and the fun atmosphere you all created. Hannah Bowery, Mark Presker, Amy Kohtz, Jen Fragale, Mohsen Omrani, Caroline Pantazis, Aida Mohammadkhani, and Nupur Jain, from the moment I started to work in the lab at Rutgers, you all have been so welcoming. From the lab dinners in Highland Park to crashing SfN parties, it’s been a blast working with and getting to know you all. To the Cohen Lab – Laura Bustamante, Sebastian Musslick, Abby Novick, Ida Momennejad – I learned so much from you all at lab meetings, and really enjoyed “foraging” around the building for free food and the occasional beer at social hour.

To all of my friends in and outside of Princeton, I am grateful for all your support. My fellow Texan Nick Davy, thank you for screaming at the TV with me through these difficult years of Texas football and for all the late-night chats. Peter Humphries, for always being willing to go to Dbar (allegedly), on good days and bad. And to the rest of our GC crew for all the sports! Cameron Ellis, Laura Bustamante, Anne Mennen, Eva Fourakis, Alex Boldin, and the entire baby lab, thank you for never letting a single minute in Princeton become boring.

Aaron, thank you for all the graduate school advice and for generally being a great older brother. Being closer to you (both figuratively and geographically) has been a major highlight of the past six years. Mom and Dad, I can’t put into words how grateful I am to have such loving and supportive parents. My entire life, you’ve provided me with everything I need to succeed and have always allowed me to follow my passions. None of this would have been possible without you. Also, I think this deserves a Braums?

Casey Lew-Williams, thank you for moving to Princeton, and especially, for dragging Jessie along with you. Jessie, who would’ve thought that a random Tuesday night at Dbar would
become one of the most important moments of my life? Since we met, every single day has been more fulfilling, whether we’re tailgating at a football game we certainly aren’t going to, or spending the night in while I’m hopped up on Benadryl. Your energy and enthusiasm makes every experience more exciting, and you have an uncanny ability to make the stresses of graduate school simply disappear. I’m excited for our next journey together!
Table of Contents

Abstract..............................................................................................................................iii

Acknowledgments .............................................................................................................v

Chapter 1: Introduction ......................................................................................................1
  1.1 Foraging behavior.........................................................................................................1
  1.2 Foraging, exploration, and the locus coeruleus.............................................................3
  1.3 Foraging, cognitive control, and the anterior cingulate cortex......................................5
  1.4 Organization of the thesis.............................................................................................7

Chapter 2: Operant-Chamber Patch Foraging Task..............................................................9

Chapter 3: Examining the Cause of Overharvesting ...........................................................15
  3.1 Rats exhibit similar time preferences in foraging and two-alternative intertemporal choice
  tasks ..................................................................................................................................17
  3.2 Suboptimal decision-making in foraging and intertemporal choice tasks is best explained
  by quasi-hyperbolic discounting.......................................................................................24

Chapter 4: The Locus Coeruleus-Norepinephrine System Regulates Task Engagement 43
  4.1 Stimulation of LC-NE neurons impairs task performance..........................................49
  4.2 Stimulation of LC-NE neurons causes disengagement from important task variables53

Chapter 5: The Anterior Cingulate Cortex Monitors Decision Variables .............................60
  5.1 Rat ACC activity increases as rewards within a patch deplete.....................................61
  5.2 Chemogenetic manipulation of ACC activity does not influence foraging decisions .71

Chapter 6: General Discussion ..........................................................................................77

References ..........................................................................................................................81
Chapter 1: Introduction

1.1 Foraging behavior

All animals must forage to obtain the resources necessary for survival. For decades, researchers have studied foraging theory – how animals efficiently acquire resources and minimize costs, such as time and effort (Stephens & Krebs 1986, Stephens et al. 2007). This has spanned numerous domains of behavior, from basic ones such as what prey to eat to more complex ones such as how to respond to risk and how to balance competing pressures such as reproduction and predation (Stephens et al. 2007). Foraging models typically assume a few constraints (Stephens & Krebs 1986). First, options are encountered sequentially – animals choose to pursue an option that’s available now or to wait for something better to come along. Second, options must be encountered according to a poisson distribution – there may be variability in the amount of time to the next option, but the probability of an encounter in a given time interval must be constant. Lastly, animals must have complete information about the environment. That is, they must know the distribution of possible outcomes for each option (all rewards and costs for pursuing an option). There may be variability in these rewards or costs, but this variability must be known to the animal. Foraging is in direct contrast with exploration, in which animals do no know the full distribution of possible outcomes, and must explore, or sample options to learn their value (Kaelbling et al. 1996, Sutton & Barto 1998).

There are two basic foraging decision paradigms: a take it or leave it paradigm, in which animals must choose to accept or reject an offer of reward now or wait for another offer later, and a stay or leave paradigm, in which animals must decide to continue harvesting reward in a current location, termed a “patch.” Reward within the patch depletes over time, and animals must decide when to leave this patch to travel to a new, potentially richer one (Stephens & Krebs 1986). This thesis focuses on the latter paradigm. The optimal behavior in patch-foraging paradigms is described by the Marginal Value Theorem (MVT): leave the current patch when it depletes to the average reward rate across all patches in an environment (Charnov 1976, Stephens & Krebs 1986). Following this strategy ensures that animals will maximize the average reward rate across all patches, earning the maximum amount of reward in a given time. MVT makes two main
predictions regarding behavior in patch foraging tasks: i) animals should stay longer in patches that provide greater rewards, as it will take longer for these patches to deplete to the average reward rate threshold, and ii) animals should stay longer in all patches when the cost of traveling to a new patch is greater, as the average reward rate of the environment will be lower (owing to the longer travel time). Many species of animals, ranging from invertebrates to birds to mammals, including rodents, non-human primates, and humans, qualitatively follow these predictions (Barack et al. 2017, Bendesky et al. 2011, Constantino & Daw 2015, Flavell et al. 2006, Hayden et al. 2011, Nonacs 2001, Stephens 2008). However, all of these species tend to overharvest, or stay in patches longer than is predicted by MVT (Constantino & Daw 2015, Hayden et al. 2011, Nonacs 2001; Stephens & Krebs 1986).

The notion that animals attempt to maximize long-term rewards and that they tend to stay too long in patches – which has been described as overly patient behavior – is in contrast with animal behavior in standard intertemporal choice tasks. In these tasks, animals are simultaneously presented with 2 offers: a smaller, more immediate reward vs. a larger, delayed reward. Animals are often described as impulsive, consistently exhibiting a preference for choosing smaller, immediate rewards even when this option does not produce the maximum reward rate (Gallistel & Gibbon, 2000; Kacelnik, 1997; Kirby, 1997). This preference is often explained by temporal discounting models, in which the value of a reward is discounted by the length of the delay to receive the reward, and the animal selects the option with the greater discounted value (Gallistel & Gibbon, 2000; Kacelnik, 1997). However, these standard temporal discounting models have had little success explaining foraging behavior (Blanchard & Hayden, 2015; Carter et al., 2015; Carter & Redish, 2016). These findings have led to the hypothesis that animals perform foraging and intertemporal choice tasks in different ways. An alternative explanation is that standard temporal discounting models assume each decision is independent, and thus only consider rewards for the current decision, whereas in foraging tasks, the reward available on the current trial depends on the decision made on the previous trial. Thus, in making the current decision, it is important to consider future rewards in models of foraging.

In a separate attempt to reconcile the differences in behavior in these tasks, it has been proposed that animals instead maximize short-term reward rates, the reward rate in the current patch or the current decision (Bateson & Kacelnik 1996, Stephens 2002, Stephens & Anderson
These strategies are suboptimal in that they don’t consider the opportunity costs of the decision at hand, such as a delay in between receiving reward and the opportunity for reward. One advantage to adopting such short-sighted strategies is that animals may be able to better estimate the value of rewards in the short-term, thus making better decisions over this time than if they considered all future rewards, under which their value estimates and decisions would be subject to more uncertainty (Stephens, 2002; Stephens et al., 2004). In support of this theory, in two-alternative choice tasks, animals are typically insensitive to delays after reward and before the next decision can be made (Bateson & Kacelnik 1996, Blanchard et al. 2013, Pearson et al. 2010, Stephens & Anderson 2001).

Although there have been many hypotheses to explain overharvesting behavior, these hypotheses have not been directly tested in the context of foraging, for example, by manipulating the length of pre-reward and post-reward delays. Furthermore, predictions of these models across foraging and standard intertemporal choice tasks where models are adapted to account for sequential foraging decisions vs. choice vs simultaneously presented options in intertemporal choice tasks has not been conducted. In Chapter 2, I will describe a novel foraging task for rats that will be used to examine patch foraging behavior. In Chapter 3, I test rats in a series of foraging experiments to characterize the cognitive biases they present in foraging tasks, including their sensitivity to pre-reward vs. post-reward delays. Then, I apply models inspired by foraging theory and temporal discounting models to examine the cause of overharvesting. Lastly, I test rats in a standard intertemporal choice task to determine whether the same biases rats exhibit in foraging can explain behavior across contexts.

1.2 Foraging, exploration, and the locus coeruleus

The locus coeruleus is a nucleus of norepinephrine producing cells in the brainstem, consisting of 15-35,000 cells in humans and ~2000 cells in rats (Aston-Jones & Cohen 2005, Goldman & Coleman 1981, Mouton et al. 1994, Ohm et al. 1997). Despite its small size, the LC is a wide reaching neuromodulatory system with efferent projections throughout the brain, and it serves as the sole source of norepinephrine for the cerebral cortex (Aston-Jones & Cohen 2005, Moore & Bloom 1979). Early studies of the physiology of LC-NE neurons revealed relatively low
spontaneous firing rates, with activity that closely followed sleep-wake cycles: LC cells are silent
during deeper stages of sleep and more active in awake animals (Aston-Jones & Bloom 1981a).
LC neurons also exhibit large phasic responses to salient sensory stimuli, suggesting that LC
neurons regulate orienting and arousal (Aston-Jones & Bloom 1981b). Further examination of the
cognitive function of LC-NE neurons in monkeys showed that these phasic bursts could be
elicited by conditioned stimuli that predict rewards (Aston-Jones et al. 1994), and that phasic
responses are more closely tied to behavioral responses, such as a lever press to receive
reward, than to the stimulus itself. These findings indicated that LC phasic responses facilitate
goal-directed behavioral actions (Bouret & Sara 2004, Clayton et al. 2004, Rajkowski et al. 2004).
Tonic activity also correlated with behavioral performance. Increased activity impairs focus on
relevant cues – tonic activity is elevated during periods of poor performance and increased
distractibility (Aston-Jones et al. 1994). But increased tonic activity may facilitate learning of new
reward contingencies. In response to changes in task reward contingencies, such as a reversal,
LC phasic responses to the former cue diminish and tonic activity increases. As the new reward
contingency is learned, tonic activity diminishes and LC neurons exhibit phasic responses to the
new rewarded cue (Aston-Jones et al. 1997).

Based on these findings, the Adaptive Gain Theory (AGT) hypothesized that LC neurons
exhibit two distinct modes of activity: phasic and tonic, which function to differentially regulate
gain or responsivity of downstream circuits (Aston-Jones & Cohen 2005). In the phasic mode, low
levels of spontaneous firing reduce responsivity of downstream circuits to distracting information
and burst responses rapidly increase responsivity facilitating behavioral responses to task-
relevant stimuli and optimizing performance within a task (Aston-Jones et al. 1994, Clayton et al.
2004, Rajkowski et al. 2004). In the tonic mode, increased levels of spontaneous firing increase
the likelihood that downstream circuits will respond to any incoming stimuli, including information
not relevant to the current task, but possibly to alternative tasks. This leads to increased
distractibility and poor performance within a task, but can facilitate exploration and learning of
other tasks (Aston-Jones et al. 1997).

AGT predicts that LC firing modes are determined by task utility: when rewards are
abundant for the current task, the LC is pushed towards the phasic mode to exploit the known
resources; when utility for a task becomes low, LC is pushed towards the tonic mode, enabling
exploration for new opportunities for reward (Aston-Jones & Cohen 2005). Computational modeling studies have shown that the balance between tonic and phasic LC firing can optimize performance in perceptual decision-making tasks and that driving changes in the balance between tonic and phasic firing can improve performance in tasks that involve changing reward contingencies (Brown et al. 2004, Eckhoff et al. 2009, Gilzenrat et al. 2002, McClure et al. 2006, Usher et al. 1999).

To date, the only empirical support for predictions of AGT have come from studies using pupil diameter measurements as an indirect measure of LC-NE activity (Eldar et al. 2013, Gilzenrat et al. 2010, Jepma & Nieuwenhuis 2011, Nassar et al. 2012). Pupil diameter tracks LC activity well – single spikes in LC neurons evoke pupil dilations, and electrical stimulation of LC neurons induces pupil dilation (Aston-Jones & Cohen 2005, Joshi et al. 2016). Pupillometry studies in humans performing cognitive tasks provide support for multiple facets of AGT, including that baseline pupil diameter tracks changes in reward rate (Gilzenrat et al. 2010), that pupil diameter may predict decisions to explore vs. exploit (Jepma & Nieuwenhuis 2011), and that pupil diameter tracks neural gain (Eldar et al. 2013). However, pupil measurements are also influenced by other neural circuits, most notably cholinergic systems (McGinley et al., 2015). To definitively test the predictions of AGT, direct measurements coupled with selective manipulation of LC-NE activity are needed. Chapter 4 will provide the first step in this direction, testing the effect of tonic LC stimulation on foraging decisions. Although stay vs. leave foraging decisions are distinct from exploit-explore decisions, the foraging task still enables testing of whether LC stimulation reduces focus on relevant task-variables. Further studies will be necessary to assess whether LC tonic activity naturally increases as utility within a task wanes to drive exploration.

1.3 Foraging, cognitive control, and the anterior cingulate cortex

The anterior cingulate cortex (ACC) is a hub of cognitive control (Heilbronner & Hayden 2016, Shenhav et al. 2013). An initial theory to explain its role in controlled behavior hypothesized that ACC monitors conflict in information processing. According to this hypothesis, as conflict increases – incoming information provides evidence for multiple potential responses – ACC activates to signal the need to exert cognitive control to reduce this interference and select the
best option (Botvinick 2007; Botvinick et al. 2001, 2004). Neuroimaging studies have widely supported this view, finding that ACC activity is increased when response conflict is greater, such as when participants must override a default response (Botvinick et al. 2001, Kerns et al. 2004, Shenhav et al. 2014). Until recently, conflict signals have rarely been found in single-unit recordings (Heilbronner & Hayden 2016, Nakamura et al. 2005), but they have been recently reported in both monkeys and rodents (Bryden et al. 2018, Ebitz & Platt 2015).

Single-unit recordings also suggest that ACC may play additional more central roles in decision-making. ACC neurons are sensitive to reward under many different circumstances across different tasks (Heilbronner & Hayden 2016, Kennerley et al. 2011, Seo & Lee 2009). Recordings as well as inactivations and lesion studies suggest a role for ACC in learning (Alexander & Brown 2011, Kennerley et al. 2006, Rushworth et al. 2003, Shima & Tanji 1998). ACC neurons have also been found to encode the value of offers of reward (Blanchard & Hayden 2014, Cai & Padoa-Schioppa 2012, Hayden et al. 2011). The hypothesis that ACC is primarily involved in decision-making has been strengthened by studies of foraging behavior.

Single-unit recordings in monkeys and neuroimaging studies in humans performing foraging tasks have both shown that ACC activity increases as the offer of a reward is low relative to the value of searching for better alternatives (Blanchard & Hayden 2014, Hayden et al. 2011, Kolling et al. 2012, Meder et al. 2016). This signal has been interpreted two different ways: that ACC’s main function is to signal the value of switching to an alternative course of action (leave a patch rather than the default option of stay; Kolling et al. 2012, 2016; Rushworth et al. 2012), or that ACC signals choice difficulty or the similarity in the value of staying vs. value of leaving a patch. The latter case is consistent with the idea that ACC signals conflict, or with an updated account: that ACC signals the value of exerting control (Shenhav et al. 2013, 2014, 2016a,b). In patch foraging tasks, it is difficult to distinguish the value of leaving from the level of conflict and/or the value of exerting control. As an agent harvests from a patch, the value of staying in the patch decreases as the value of leaving the patch increase. As the value of staying and leaving approach one another, the choice to stay vs. leave becomes more difficult, and more control is needed to make adaptive foraging decisions (Shenhav et al. 2013, 2014). The value of leaving and choice difficulty increase together until reward depletes to the animal’s leaving threshold (Shenhav et al. 2014). This is the point of maximal choice difficulty, but also the point at which the
MVT predicts an agent should leave the patch, and thus is the maximum observed value of leaving. To distinguish these quantities, agents need to be dropped into patches that begin with rewards that are well below their leaving threshold. In this scenario the value of leaving the patch is high, but the decision to leave should be easy. Shenhav et al. (2014, 2016b) designed a foraging task for which the value of leaving was high, but choice difficulty was low on 50% of trials. In this task, ACC activity was greatest at the point of maximal choice difficulty, not the point of maximum value of leaving, indicating that ACC tracks choice difficulty, not value of leaving. This evidence is still only correlational; manipulation of ACC activity will be necessary to determine the causal function of the ACC.

There are many tools to selectively alter neural activity in rodent models, but whether rodents have a region homologous to monkey and human ACC, and which area of cortex is most homologous is still a topic of active debate (Heilbronner & Hayden 2016). The most likely candidate region is cingulate area (Cg) of the rodent cortex (Heilbronner & Hayden 2016, Passingham & Wise 2012). There is some new evidence for similar functions in rat and primate ACC. A recent study found that rat ACC neurons exhibit a single-unit correlate of the error-related negativity signal, a canonical signal attributed to the ACC, and consistently found in humans (Botvinick et al. 2001; Hyman et al. 2017). As mentioned above, conflict signals have been found in single neurons in a similar area of rat ACC (Bryden et al. 2018). In Chapter 5, I seek to confirm whether rodent ACC performs a similar function in the foraging task using single and multi-unit recordings, and I test the causal role of the ACC in foraging by performing selective activations and inhibitions in behaving rats.

1.4 Organization of the thesis

In Chapter 2, I present a novel, operant chamber based patch foraging task for rats. Similar to previous studies, rats were found to follow qualitative predictions of MVT, but overharvest. In Chapter 3, the cause of overharvesting behavior was examined. Rats were tested in a series of experiments with a variety of manipulations to the foraging environment. Rats not only overharvested, but also exhibited time preferences: they preferred immediate rewards over delayed rewards, even when the delays had no effect on reward rate. Rats exhibited similar
biases in a two-alternative intertemporal choice task. I examined the ability of many models of intertemporal choice behavior to explain suboptimal decision-making in both paradigms, finding that maximizing discounted future rewards via a hyperbolic discounting function provided the best explanation across both tasks. In Chapter 4, I begin to explore the neural mechanisms of foraging, testing the role of the LC. As predicted by AGT, stimulating LC-NE neurons increased decision noise – decisions to stay vs. leave were more random and less tied to the MVT. LC stimulation also impaired rats’ ability to perform the task in general. In chapter 5, I examine the function of the ACC in foraging. Rat ACC neurons, similar to ACC in human and monkeys, increased in activity as reward within a patch depleted and rats were more likely to leave a patch. When the function of the ACC was tested using selective manipulations, it was found that rat ACC was neither necessary for adaptive foraging decisions nor sufficient to drive decisions to leave patches, indicating that ACC may only perform a monitoring role in foraging decisions. Finally, in Chapter 6, I discuss the implications of these studies.
Chapter 2: Operant-Chamber Patch Foraging Task

To examine how animals make stay-or-leave decisions and the neural mechanisms supporting this behavior, I developed an operant conditioning chamber based patch-foraging task for rats. In this task, rats performed a series of trials in which they had to decide to stay in or leave a “patch”. Rats harvested from a patch by pressing a lever at the front of the chamber. Each patch started at a particular reward volume and depleted – it yielded smaller reward volume – with each consecutive lever press. To leave a patch, rats used a nose poke at the back of the chamber. Nose poking caused the lever to retract and initiated a delay period simulating the time to travel to a new patch. After the delay, a lever on the opposite side extended (i.e. if the left lever was extended previously, the right lever extended now), and rats could resume “harvesting” from this new patch, with replenished reward volume. In an initial experiment, I tested whether rats would follow the basic predictions of MVT: i) in a given environment, animals should stay longer in patches that yield more rewards, as these patches will take longer to deplete to the average reward rate across all patches, and ii) if the cost of traveling to a new patch is greater, animals should stay longer in all patch types, as the average reward rate across all patches will be lower (Charnov 1976).

Methods

Animals

Adult Long-Evans rats (Charles River, Kingston, NY; n = 32) were used. Rats were housed on a reverse 12 h/12 h light/dark cycle (lights off at 7 a.m.). All behavioral testing was conducted during the dark period. Throughout behavioral testing, rats were food restricted to maintain a weight of 85% to 90% ad-lib feeding weight, and were given ad-lib access to water. All procedures were approved by the Princeton University Institutional Animal Care and Use Committee.

Operant Training

Rats were initially trained to press a lever for a 10% sucrose water reward on an FR1 paradigm, in which each lever press was rewarded with 100 μL of sucrose water. Next, rats were
trained to nose poke when reward was unavailable. In this stage, the lever stopped yielding reward on randomly selected intervals between 4-12 lever presses. To regain access to a rewarding lever, rats were required to travel to the back of the chamber and use a nose poke, which caused the initial lever to retract and the lever on the opposite side of the chamber to extend, which yielded reward. Next, a delay (first 5 s, then 10 s) that simulated the time to travel between patches was introduced between the nose poke and the opposite lever extending. Finally, rats were trained and tested on the foraging task (described below). To move on to the next stage of training, rats were required to pass a criterion of at least 100 rewarded lever presses in a 1 h session.

**Task Design**

This task simulated foraging in a patchy environment, resembling the task used for monkeys in Hayden et al. (2011). In a series of trials, rats repeatedly decided to stay in the patch to harvest a depleting reward source by pressing a lever or to leave the patch by traveling to the back of the chamber to nose poke (incurring a cost of time to travel to a new patch). At the beginning of each trial, cue lights above the activated lever and the nose poke illuminated, indicating that the rat could make a decision to harvest reward from the patch (lever press) or to travel to a new patch (nose poke). If rats pressed the lever to harvest from the patch, a cue light turned on in a reward port adjacent to the lever. Reward (10% sucrose water) was delivered as soon as the rat’s head entered the reward port. The size of the reward was cued using a tone, with longer duration signaling greater reward. If rats did not enter the reward port within 5 s, the cue light turned off, rats lost the opportunity to receive reward on the trial, and the trial was scored as an omission. After reward consumption, there was an inter-trial delay period, after which the next trial would begin. With each consecutive harvest trial, the reward volume was reduced via an exponential decay to simulate patch depletion. If rats nose poked to travel to a new patch, the lever retracted for a delay period, simulating the time to travel to a new patch. After this delay, the opposite lever extended, and rats could begin to harvest from a new replenished patch (Figure 2.1).
Figure 2.1. Diagram of the foraging task. Rats press a lever to harvest reward from the patch then receive reward in an adjacent port following a handling time delay. After receiving reward, there is an inter-trial interval (post-reward delay) before rats can make their next decision. Rats can leave the patch by nose poking in the back of the chamber (trial n+2), which initiates a delay simulating time to travel to the next patch, after which, rats can harvest from a new replenished patch.
In this initial experiment, rats encountered three patch types at random, starting with 60, 90, or 120 \( \mu \text{L} \) of reward, each depleting according to the same exponential decay function. In separate sessions, rats were tested on two travel time delays: 10 or 40 s.

**Data analysis**

To examine whether rats followed qualitative predictions of MVT, I examined the number of harvests per patch and the reward rate at which rats left patches. As each decision to stay or leave occurs within a defined trial, the number of harvests serves as a good proxy for time spent in the patch. The reward rate at which rats leave each patch is a measure of rats leaving threshold. MVT predicts that the threshold should be constant across all patches within an environment. I tested for changes in harvests per patch and in the reward rate threshold at which rats left patches due to patch type and travel time using mixed effects models, with random effects for both patch type and travel time. I tested for deviations from optimal number of harvests per patch using a single sample t-test of observed number of trials per patch against the optimal, averaging over patch types and travel times.

**Optimal behavior**

The optimal behavior is that which maximizes the average reward rate across all patches. Per MVT, the average reward rate across patches, \( R \), is defined by the sum of rewards across all patch types over the total time spent foraging:

\[
R = \frac{1}{n} \sum_{i=1}^{n} g_i(t_i) \left( \frac{1}{\tau} + \frac{1}{n} \sum_{i=1}^{n} t_i \right),
\]

where \( n \) is the number of patch types (\( n = 3 \)), \( t_i \) is the time spent in patch type \( i \), and \( g_i(t_i) \) is the cumulative reward obtained in patch type \( i \) over the course of time \( t_i \), and \( \tau \) the travel time between patches. The optimal behavior is to choose the times \( t_i \) that maximize reward rate:
In the present experiment, \( r_i \) represented the number of harvests in each patch; the \( t_i \) that maximize reward rate were found using exhaustive search over all possible \( t \).

**Results**

Rat foraging behavior qualitatively followed predictions of MVT: i) rats stayed for more trials in patches that started with greater reward volume (\( \beta = .826, SE = .075, p < .001 \)), and ii) rats stayed longer in all patch types in the 40 s versus the 10 s travel time (\( \beta = 1.265, SE = .144, p < .001 \); Figure 2.2A). Consistent with MVT, rats adopted a single reward rate threshold to leave patches – there was no effect of patch type on the patch leaving threshold (\( \beta = .004, SE = .010, p = .647 \)) – but the threshold was reduced due to the longer travel time (\( \beta = .180, SE = .019, p < . \).

\[
R^* = \arg \max_{t_i} \frac{1}{n} \sum_{i=1}^{n} g_i(t_i) - \tau + \frac{1}{n} \sum_{i=1}^{n} t_i.
\]

**Figure 2.2.** A) The number of harvests per patch in the three different patch types (x-axis) and two travel time conditions (solid vs. dashed line). B) The reward rate in uL/s at which rats left patches in the three different patch types and two travel time conditions. Points and error bars represent mean and standard error of rat behavior. Red lines indicate optimal behavior, as predicted by MVT.
However, rats overharvested, staying in all patch types for longer than is quantitatively predicted by the MVT ($t(7) = 11.341, p < .001$).

Ultimately, rats performed this foraging task – following qualitative predictions of MVT but overharvesting – similar to a variety of other species (Calhoun & Hayden 2015, Constantino & Daw 2015, Pyke 1978, 1984; Stephens & Krebs 1986). In Chapter 3, I will use this task to examine the cause of overharvesting, then again in Chapters 4-5 to study the function of the locus coeruleus and anterior cingulate in foraging.
Chapter 3: Examining the Cause of Overharvesting

Introduction

In foraging paradigms, animals must choose to stick with a depleting reward within a patch, or incur a cost of time and effort to travel to a new patch. In these paradigms, animals qualitatively follow predictions of the optimal behavior described by the Marginal Value Theorem (MVT; Charnov 1976) – leave a patch when it depletes to the average reward rate across all patches. However, animals consistently overharvest, or stick with the depleting reward longer than is optimal (Constantino & Daw 2015, Hayden et al. 2011, Nonacs 1991, Stephens & Krebs 1986). Many hypotheses to explain overharvesting have been proposed, including subjective costs, such as an aversion to leaving a patch (Carter & Redish 2016, Wikenheiser et al. 2013) or diminishing marginal utility, by which larger rewards in a new patch are not perceived as proportionally larger than smaller depleted rewards in the current patch (Constantino & Daw 2015). However, these hypotheses have not been directly tested in the context of foraging. Furthermore, these hypotheses fail to generalize to contexts outside of foraging, such as standard intertemporal choice tasks, in which animal exhibit a consistent preference to accept smaller, more immediate rewards compared to larger, delayed rewards.

One potential explanation for such time preferences is that instead of maximizing reward rate across all patches or all decisions (per MVT), animals seek to maximize reward rate in the current patch or on the current decision, ignoring opportunity costs of the decision at hand. These rules are often termed short-term rate maximization or short-sighted rules (Bateson & Kacelnik 1996; Stephens 2002, 2008; Stephens & Anderson 2001, Stephens et al. 2004). In support of this hypothesis, animals typically only attend to delays between making decisions and receiving rewards, and are insensitive to post-reward delays – that is, delays between receiving reward and making the next decision (Bateson & Kacelnik 1996, Blanchard et al. 2013, Gallistel & Gibbon 2000, Kacelnik 1997, Pearson et al. 2010, Stephens & Anderson 2001). One advantage to adopting such short-sighted rules is that animals may be able to better estimate the value of rewards in the short-term, thus making better decisions over this time than if they considered all future reward (Stephens 2002, Stephens et al. 2004).
Preference for small, immediate rewards over larger, delayed rewards in two-alternative choice tasks are commonly explained using delay discounting models, which assume the value of a reward is discounted due to the time the animal must wait to receive the reward. Normative discounting models suggest that rewards should be discounted according to an exponential function, one for which the value of a reward decreases by a constant fraction per unit time. However, animal behavior is better described by hyperbolic discounting, in which the value of a reward decreases more in the short-term than in the long-term, leading to inconsistent time preferences (Ainslie 1992, Gallistel & Gibbon 2000, Kacelnik 1997, Kirby 1997). Hyperbolic discounting can be formalized in a number of ways. Its most common form, \[ V = \frac{A}{1 + kD} \], has been applied to animal choices in numerous behavioral studies (Gallistel & Gibbon 2000, Kacelnik 1997, Kirby 1997, Mazur 1984). Another method, quasi-hyperbolic discounting (Laibson 1997), hypothesizes two-systems: a shallow exponential discounting system that captures long-term time preferences and a constant that determines an additional weight given to immediate reward. There is neural evidence for the existence of two discounting systems: limbic activation correlates with choices of smaller immediate rewards whereas frontoparietal activation is associated with choices for delayed options (McClure et al. 2004, 2007). Temporal discounting may explain overharvesting in foraging behavior – discounting reward received in future patches may lead animals to over-value reward in the current patch.

Lastly, uncertainty in interval timing has been proposed as a mechanism for suboptimal decision-making, and may play an important role in intertemporal choice. Animals’ time perception follows the scalar property: accuracy in estimating time intervals decreases with the length of the interval (Gallistel & Gibbon 2000, Gibbon 1977, Gibbon et al. 1984). In timing tasks, including foraging tasks for which decisions to stay or leave directly depend on the estimation of elapsed time, animals account for their uncertainty in estimating time intervals in a near optimal fashion (Balci et al. 2009, Balci et al. 2011, Brunner et al. 1992, Çavdaroğlu et al. 2014, Freestone et al. 2015). In foraging tasks with exogenous variability in time intervals (e.g. variable travel time or post-reward delay), animals similarly adjust their behavior in response to timing uncertainty (Brunner et al. 1996, Kacelnik & Todd 1992). Furthermore, suboptimal perceptual decision-making in humans is influenced by timing uncertainty (Balci et al. 2011). Uncertainty in
estimation of reward rates due to inaccurate timing may cause animals to overharvest, however this has not been tested in the absence of exogenous variability in time intervals.

To determine the cause of overharvesting, I characterized behavior of rats in a series of patch foraging experiments. Consistent with previous studies, rats followed qualitative predictions of MVT but overharvested. In certain environments, rats deviated from predictions of MVT, changing their behavior based on the placement of delays (pre- vs. post-reward). I further tested rats time preferences, examining their sensitivity to post-reward delays in a foraging task and a two-alternative choice task. Contrary to previous studies, rats were sensitive to post-reward delays in both contexts. To understand the mechanism driving these decision biases, I used computational model comparison to test quantitative predictions of several possible accounts for this pattern of behavior, including all of the hypotheses described above. Quasi-hyperbolic discounting provided the best explanation for rat behavior across all experiments. These data suggest that rats exhibit similar biases in foraging and two-alternative choice paradigms, and maximization of discounted future rewards may be a common mechanism for suboptimal decision-making across paradigms.

3.1 Rats exhibit similar time preferences in foraging and two-alternative intertemporal choice tasks

Methods

Animals

Adult Long-Evans rats were used (Charles River, Kingston, NY). One group of eight rats participated in four foraging experiments, a different set of eight rats were tested on a final foraging experiment then the two-alternative choice task. Rats were housed on a reverse 12 h/12 h light/dark cycle. All behavioral testing was conducted during the dark period. Rats were food restricted to maintain a weight of 85-90% ad-lib feeding weight, and were given ad-lib access to water. All procedures were approved by the Princeton University and Rutgers University Institutional Animal Care and Use Committee.
Foraging Task

Animals were trained and tested as in Chapter 2. Rats were first trained to lever press for 10% sucrose water on an FR1 reinforcement schedule. Once exhibiting 100+ lever presses in a one hour session, rats were trained on a sudden patch depletion paradigm – the lever stopped yielding reward after 4-12 lever presses – and rats learned to nose poke to reset the lever. Next rats were tested on the full foraging task.

A diagram of the foraging task is shown in Figure 2.1. On a series of trials, rats had to repeatedly decide to lever press to harvest reward from the patch or to nose poke to travel to a new, full patch, incurring the cost of a time delay. At the start of each trial, a cue light above the lever and inside the nose poke turned on, indicating rats could now make a decision. The time from cues turning on until rats pressed a lever or nose poked was recorded as the decision time (DT). A decision to harvest from the patch (lever press) yielded reward after a short handling time delay (HT). Reward (sucrose water) was delivered when the rat entered the reward magazine. The next trial began after an inter-trial interval (ITI). In the first 4 foraging experiments, to control the reward rate within the patch, the length of the ITI was adjusted based on the DT of the current trial, such that the full duration of all harvest trials was equivalent. With each consecutive harvest, the rat received smaller volume of reward to simulate depletion from the patch. A nose poke to leave the patch caused the lever to retract for a delay period simulating the time to travel to a new patch. After the delay, the opposite lever extended, and rats could harvest from a new, replenished patch.

In four separate experiments, I manipulated different variables of the foraging environment. In the travel time experiment, rats encountered three different patch types within each session, starting with 60, 90, or 120 μL of reward. In separate sessions, rats were tested on either a 10s or 30s travel time. After initial training, rats were trained for 5 consecutive sessions and tested on an additional 5 consecutive sessions for each travel time, counterbalanced across rats. In the depletion rate experiment, rats encountered one patch type in each session, starting with 90 μL and depleting at either 8 or 16 μL per trial. Different depletion rates were tested in separate sessions in a similar design to the travel time experiment. In the scale experiment, the magnitude of rewards and delays were manipulated, such that in one condition, the size of
rewards and the length of delays was twice that of the other. In the handling time experiment, the placement of delays was manipulated. The total time to harvest reward remained constant, but in one condition there was no pre-reward delay and ~13s post reward delay, and the other, there was a 3s pre-reward delay, simulating the time to obtain reward once deciding to harvest, and ~10s post-reward delay. Parameters for each of the first four experiments can be found in Table 3.1. In the final foraging experiment (post-reward delay experiment), the post-reward delay or ITI was manipulated 3 s in one condition vs. 12 s in the other. In this experiment, the length of harvest trials was longer due to the longer post-reward delay, such that reward rate in the patch was reduced in this condition. For each experiment, rats were trained on a specific condition for 5 days, then tested for 5 days. Conditions within experiments were counterbalanced. Rat foraging behavior was assessed using mixed effects models. In the travel time experiment, I assessed the effect of starting volume of the patch and the travel time on number of harvests per patch, with random intercept for each subject and random slope for both variables. In all other experiments, I assessed the effect of experimental condition on harvests per patch, with random intercept for each subject and random slope for the effect of experimental condition.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Condition</th>
<th>Start Reward</th>
<th>Depletion Rate</th>
<th>HT</th>
<th>Harvest Time</th>
<th>Travel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>travel time</td>
<td>10s 30 s</td>
<td>60, 90, or 120 µL</td>
<td>-8 µL</td>
<td>0 s</td>
<td>10 s</td>
<td>10 s 30 s</td>
</tr>
<tr>
<td>depletion rate</td>
<td>-8 µL -16 µL</td>
<td>90 µL</td>
<td>-8 µL -16 µL</td>
<td>0 s</td>
<td>12 s</td>
<td>12 s</td>
</tr>
<tr>
<td>scale</td>
<td>90 µL/10 s 180 µL/20 s</td>
<td>90 µL 180 µL</td>
<td>-8 µL -16 µL</td>
<td>0 s</td>
<td>10 s 20 s</td>
<td>10 s 20 s</td>
</tr>
<tr>
<td>handling time</td>
<td>0 s 3 s</td>
<td>90 µL</td>
<td>-8 µL</td>
<td>0 s 3 s</td>
<td>15 s</td>
<td>15 s</td>
</tr>
</tbody>
</table>

Table 3.1. Parameters for each of the first 4 foraging experiments. HT = handling time. Harvest time = time to make a decision to harvest + HT + inter-trial interval. To control reward rate in the patch, the inter-trial interval was adjusted relative to the decision time to hold the harvest time constant.
Two-alternative choice task

Rats were immediately transferred from the foraging task to the two alternative choice task with no special training. This task consisted of a series of episodes that lasted 20 trials. At the beginning of each episode one lever was randomly selected as the shorter-sooner lever, yielding 40 μL of reward following a 1 s delay. The other lever (longer-later lever) was initialized to yield a reward of 40, 80, or 120 μL after a 1, 2, 4 or 6 s delay. For the first 10 trials of each episode, only one lever extended, and rats were forced to press that lever to learn the reward value and delay associated with the lever. The order of forced trials was randomly selected such that, within the first six trials, rats experienced three trials on one lever and three trials on the other, and within the last four trials (trials 7-10), rats experienced two trials on one and two trials on the other lever. This series was chosen to reduce the possibility of rats developing a perseveration bias. For the remaining 10 trials of each episode, both levers extended, and rats were free to choose the option they prefer. At the beginning of each trial, cue lights turned on above the lever indicating rats could now make a decision. Once the rat pressed the lever, the cue light turned off, and the delay period was initiated. A cue light turned on in the reward magazine at the end of the delay period, and rats received reward as soon as they entered the reward magazine. Reward magnitude was cued by light and tone. Following reward delivery, there was an ITI before the start of the next trial. At the completion of the episode, the levers retracted, and rats had to nose poke to begin the next episode (with a different larger-later reward and delay).

Two-alternative choice data was analyzed using a mixed effects logistic regression, examining the the effect of larger-later reward value and larger-later delay, both as categorical variables, on rats choices. Post-hoc tests were conducted using the phia package in R (de Rosario-Martinez 2015).

Results

With longer travel time, the long-run average reward rate is lower, thus MVT predicts rats should stay in patches longer. In the travel time experiment, rats encountered three different patch types within sessions, which started with varying amount of reward (60, 90, or 120 μL) and
depleted by the same rate (8 μL/trial). Between sessions, rats were tested on either a 10 s or 30 s travel time delay following their decision to leave the patch. As predicted by MVT, rats stayed longer in patch types that started with larger reward volume, indicated by more harvests per patch ($β = 118.266$, $SE = 2.753$, $p < .001$). Rats stayed longer in all patches with longer travel time ($β = 1.907$, $SE = .306$, $p < .001$). However, rats overharvested relative to predictions of MVT, staying in patches for an average of 3.25 and 3.55 trials longer in the 10 s and 30 s travel conditions (Figure 3.2A).

Quicker reward depletion causes the local reward rate to deplete to the long-run average reward rate quicker, such that MVT predicts earlier patch leaving. Within sessions, rats encountered a single patch type (starting volume of 90 μL), which depleted at a rate of either 8 or 16 μL/trial, tested between sessions. As predicted by MVT, rats left patches earlier when they depleted more quickly ($β = 2.589$, $SE = .155$, $p < .001$). But, again, rats stayed in patches longer than is predicted by MVT by 2.80 trials with 8 μL depletion and 1.21 trials with 16 μL depletion (Figure 3.2B).

Since MVT is concerned with reward rates, doubling the size of rewards and the time to obtain those rewards while keeping reward rate across trials equivalent, has no effect on trial-by-trial behavioral predictions. The scale of rewards and delays were manipulated in this manner: in one condition, patches started with 90 μL, depleted at a rate of 8 μL/trial, and the duration of harvest trials and travel time between patches was 10 s. In the other condition, patches started with 180 μL, depleted at a rate of 16 μL/trial, and the duration of harvest trials and the travel time between patches was 20 s. Rats overharvested in both conditions, and contrary to predictions of MVT, rats stayed in patches significantly longer when given larger rewards with longer delays ($β = 1.937$, $SE = .193$, $p < .001$; Figure 3.2C).

Lastly, this cohort of rats was tested in two environments for which the harvest time and travel time were matched trial-by-trial, but one condition had a 3 s handling time delay (pre-reward delay) with shorter post-reward delay. MVT is not sensitive to the placement of delays, and predicts no change in behavior across these environments, however, rats left patches earlier
in the environment with a 3s pre-reward delay and shorter post-reward delay ($\beta = 2.345$, SE = .313, p < .001; Figure 3.2D).

There are a number of explanations for the overharvesting observed in the travel time and depletion rate experiments. Staying longer in patches with larger rewards and longer delays but equivalent reward rate suggests some bias in perception of reward value or time. A constant subjective cost, as if the effort for leaving a patch costs additional units of reward, could not account for the change in the degree of overharvesting across conditions. Furthermore, the change in harvests per patch due to the handling time indicates that rats exhibit a bias in processing pre- vs. post-reward delays. To further examine time preferences, a new cohort of rats was tested in two foraging environments, one with a short post-reward delay (3 s), and the other with a long post-reward delay (12 s). The time of harvest trials was not held constant; the longer

**Figure 3.2.** Rat foraging behavior in the A) travel time, B) depletion rate, C) scale, and D) handling time (pre-reward delay) experiments. In A, points and error bars represent mean and standard error. In B-D, points and connecting lines represent behavior of each individual rat. Red lines indicate optimal behavior (per MVT).
post-reward delay reduced reward rate within patches, such that MVT predicts that rats should leave patches earlier, whereas short-term rules that ignore post-reward intervals predict no effect of the increased post-reward delay. As in other experiments, rats overharvested, staying in patches for 3.33 trials and 3.66 trials longer than predicted by MVT in the 3 s and 12 s post-reward delay conditions, respectively. As predicted by MVT and not short-term rules, rats were sensitive to the post-reward delay, leaving patches earlier in the 12 s delay environment ($\beta = 1.411$, SE $= .254$, $p < .001$; Figure 3.3A).

Lastly, I examined whether the same rats exhibit similar time preferences in a two-alternative choice task. On each trial, rats pressed either the left or right lever to receive a smaller-sooner reward of 40 $\mu$L after a 1 s delay or a larger-later reward of either 40, 80, or 120 $\mu$L after a 1, 2, 4, or 6 s delay. The task consisted of a series of 20 trial episodes. At the start of the episode, the larger-later reward value and delay, and lever (left or right) were randomly selected. For the first 10 trials of each game, rats were forced to press either the left or right lever to learn the value and delay associated with that lever (only one lever extended on each of these trials). For the last 10 trials, both levers extended and rats were free to choose. Rats were tested on two different versions of this task: one in which the post-reward delay was held constant, such that longer pre-reward delays reduced reward rate (const. delay), the other in which the time of the trial was held constant, such that longer pre-reward delays resulted in shorter post-reward delays to keep reward rate constant (const. trial). Long-run rate maximizers would be sensitive to the pre-reward delay in the constant delay condition only, but not the constant trial condition, as reward rate is held constant. Short-sighted animals would be sensitive to the pre-reward delays in both conditions.

Rats exhibited a change in preference over the course of the 10 free choice trials, indicating that they may still be learning reward values in early free choice trials. Accordingly, I focused analysis on the final 5 free choice trials (statistical analysis were robust to including all 10 free choice trials). In both conditions, rats were more likely to select the larger-later option when the larger-later reward was 120 $\mu$L vs. 80 $\mu$L vs. 40 $\mu$L (Figure 3.3B; effect of reward size across conditions: $\chi^2(1) = 12.953$, $p_{40-80} < .001$; $\chi^2(1) = 22.626$, $p_{40-120} < .001$; $\chi^2(1) = 18.650$, $p_{80-120} < .001$), and rats were less likely to select the larger-later reward with longer pre-reward delays.
(Figure 3B; effect of delay within conditions: $\chi^2_{\text{const delay}}(3) = 28.633, p_{\text{const delay}} < .001; \chi^2_{\text{const trial}}(3) = 12.946, p_{\text{const trial}} = .004$). However, rats were less sensitive to the pre-reward delay in the constant trial condition, in which it did not affect reward rate (effect of delay between conditions: $\chi^2(3) = 9.437, p = .024$). This effect was most evident at the longest pre-reward delay (Figure 3.3C; $\chi^2(1) = 6.453, p_{\text{const delay:6s vs. const time:6s}} = .044$).

3.2 Suboptimal decision-making in foraging and intertemporal choice tasks is best explained by quasi-hyperbolic discounting

Behavior in the foraging and two-alternative choice tasks indicates that rats consider future rewards, but exhibit a bias in processing time delays before obtaining rewards (while anticipating reward delivery) vs. after obtaining rewards (before making a subsequent decision). To test which specific hypotheses could account for this bias, I modeled both tasks as continuous
time semi-markov processes. The time between each event in the task was represented as a state (e.g. cues turning on/off, lever press, reward delivery; state space diagram of foraging task in Figure 3.4). The value of a state was defined as the discounted value of all future rewards. As the discount factor approached 1 (no temporal discounting), this model converged to long-run reward maximization, equivalent to MVT. Models representing each of the hypotheses described above were fit to behavioral data, and the predictions of each model were compared against rat behavior. Full details for all models can be found in the methods below.

Methods

Foraging Models

All models were constructed as continuous time semi-markov processes. This provided a convenient way to capture the dynamics of timing in both tasks, including slow delivery and consumption of reward (up to 6 s for the largest rewards), and test hypotheses regarding uncertainty in time intervals. To model the foraging task, each event within the task (e.g. cues turning on/off, lever press, reward delivery, etc.) marked a state transition (abbreviated state space diagram in Figure 3.4). All state transitions were deterministic, except for decisions to stay in vs. leave the patch, which occurred in decision states (the time between cues turning on at the start of the trial and rats performing a lever press or nosepoke). In decision states, a decision to stay in the patch transitioned to the handling time state, then reward state, ITI state, and to the decision state on the next trial. A decision to leave transitioned to the travel time state, then to the first decision state in the patch. Using the notation of Bradtke & Duff (1995), the reward for staying in state \( s \), \( Q(stay, s) \) is the reward provided for staying in state \( s \), \( R(stay, s) \), plus discounted value of the next state:

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

where \( \gamma(stay, s) \) is the discount factor applied to the value of the next state for staying in state \( s \), and \( V(s_{next}) \) is the value of the next state in the patch. For all non-decision states, rats did not have the option to leave the patch, so for these states, \( V(s) = Q(stay, s) \). For decision states, the value of the state was the greater of \( Q(stay, s) \) and \( Q(leave) \).
For simplicity, in most models I assume the time spent in a given state is constant. Under this assumption, \( R(\text{stay}, s) \) is the reward rate provided over the course of the state, \( r(s) \), multiplied by the time spent in that state \( T(s) \), discounted according to discount factor \( \beta \):

\[
R(\text{stay}, s) = \frac{1 - e^{-\beta \cdot T(s)}}{\beta} \cdot r(s),
\]

and discount factor for staying in state \( s \), \( \gamma(\text{stay}, s) = e^{-\beta \cdot T(s)} \). The value of leaving a patch, \( Q(\text{leave}) \), was equal to the discounted value of the first state in the next patch, \( V(s_{\text{first}}) \):

\[
Q(\text{leave}) = \gamma(\text{leave}) \cdot V(s_{\text{first}}),
\]

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

![State space diagram](image)

**Figure 3.4.** State space diagram for the semi-markov model of the foraging task. Decisions to stay vs. leave are made in Decision states. Decision to stay transitions to handling time, then reward, ITI, and to Decision state on the next trial. Reward is delivered throughout time spent in the each reward state. Reward depletion is achieved via shorter time spent in reward state (resulting in longer stay in the ITI state). Decision to leave transitions to the travel state, then to the first trial of the patch.
state in the patch in which $Q(stay, s) \leq Q(leave)$. To capture variability in the trial at which rats left patches, I added gaussian noise to $Q(leave)$. As decisions within each patch are not independent, the patch leaving threshold $Q(leave)$ did not vary trial-by-trial, but rather patch by patch, such that the cumulative probability that a rat has left the patch by state $s$, $\pi(leave, s)$, was the probability that $Q(stay, s) \leq Q(leave) + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$, with free parameter $\sigma$.

The optimal policy was found using dynamic programming. Optimal foraging behavior is to maximize undiscounted long-term reward rate (discount rate $\beta = 0$). As both $Q(stay)$ and $Q(leave)$ are infinite sums with $\beta = 0$, optimal behavior was determined by fixing $\beta = .001$ and assuming no decision noise. Optimal behavior was found for each rat, with the time spent in each state taken each individual rats’ data. For each model, I fit group level parameters and individual parameters for each rat using an expectation-maximization algorithm (Huys et al. 2011, MacKay 2003).

To model subjective costs, I assumed rats had an aversion to leaving a patch (Carter & Redish 2016, Wikenheiser et al. 2013), possibly due to the cost of increased cognitive effort required to make a decision to leave. I formalized subjective costs using a free parameter, $c$, representing a penalty applied to the patch leaving threshold:

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

For the diminishing marginal utility models, I assume rats maximized long-term subjective utility rather than reward. Subjective utility was defined by following the utility functions, with free parameter $\eta$:

$$Q_{utility}(stay, s) = U(stay, s) + \gamma(stay, s) * V_{utility}(s_{next}),$$

$$U_{power}(stay, s) = R(stay, s)^\eta, \text{ or}$$

$$U_{isoelastic}(stay, s) = \frac{R(stay, s)^{1-\eta} - 1}{1 - \eta}.$$
To test the effects of uncertainty in time intervals, the time spent in each state, \( T_{\text{uncertain}}(s) \), was normally distributed with variance proportional to the mean to account for the scalar property of animal time perception:

\[
T_{\text{uncertain}}(s) = T(s) + e_{\text{time}}(s), \text{ where } e_{\text{time}}(s) \sim \mathcal{N}(\mu = 0, \sigma = T(s) \cdot c_v),
\]

and \( c_v \) is the coefficient of variation in time perception. To determine the optimal behavior in the face of uncertainty, the reward function, \( R_{\text{uncertain}}(\text{stay}, s) \), was the average reward over distribution of time in the state, and discount functions, \( \gamma_{\text{uncertain}}(\text{stay}, s) \) and \( \gamma_{\text{uncertain}}(\text{leave}) \), the average discount over the distribution of time in the state. I approximated these three functions by taking the average of 10,000 samples from \( T_{\text{uncertain}}(s) \).

To determine behavior of a risk-averse agent practicing a maximin strategy, one that seeks to maximize the worst case scenario in the face of uncertainty, I assumed rats encode a distribution over the rewards and discount functions, and choose the best option assuming that rewards and discount functions fall within the bottom 10\% of their respective distributions.

To examine linear and sublinear underestimation of post-reward delays, respectively, the time spent in post-reward delay (ITI) states was transformed, with free parameter \( \alpha \):

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

\[
T_{\text{post-sublinear}}(s_{\text{ITI}}) = \frac{1 - e^{-\alpha T(s_{\text{ITI}})}}{\alpha}, \text{ where } \alpha > 0.
\]

Similarly, for overestimation of pre-reward delay, the handling time and travel time were transformed:

\[
T_{\text{pre-delay}}(s_{\text{HT}}) = \alpha T(s_{\text{HT}}), \text{ and }
\]

\[
\tau_{\text{pre-delay}} = \alpha \tau, \text{ where } \alpha > 1.
\]

For exponential discounting, \( \beta \) was fit as a free parameter.

Standard hyperbolic discounting was implemented using the \( \mu \)Agents model described by Kurth-Nelson & Redish (2009). The value functions of the overall model \( Q^\mu(\text{stay}, s) \) and \( Q^\mu(\text{leave}) \) were the average of 100 \( \mu \)Agents, each with their own exponential discount factor \( \beta_i \).
and thus individual reward functions \( R_i(stay, s) \) and discount functions \( \gamma_i(stay, s) \) and \( \gamma_i(leave) \):

\[
Q_i(stay, s) = R_i(stay, s) + \gamma_i(stay, s) \cdot V_i(s_{next})
\]

\[
Q^{\mu_{Agent}}(stay, s) = \frac{1}{100} \sum_i R_i(stay, s) + \gamma_i(stay, s) \cdot V_i(s_{next})
\]

\[
Q_i(leave) = \gamma_i(leave) \cdot V_i(s_{first})
\]

\[
Q^{\mu_{Agent}}(leave) = \frac{1}{100} \sum_i \gamma_i(leave) \cdot V_i(s_{first}).
\]

If the \( \mu_{Agent} \) discount factors, \( \beta_i \), were drawn from an exponential distribution with rate parameter \( \lambda > 0 \), the discounting function of the model approximated the standard hyperbolic discount function, \( \text{reward}/(1 + k \cdot \text{delay}) \), with discount rate \( k = 1/\lambda \). This relationship is presented in Figure 3.5. \( \lambda \) was fit as a free parameter.

**Figure 3.5.** Discount function of the \( \mu_{Agent} \) hyperbolic discounting model vs. standard hyperbolic discounting. Lines represent standard hyperbolic discounting function, \( 1 / (1 + k \cdot \text{time}) \). Ribbon represents the mean ± standard deviation of 100 simulations of the \( \mu_{Agent} \) model in which the discount factor for each of the 100 agents was sampled from an exponential distribution with rate parameter \( \lambda = 1/k \).
Quasi-hyperbolic discounting was originally formulated for discrete time applications (Laibson 1997). I used the continuous time formulation from McClure et al. (2007), in which the value functions of the overall model were the weighted sum of two exponential discount systems, a steep discounting $\beta$ system that prefers immediate rewards and a slower discounting $\delta$ system, each with their own reward functions $R_\beta(stay, s)$ and $R_\delta(stay, s)$ and discount functions $\gamma_\beta(stay, s)$, $\gamma_\beta(leave)$, $\gamma_\delta(stay, s)$, and $\gamma_\delta(leave)$:

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

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

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

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

The value function of the overall quasihyperbolic discounting model were:

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

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

where $0 < \omega < 1$ was the weight of the $\beta$ system relative to the $\delta$ system. I fit the parameters $\beta$, $\delta$, and $\omega$.

Two-Alternative Choice Models

Similar to the foraging task, events within the task marked state transitions, and all state transitions were deterministic except for decisions to choose the shorter-sooner option (SS) or longer-later option (LL) which occurred only in decision states (abbreviated state space diagram in Figure 3.6). From decision states, animals transitioned to delay, reward, and post-reward delay (ITI) states for the chosen option – the delay, reward and ITI for the SS and LL options were represented by separate states. The value of choosing SS or LL in decision state $s$ is the discounted value of the next state, the following delay state:

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

$$Q(LL, s) = \gamma(s) \cdot Q(LL \ Delay).$$
The value of delay states were the discounted value of the reward state for that action, the value of reward states were the reward for that action plus the discounted value of the ITI state for that action, and the value of ITI states were the discounted value of the decision state for the next action:

\[ Q(SS \text{ Delay}) = \gamma(SS \text{ Delay}) * Q(SS \text{ Reward}), \]
\[ Q(SS \text{ Reward}) = R(SS \text{ Reward}) * \gamma(SS \text{ Reward}) * Q(SS \text{ ITI}), \]
\[ Q(SS \text{ ITI}) = \gamma(SS \text{ ITI}) * V(s_{\text{next dec}}), \]

where the value of the next decision state, \( V(s_{\text{next dec}}) \) is the greater of the value of choosing SS or choosing LL in that decision state. Decisions were made using a softmax rule, with the probability of choosing the LL option in decision state \( s \) defined as:

\[ p(\text{choose LL}, s) = \frac{1}{1 + e^{Q(SS, s) - Q(LL, s)}/\sigma}, \]

with temperature parameter \( \sigma \), a free parameter that determines decision noise. The underestimating post-reward delays and temporal discounting models were implemented as they were in the foraging task.

Figure 3.6. Abbreviated state space diagram for the semi-markov model of the two-alternative choice task.
Model Comparison

All models had two parameters except for the quasihyperbolic discounting model, with four. To determine the model that fit provided the best fit to the data, while accounting for the increased flexibility of the quasihyperbolic discounting model, I calculated the integrated Bayes information criterion over the group level parameters (referred to as BIC for the rest of the chapter, Huys et al. 2011). BIC penalizes the log marginal likelihood, \( \log p(D \mid \theta) \), which is the log probability of all the data, \( D \), given group level parameters, \( \theta \), for model complexity, determined by the number of parameters \( k \), with the penalty being dependent on the total number of data points, \( n \):

\[
BIC = \log p(D \mid \theta) + \frac{k}{2} \log(n).
\]

I use a Laplace approximation to the log marginal likelihood.

\[
\log p(D \mid \theta) = -\frac{n}{2} \log(2\pi) * s + \sum_{i=1}^{s} p(D_i \mid \theta_i) p(\theta_i \mid \theta) - \frac{\sum_{i=1}^{s} \log\det(H_f(\theta_i))}{2}
\]

where \( s \) is the number of subjects, and \( H_f(\theta_i) \) is the Hessian matrix of the likelihood for subject \( i \) at the individual parameters \( \theta_i \).

Results

First, I confirmed that subjective costs and diminishing marginal utility could explain overharvesting, but not time preferences observed in the foraging task. Subjective costs were formalized as an aversion to leaving the patch (Carter & Redish 2016, Wikenheiser et al. 2013) – a constant cost term reduced the value of leaving. This model captured effects of overharvesting in the travel time and depletion rate experiments. As the cost was constant across environments, it failed to predict behavior in the scale experiment, in which rats overharvested to a greater degree when given larger rewards with longer delays. The cost model was also insensitive to the placement of delays, thus it failed to predict earlier patch leaving due to the pre-reward delay in the handling time experiment (Figure 3.7A).
To model diminishing marginal utility, the utility of a reward received in the task followed a sublinear function. I tested two different utility functions, a power law function and a steeper

![Graphs showing predictions of the best fit subjective cost and diminishing marginal utility models.](image)

**Figure 3.7.** A) Predictions of the best fit subjective cost and diminishing marginal utility models (power law = util-pwr; isoelastic utility = util-iso). B) BIC for the subjective cost and diminishing marginal utility models.
hyperbolic utility function, which becomes increasingly risk averse with larger rewards. Both utility models captured overharvesting in the travel time, depletion rate, and post-reward delay experiments. Furthermore, they predicted later patch leaving due to larger rewards with longer delays in the scale experiment, as larger rewards were not viewed as proportionally more valuable, but they failed to capture the magnitude of this effect. As the utility models were insensitive to the placement of delays, neither model predicted earlier patch leaving due to the pre-reward delay (Figure 3.7A).

I next considered whether deviations from MVT could be explained by timing uncertainty. If timing uncertainty causes animals to overharvest, then I would expect animals to overharvest more when uncertainty is greater (e.g. with longer durations in the scale experiment). Furthermore, if delays are broken into pre- vs. post-reward intervals, animals may be able to better estimate two shorter delays than one longer post-reward interval, so that animals may overharvest less when the intervals between events are shorter. However, it is unknown whether overharvesting is an optimal response to timing uncertainty. Uncertainty in time intervals can be explicitly modeled using semi-markov dynamics, where the time spent in a given state is drawn from a distribution with variance proportional to the mean time spent in the state (Bradtke & Duff 1995, Daw et al. 2006). In a simple environment with a single patch type, I tested whether increasing the coefficient of variation in timing altered the optimal trial at which an agent should leave a patch across various levels of temporal discounting. Increasing the coefficient of variation had no effect on the optimal time to stay in a patch (Figure 3.8). I also considered that rats may not try to optimize decision-making in the face of uncertainty, but instead adopt a robust strategy, seeking to obtain a minimum acceptable amount of reward. I tested whether a maximin strategy, that rats would choose the action that maximizes the worst-case scenario for each state in the face of uncertainty (Balci et al. 2011, Brunner et al. 1996, Zacksenhouse et al. 2010) would cause overharvesting. Using the maximin strategy, timing uncertainty had no effect at low discount factors, but with greater levels of discounting and large coefficient of variation, it predicted that rats would leave patches one trial later (Figure 3.8). Based on these results, it is possible that timing uncertainty plays a small role in overharvesting, but it alone cannot account for the degree of overharvesting exhibited by rats.
Next, two classes of models were tested that assume biased perception of reward value as a function of time: ignoring (or underestimating) post-reward delays and temporal discounting. Three versions of the idea that animals are insensitive to post-reward delays were tested: i) that rats are perceive post-reward delays as shorter than they actually are, in a linear fashion, ii) that rats perceive post-reward delays in a sublinear fashion, and a complementary model: iii) that rats perceive post-reward delays accurately, but overestimate pre-reward delays. The last model represented an aversion to time spent in anticipation of receiving reward. Sublinear underestimation of post-reward delays provided the best fit to the first four experiments, predicting rat behavior within 1 standard deviation in each experiment (Figure 3.9A-D). However, this model largely ignored longer post-reward delays, and could not account for sensitivity to post-reward delays exhibited by rats (Figure 3.9E). The linear underestimation of post-reward delays and overestimation of pre-reward delays predicted the direction of all effects: general overharvesting, staying longer in patches due to larger rewards with longer delays, earlier patch leaving due to a pre-reward delay, and sensitivity to post-reward delays, but failed to predict the magnitude of the changes in the scale and handling time experiments (Figure 3.9C-D).

Figure 3.8. Left) Optimal behavior for the scalar timing uncertainty model at different discount rates. Right) Behavior of an agent using a maximin strategy in the face of scalar timing uncertainty at different discount rates. cv = coefficient of variation.
The last models tested were three temporal discounting models: an exponential, a hyperbolic, and a quasihyperbolic discounting model. To implement the standard form of hyperbolic discounting in the semi-markov model, I used an approach previously described by Kurth-Nelson & Redish (2009). This model contained a number of “Agents,” each with their own exponential discount factor and value function. The overall value function of the “macroAgent” was the average of the Agents. The rate of hyperbolic discounting exhibited by the macroAgent is determined by the distribution of exponential discount factors of the Agents. The quasihyperbolic discounting model was a more flexible form of hyperbolic discounting, allowing for steeper discounting of rewards obtained in the near future relative to rewards obtained further in the future. This model consisted of a weighted-sum of two exponentially discounted values (McClure et al. 2007). All three discounting models explained overharvesting in the travel time and

Figure 3.9. Predictions of the best fit models of overestimation of pre-reward delays (pre-delay), linear underestimation of post-reward delays (post-delay), and sublinear underestimation of post-reward delays (post-delay-sublinear). Points and errorbars are the mean and standard deviation of rat behavior, colored lines represent the mean predicted number of harvests across all rats.

The last models tested were three temporal discounting models: an exponential, a hyperbolic, and a quasihyperbolic discounting model. To implement the standard form of hyperbolic discounting in the semi-markov model, I used an approach previously described by Kurth-Nelson & Redish (2009). This model contained a number of “Agents,” each with their own exponential discount factor and value function. The overall value function of the “macroAgent” was the average of the Agents. The rate of hyperbolic discounting exhibited by the macroAgent is determined by the distribution of exponential discount factors of the Agents. The quasihyperbolic discounting model was a more flexible form of hyperbolic discounting, allowing for steeper discounting of rewards obtained in the near future relative to rewards obtained further in the future. This model consisted of a weighted-sum of two exponentially discounted values (McClure et al. 2007). All three discounting models explained overharvesting in the travel time and
depletion rate experiments, and that rats would stay longer in patches yielding larger rewards with longer delays, as the delay to the next patch was longer (Figure 3.10A-C). Surprisingly, the exponential and hyperbolic discounting models did not predict that rats would leave earlier due to a pre-reward delay. For both of these models, the discounting function was not steep enough to explain how a 3 s pre-reward delay would influence foraging behavior to this degree. A steeper discounting function may be able to explain the influence of the 3 s pre-reward delay, but steeper discounting would also predict greater levels of overharvesting. As the quasihyperbolic discount function enables steep discounting of rewards received in the near future with shallow discounting of rewards obtained further in the future, this model was able to explain the degree of overharvesting and rats’ bias to leave patches earlier due to the 3 s pre-reward delay. All three models predicted rats would be sensitive to the post-reward delay (Figure 3.10E).

Figure 3.10. Predictions of the best fit exponential discounting model (disc-exp), hyperbolic discounting model (disc-hyp), and quasi hyperbolic discounting model (disc-quasi). Points and errorbars are the mean and standard deviation of rat behavior, colored lines represent the mean predicted number of harvests across all rats.
The quasihyperbolic discounting model was the only model tested that could capture rat behavior within one standard deviation across all foraging tasks. I performed quantitative model comparison, examining the BIC of all models in each of the foraging tasks (Figure 3.11, 3.7B). The quasihyperbolic discounting model scored the lowest in BIC in all tasks except for the depletion rate experiment, in which it ranked second to the hyperbolic discounting model with a difference of 6. I next determined whether the same quasihyperbolic discounting model could explain behavior in the two-alternative choice task.

For the two-alternative choice task, I modeled each series of 10 free choices as separate episodes, such that the value of the first choice was equal to discounted reward across all 10 episodes.

**Figure 3.11.** BIC for each of the underestimating post-reward delays models and temporal discounting models for each foraging experiment and two-alternative choice experiment. Pre-del = linear overestimation of pre-reward delays, post-del = linear underestimation of post-reward delays, post-del-sub = sublinear underestimation of post-reward delays, disc-exp = exponential discounting, disc-hyp = hyperbolic discounting, disc-quasi = quasi-hyperbolic discounting.
choices, the value of the second choice equal to discounted reward across the remaining 9 choices, and so on. The episode ended at the 10th choice, and did not consider rewards on future games. I tested the underestimating post-reward delays models and temporal discounting models in this task, finding that the quasihyperbolic discounting model had the lowest BIC (Figure 3.11, 3.12). To further test whether quasi-hyperbolic discounting may represent a common mechanism for suboptimal decision-making across tasks, I examined whether the best-fit discounting function to the two-alternative choice task could explain behavior in the foraging task, and whether the best fit discounting function from the foraging task could explain behavior in the two-alternative choice task. As the same rats performed both the post-reward delay foraging task and the intertemporal choice task, I focused on this foraging experiment. The best fit discounting functions for each individual rat from the intertemporal choice task predicted that rats would overharvest and leave patches earlier due to the longer pre-reward delay, but it predicted that rats would leave patches somewhat earlier than they did (0.89 trials earlier in the short delay condition and 0.27 trials earlier in the long delay condition; Figure 3.13A). The best fit discounting function from the foraging task provided a good fit to the two-alternative choice data, predicting that rats would be less likely to select the larger-later option due to the pre-reward delay, within the range of rats performance (Figure 3.13B). The ability of to predict two-alternative choice data from

![Figure 3.12. Predictions of the quasi-hyperbolic discounting model. Points and error bars represent mean and standard error of rat behavior, lines represent mean model prediction.](image-url)

39
foraging behavior suggests that rats approach these tasks in a similar manner, and that quasi-hyperbolic discounting may represent a common mechanism for suboptimal behavior in these two paradigms.

**Figure 3.13.** A) Predictions for the best fit parameters from the two-alternative choice task, applied to the foraging task. B) Predictions for the best fit parameters from the post-reward delays foraging experiment, applied to the two-alternative choice task.

### Discussion

In the present study, rat foraging behavior was consistent with a number of previous studies (Constantino & Daw 2015, Hayden et al. 2011): rats followed basic predictions of MVT, staying longer in patches that yielded greater rewards, staying longer in all patch types when the cost of traveling to a new patch was greater, and leaving patches earlier when rewards depleted quicker. Furthermore, rats overharvested, staying longer in patches than is predicted by MVT. These findings were extended to show that in certain environments, rats violated qualitative predictions of MVT, overharvesting to a greater degree when given larger rewards with longer delays despite equivalent reward rates, and rats were sensitive to the placement of delays (before vs. after receiving reward). These data indicate that overharvesting may be a by product
of time preferences. When I directly tested rats sensitivity to post-reward delays, I found that contrary to a variety of studies on intertemporal choice (Bateson & Kacelnik 1996, Blanchard et al. 2013, Stephens & Anderson 2001), rats were sensitive to post-reward delays in both foraging paradigm and two-alternative choice paradigms.

Whether animals maximize long-term rewards, the reward rate across all patches or all future decisions, has long been under debate (Bateson & Kacelnik 1996, Denny 2017, Templeton & Lawlor 1981, Turelli et al. 1982). Impulsivity observed in two-alternative choice tasks suggests that animals may instead maximize local reward rate, reward rate in the current patch or on the current decision, while failing to fully consider opportunity costs of that decision (such as a long post-reward delay; Bateson & Kacelnik 1996, Gallistel & Gibbon 2000, Kacelnik 1997). This rule is similar to standard delay discounting models (Bateson & Kacelnik 1996, Kacelnik 1997).

Beyond two-alternative choice tasks, local reward maximization has been applied recently to explain behavior of monkeys in a patch leaving task – in an environment with variable travel time, monkeys stayed longer in patches when the impending travel time to the next patch was longer, rather than adopt a single leaving threshold across all patches (Hayden et al. 2011). However, rat behavior in the present study provides multiple lines of evidence for long-term rate maximization: i) in an environment with multiple patch types (travel time experiment), rats adopt a single reward rate threshold at which all patches are left, staying longer in patches that yield greater rewards, and ii) rats are sensitive to post-reward delays in both foraging and intertemporal choice context, confirming that they consider future rewards. Although rats did not behave optimally in either paradigm, observed time preferences were more consistent with temporal discounting than local-reward maximization.

The idea that animals exhibit similar decision biases in foraging and two-alternative choice paradigms, and that these biases can be explained by a common mechanism is in conflict with prior studies that found that animals are better rate maximizers in foraging tasks and that delay discounting models of two-alternative choice behavior were poor predictors of foraging behavior (Blanchard & Hayden 2015, Carter & Redish 2016, Carter et al. 2015, Stephens 2008). It has been argued that animals may perform better in foraging tasks because decision-making systems have evolved to solve foraging problems rather than two-alternative choice problems (Blanchard & Hayden 2015, Stephens 2008, Stephens et al. 2004). Results from the present
study suggest that animals use similar rate maximizing strategies across paradigms. One potential explanation for why delay discounting models fail to predict foraging behavior is that typically, delay discounting models only consider the decision at hand (local reward maximization). This is reasonable, as in most two-alternative choice tasks, decisions are independent of one another, such that future expected rewards are equivalent regardless of the current decision. However, in patch leaving tasks, decisions are not independent – a decision to stay in the patch right now affects the potential reward on the next decision. Optimal foraging models, such as MVT, assume that animals consider future rewards, and behavior of our rats confirms this prediction.

Maximization of discounted future rewards best explained rat behavior across foraging and two-alternative choice paradigms, but many of the models tested were capable of explaining some of the biases exhibited by rats. I cannot exclude the possibility of subjective costs, diminishing marginal utility, or biased estimation of time intervals from contributing to suboptimal decision-making. Importantly, however, our data indicate that maximization of discounted future rewards may provide a link between foraging and two-alternative choice paradigms, and it highlights the importance of future work considering the source of time preferences. Although quasi-hyperbolic discounting is necessary to explain to foraging behavior across a range of contexts, within a single task, discounting and simpler models based on the MVT, such as the subjective costs model, make similar predictions and have been more widely used in previous studies. For simplicity and consistency with previous studies, the MVT will remain the focus of analysis in the next two chapters.
Chapter 4: The Locus Coeruleus-Norepinephrine System Regulates Task Engagement

Introduction

Locus coeruleus-norepinephrine (LC-NE) neurons exhibit two distinct modes of activity: phasic and tonic (Aston-Jones & Bloom 1981, Aston-Jones & Cohen 2005). In the phasic mode, LC-NE neurons exhibit moderate baseline activity, with short bursts of higher frequency firing (10-15 Hz) associated with salient stimuli or decisions. These burst responses are thought to optimize performance within a cognitive task, promoting appropriate actions that lead to reward (e.g., lever presses) (Bouret & Sara 2004, Clayton et al. 2004, Shea-Brown et al. 2008, Usher et al. 1999). In contrast, in the tonic mode LC-NE neurons exhibit elevated and more irregular baseline activity (2-6 Hz) with few evoked (burst) responses (Aston-Jones & Cohen 2005, Aston-Jones et al. 1994). Periods of high tonic activity are associated with poor performance and increased distractibility within a given task, as animals participate in fewer trials and commit more errors (Aston-Jones et al. 1994).

Although increased tonic activity can be detrimental to performance within a task, Adaptive Gain Theory (AGT) proposes that increased tonic activity may promote behavioral flexibility when utility in a given task is low (Aston-Jones & Cohen 2005). LC-NE tonic activity increases when reward contingencies change (Aston-Jones et al. 1997), and increasing NE via pharmacological manipulations improves behavioral flexibility in rats (McGaughey et al. 2008). Furthermore, baseline pupil diameter, an indirect measure of LC-NE tonic activity (Aston-Jones & Cohen 2005, Joshi et al. 2016, Varazzani et al. 2015), increases as reward rate within a task wanes (Gilzenrat et al. 2010). However, a causal relationship between LC tonic activity and propensity to change behavior has yet to be established.

To test for a causal relationship, I selectively stimulated LC tonic activity using designer receptors (DREADDs; Alexander et al. 2009, Armbruster et al. 2007) as rats performed a patch foraging task, predicting that increasing LC tonic activity would promote disengagement from current behavior (i.e., exploiting a particular patch). Specifically, I predicted that disengagement from exploiting a patch would result in earlier patch leaving, and that rats’ decisions to stay in the patch vs. leave the patch would be less correlated with reward rate (i.e. would reflect an increase
in decision noise rather than a systematic bias to leave earlier). Our experimental results showed that DREADD stimulation of LC-NE neurons, which increases LC tonic firing (Vazey & Aston-Jones 2014), caused rats to leave patches earlier and increased decision noise, dissociating rats' decisions from the reward rate within the patch. Additionally, DREADD stimulation impaired performance in the foraging task, reducing task participation, increasing omission rates, and – after prolonged stimulation – inducing long bouts of immobility that resembled previously reported behavioral arrest (Carter et al. 2010). Aside from bouts of immobility observed following prolonged stimulation of LC-NE neurons, these findings are consistent with the hypothesis, derived from AGT, that increases in LC tonic firing rate favor task disengagement and the pursuit of alternative behaviors by global modulation of gain that increases processing noise.

Methods

Animals

Adult Long-Evans rats (Charles River, Kingston, NY; n = 32) were used. Rats were housed on a reverse 12 h/12 h light/dark cycle (lights off at 7 A.M.). All behavioral testing was conducted during the dark period. Throughout behavioral testing, rats were food restricted to maintain a weight of 85-90% ad-lib feeding weight, and were given ad-lib access to water. All procedures were approved by the Princeton University Institutional Animal Care and Use Committee.

Operant Training

Rats were initially trained to press a lever for 10% sucrose water reward on an FR1 paradigm, in which each lever press was rewarded with 100 µL of sucrose water. Next, rats were trained to nose poke when reward was unavailable. In this stage, the lever stopped yielding reward on randomly selected intervals between 4-12 lever presses. To regain access to a rewarding lever, rats were required to nose poke at the back of the chamber, which caused the lever to retract, and another lever on the opposite side of the chamber to extend, which yielded reward. Next, a delay (first 5 s, then 10 s) that simulated the time to travel between patches was introduced between the nose poke and the opposite lever extending. Finally, rats were trained
and tested on the foraging task (described below). To move on to the next stage of training, rats were required to pass a criterion of at least 100 rewarded lever presses in a 1 h session.

**Foraging task**

Rats performed the foraging task as in Chapter 2. In this task, rats were tested on a wide range of patch types within each behavioral session, starting with 30-150 μL of reward and depleting at the same rate. Only one travel time was tested: 10 s. Rats were trained until reaching a criterion of stable behavior, then tested for 3-5 sessions for each drug treatment, with at least one day with no drug treatment in between each testing day.

**Foraging Video Analysis**

A single behavioral session for each condition (saline, 0.3 and 1 mg/kg CNO) of hM3Dq-expressing animals was recorded and analyzed. The frequency and duration of periods of immobility (no detectable movement for at least 10 s) were scored.

**Model Fitting**

To determine how well reward rate predicted decisions to leave patches, I used logistic decision models to fit choices as a function of the difference between the cumulative reward rate across patches, $R$, and current reward rate within the patch, $r_i$. I refer to this difference as the value of leaving the patch, $V_{leave}$:

$$r_i = \frac{reward_i}{time_i},$$

$$R = \frac{\sum_i reward_i}{\sum_i time_i},$$

$$V_{leave} = R - r_i.$$ 

The cumulative reward rate, $R$, was measured as the average reward rate across the entire
session. As rats were trained for many days and maintained stable behavior, I assume they had an estimate of the average reward rate in the environment from the beginning of the session.

Logistic decision curves were fit as generalized linear mixed effects models, with a fixed effect of group (mCherry vs. hM3Dq), random effects for drug treatment, \( V_{\text{leave}} \), and interactions among all variables as fixed and random effects. From this model, I am interested in two parameters: the indifference point and the slope. The indifference point, the point at which rats were equally likely to stay in the patch vs. leave, represents the bias towards staying in vs. leaving the patch. If the indifference point occurs at a positive value of leaving (i.e. reward rate within the patch is greater than the cumulative reward rate), it indicates overharvesting (i.e., staying too long in patches). The indifference point is measured from the intercept of the model, and change in the intercept due to group and/or drug treatment is given by the fixed effect of group, treatment, and their interaction. The slope provides a measure of how well rats’ decisions correlate with the value of leaving the patch, and serves as an index of decision noise; a smaller slope (more shallow curve) indicates that the value of leaving the patch is less predictive of rats’ decisions, and thus presumably more subject to noise. This parameter is given by the effect of \( V_{\text{leave}} \), and changes in the slope due to group and/or drug treatment are given by the interactions among the \( V_{\text{leave}} \), group, and drug treatment.

Two potential concerns with this analysis are that i) parameters may change over the course of the session; and ii) the model assumes all foraging decisions are independent and identically distributed, whereas there may have been dependency on decisions within the same patch. To address these concerns, I fit an additional GLM, adding a fixed and random effect of time into the session, and random intercept for each patch visited. The slope and indifference point for each group and drug treatment were estimated as described above.

**Virus injection surgery**

Anesthesia was induced and maintained using isoflurane (5% and 2% respectively). Rats were aligned in a stereotaxic frame such that bregma was 2 mm below lambda (~15° head angle). Rats received bilateral injections of an AAV-PRSx8-hM3Dq-HA or AAV-PRSx8-mCherry virus into LC at the following coordinates from lambda: AP: -4.0 mm and ML: ± 1.35 mm from
lambda, DV: 6-6.5 from skull surface. Total injection volume was 1.2 µL per hemisphere, at a rate of 100 nL/min (Vazey & Aston-Jones 2014). After surgery, rats received meloxicam (0.1 mg/kg i.p.). for post-operative analgesia daily for 3 days, and were given 7 days to recover before further testing.

**CNO administration**

DREADDs were activated by the selective ligand clozapine-N-oxide (CNO; NIMH Chemical Synthesis and Drug Supply Program), which was dissolved in 2% DMSO and diluted in sterile saline to a final injection volume of 1 mL/kg. CNO (0.3 mg/kg or 1 mg/kg, i.p.) or saline was administered immediately prior to the start of behavioral testing, and CNO (1 mg/kg, i.p.) was administered 2 hours before perfusion for Fos immunohistochemistry.

**Histology**

After a final CNO injection, animals were transcardially perfused with PBS then 4% paraformaldehyde. Brains were extracted, post-fixed overnight in 4% paraformaldehyde, and cryoprotected in PBS with 30% sucrose for 72 hr. Brains were then embedded in OCT compound and frozen using isopentane, then stored at -80° C. Coronal sections (35 µm) throughout LC were cut on a cryostat. To confirm expression of hM3Dq and mCherry, a 1-in-4 series of 35 um-thick sections was labeled with antibodies against tyrosine hydroxylase (TH), as a marker of LC-NE neurons, and the HA DREADD-tag or mCherry. Sections were first blocked with 5% normal donkey serum (NDS) in PBS with .25% Triton X-100 for 1 hr, then incubated overnight in primary antibodies: mouse anti-TH (1:1000; Immunostar, cat no.: 22941), rabbit anti-HA (1:1000, Cell Signaling Technologies, cat no.:3724S), or rabbit anti-DsRed (1:1000, Clontech: cat no.: 632496). Sections were then washed 3 x 5 min in PBS, and incubated for 2 hr in secondary antibodies (conjugated to Alexa Fluor dyes): donkey anti-mouse 488 (1:250, Invitrogen, cat no.: A21202) and donkey anti-rabbit 568 (1:250, Invitrogen, cat no.: A10042). Finally, sections were washed again for 3 x 5 min in PBS.

To examine expression of the immediate early gene product Fos in LC-NE neurons following CNO administration, an additional series of sections was labeled with antibodies against TH and Fos. Sections were first blocked in 10% NDS in PBS with .25% Triton X-100, then
incubated overnight with primary antibodies: mouse anti-TH (1:500) and goat anti-Fos (1:100, Santa Cruz, cat no.: SC-52-G). Sections were then washed for 3 x 5 min in PBS and incubated for 2 hr in secondary antibodies: donkey anti-mouse biotin (1:100, JAX, cat no.: 715-065-150) and rabbit anti-goat 488 (1:100, Invitrogen, cat no.: A11078). Sections were washed again for 3 x 5 min in PBS and then incubated for 2 hr in streptavidin 350 (1:100, Invitrogen, cat no: S11249) and donkey anti-rabbit 488 (1:100, Invitrogen, cat no.: A21206).

**Histological Analysis**

To measure the density of hM3Dq-HA and mCherry staining, two images of TH and hM3Dq-HA or mCherry staining, ~mid-LC along the anterior-posterior axis, were acquired using a Zeiss LSM 700 confocal microscope, and analyzed using Fiji (Schindelin et al. 2012). In each image, TH and transgene staining were binarized using the same threshold for each image, and I measured the percentage of pixels positively labeled with transgene within the pixels positively labeled with TH. The density of Fos+ cells in the LC was determined using the optical fractionator method in StereoInvestigator (MBF Bioscience, Williston, VT) with an Olympus BX-60 microscope. In three sections for each animal, a contour was drawn around the boundary of LC in each hemisphere. Fos+ cells were counted using a 40X objective. The size of the grid (70 x 70 µm) and counting frame (110 x 110 µm) were set to count ~ 40% of the volume of each contour. Due to the small size of the LC, differences in the volume of LC in counted sections had a great influence on the stereological estimate of Fos+ cells. Therefore, I normalized the number of Fos+ cells in counted sections to the volume of counted sections (density = Fos+ count/volume).

**Data Analysis**

All analyses were performed in R (R Core Team 2015). To examine changes in the number of trials spent in patches, a two-way ANOVA with patch type and travel time as random effects was used for the initial behavioral experiment. A t-test was used to compare rat behavior with optimal behavior. To examine differences in Fos expression due to LC stimulation, a t-test on the density of cells expressing Fos between mCherry and hM3Dq animals was used. Additionally, I regressed the density of Fos+ cells by the level of mCherry or hM3Dq expression to examine if more DREADD expression leads to greater activation of LC. For the DREADD stimulation
foraging experiment, a mixed effects model with fixed effects for animal group (mCherry vs. hM3Dq), patch type, and drug treatment, and random effects for patch type and drug treatment, was used to examine the effects of LC stimulation on foraging behavior. To examine changes in logistic decision model parameters, I used two-way ANOVAs, with animal group as a fixed effect and drug treatment as a random effect. Mixed effects models were conducted using the lme4 package (Bates et al. 2015). To examine simple and interaction effects, I tested contrasts using the phia and lsmeans packages in R, with Bonferroni correction for multiple comparison in all post-hoc contrasts (de Rosario-Martinez 2015).

4.1 Stimulation of LC-NE neurons impairs task performance

DREADD stimulation of LC-NE neurons increases Fos expression

To stimulate LC tonic firing, I expressed the Gq-coupled DREADD hM3Dq or the fluorescent protein mCherry to control for potential effects of virus expression, selectively in LC-NE neurons. Expression was restricted to LC-NE neurons by including the synthetic dopamine-β-hydroxylase (DBH) promoter, PRSx8 (Hwang et al. 2001, 2005), in an AAV that encoded either hM3Dq (hM3Dq-HA; n = 15) or mCherry (control, n = 9). Prior studies using these vectors found that 97% of cells expressing hM3Dq-HA or mCherry co-localize with TH following injection of these AAVs into LC (Vazey & Aston-Jones 2014). Figure 4.1 shows expression of hM3Dq-HA and mCherry in LC. Because hM3Dq-HA is heavily trafficked to the plasma membrane and dendrites (Vazey & Aston-Jones 2014), counting the number of LC-NE cell bodies expressing hM3Dq-HA would not provide the most accurate measure of expression levels. Instead, I examined hM3Dq and mCherry expression by quantifying the density of hM3Dq-HA and mCherry immunofluorescence (Figure 4.1A and B), and confirmed activation of LC-NE neurons by measuring Fos expression 2 h after CNO injection (1 mg/kg). Results showed that 13/15 hM3Dq-HA rats had bilateral DREADD expression in LC, and 2/15 unilateral; 9/9 mCherry rats had bilateral expression. I counted the density of Fos+ cells in LC in a subset of these rats (10 hM3Dq-HA rats – 8 bilateral/2 unilateral – and 6 mCherry rats). All 10 hM3Dq-HA expressing animals exhibited greater Fos expression in LC than the 6 mCherry animals, t(14) = 4.83, p < .001, confirming our ability to stimulate LC-NE neurons. Additionally, Fos expression increased
with greater levels of hM3Dq-HA staining, but not with greater levels of mCherry staining (hM3Dq: \( \beta = 6734, p = .004 \); mCherry: \( \beta = -16.15, p = .932 \); Figure 4.1C).

**Figure 4.1.** A) Images of TH (top), hM3Dq-HA and Fos (middle), and merged images, showing expression of hM3Dq-HA and Fos in LC. B) Images showing expression of mCherry and Fos in LC. C) The density of Fos+ cells in the LC as a function of the level of expression of hM3Dq-HA or mCherry. Because hM3Dq-HA and mCherry exhibited different expression patterns, the density of expression was normalized (z-scored) for comparison. The density of Fos+ cells in LC increased with more expression of hM3Dq-HA, but not with mCherry.
DREADD stimulation impairs performance in the foraging task

hM3Dq-HA and mCherry rats were tested in the foraging task with 9 different patch types, which started with a range of 30-150 μL of reward, and a 10 s travel time. In typical foraging tasks, reward rate within the patch starts above rats’ threshold for leaving patches (the cumulative reward rate per MVT) and depletes to this threshold, such that agents experience few trials in which the reward rate within the patch is below their threshold for leaving patches. Using this design – testing rats on a wider variety of patch types, including patches that start with very little reward (30-45 μL) – rats encountered some patches that started with reward rate below their typical threshold for leaving patches. In this version of the task, rats stayed significantly longer in patches that started with larger rewards, F(8, 184) = 447.6, p > .001.

Rats were injected with either saline or CNO (0.3 or 1 mg/kg), and tested in the foraging task immediately following the injection. I first assessed general performance, including the number of trials completed per 1 h session, response times, and the rate of omissions (defined as trials in which the lever was pressed to harvest from the patch without then entering the reward magazine to obtain the reward). Behavioral data were binned into 10 min intervals to look at the time course of these behaviors within sessions. Following saline injections, hM3Dq-HA rats participated in an average of 361 ± 31 trials, and mCherry rats in 358 ± 31 trials. The number of completed trials declined in the last 20 min of the session for both LC-hM3Dq-HA and mCherry rats (mean trials in first 10 min = 61.93, SD = 4.32; mean trials in last 10 min = 52.33, SD = 9.78; main effect of time on trials: t(396) = 3.497, p < .001), and RTs slightly increased in the last 20 min, but the main effect of time on RT was not significant (mean RT in first 10 min = 2.06, SD = .70 s; mean RT in last 10 min = 3.34, SD = 1.79 s; main effect of time on RT: t(394) = .386).

Omission rates remained constant throughout the session, at an average of 0.1% omissions for both hM3Dq-HA and mCherry rats (main effect of time on omissions: t(396) = .01, p = .992).

Consistent with prior studies in which high tonic LC activity correlated with poor performance (i.e., more errors and slower RTs; Aston-Jones et al., 1994), performance of hM3Dq-HA rats sharply declined within 30 min following CNO injection. hM3Dq-HA, but not mCherry rats, exhibited a significant reduction in trials (effect of treatment within group: hM3Dq– p < .001; mCherry– p = .821; effect of treatment between groups: p < .001; Figure 4.2A), a significant increase in RT (effect of treatment within group: hM3Dq– p < .001, mCherry– p = 1; effect of
treatment between groups: $p < .001$; Figure 4.2B), and increased omission rates (effect of treatment within group: hM3Dq– $p < .001$, mCherry– $p = 1$; effect of treatment between groups: $p$

Figure 4.2. CNO administration caused impairment in performance in the foraging task in hM3Dq-HA, but not mCherry animals, consisting of a reduction in the number of trials that rats participated in (A), an increase in response time (B), and increase in the rate of omissions, i.e. trials in which rats pressed the lever but did not retrieve reward (C). Testing began immediately following drug treatment, so time along the x-axis represents both time within the session and time following drug administration. Values represented as mean ± SEM; $n = 15$ hM3Dq, 9 mCherry.
Additionally, in some animals, prolonged stimulation of LC-NE neurons induced long bouts of complete disengagement and immobility (lasting 10 s-5 min) that resembled behavioral arrest (Carter et al. 2010). Bouts of immobility were first observed ~20 min following CNO administration, and occurred more frequently and for longer duration as time since CNO administration elapsed. I analyzed videos of hM3Dq-HA rats performing the task following saline, 0.3, and 1 mg/kg CNO injections to examine the times at which rats were immobile. hM3Dq-HA rats exhibited bouts of immobility on a higher percentage of trials following both 0.3 and 1 mg/kg CNO than following saline injections (saline: M = 0.06%, SD = 0.12% of trials; CNO 0.3 mg/kg: M = 2.53%, SD = 3.11%; CNO 1 mg/kg: M = 3.06%, SD = 3.88%; saline vs. .3 mg/kg: p < .001, saline vs. 1 mg/kg: p = .031, .3 mg/kg vs. 1 mg/kg: p = .679). I next examined the total duration of bouts of immobility in the first 20 min, middle 20 min, and final 20 min of behavioral sessions. Following CNO injections, rats spent more time immobile in the middle and final 20 min of the session than in the first 20 min (effect of time on immobility following saline: p = 1.00; .3 mg/kg CNO: p = .016; 1 mg/kg CNO: p < .001).

4.2 Stimulation of LC-NE neurons causes disengagement from important task variables

**DREADD stimulation causes earlier patch leaving**

Because task participation waned and omission rates and bouts of immobility increased over the course of the session, I restricted analyses of rats’ patch-leaving behavior to the first 30 minutes of testing. As rats general behavior changed during this period, I included time into the session as a predicting variable in the mixed effects model used to examine the number of trials spent in each patch across animal group (mCherry vs. hM3Dq) and drug treatment. Overall, rats spent more trials in patches that contained more reward (main effect of patch type: t(31) = 25.598, p < .001). As predicted by AGT, hM3Dq-HA rats left patches earlier following both .3 mg/kg and 1 mg/kg CNO when compared to saline administration (saline vs. .3 mg/kg: p = .003; saline vs. 1 mg/kg: p < .001, .3 mg/kg vs. 1 mg/kg: p = .460; Figure 4.3). Neither dose of CNO affected patch leaving behavior in mCherry rats (saline vs. .3 mg/kg: p = .416; saline vs. 1 mg/kg:
Additionally, I directly tested the change in behavior due to drug administration between groups, finding that the change in behavior from saline to 1 mg/kg CNO was greater in hM3Dq rats than in mCherry rats (mCherry vs. hM3Dq comparisons: saline vs. 1 mg/kg – p = .045, saline vs. .3 mg/kg – p = .197, .3 mg/kg vs. 1 mg/kg – p = .369). This patch leaving behavior was not directly the result of bouts of immobility as these effects were robust to removing patches in which a bout of immobility occurred.

**DREADD stimulation increases decision noise**

To examine trial-by-trial choices, I computed the psychometric function for decisions to stay vs. leave the patch suggested by MVT: that is, the probability of leaving as a function of the difference between the cumulative reward rate across patches and current reward rate within the patch, which I refer to as the value of leaving the patch. Logistic decision curves were fit to the data as generalized linear mixed effects models (GLMMs), with a fixed effect for group (hM3Dq vs. mCherry), and random effects for drug treatment and the value of leaving the patch. From this model, I assessed i) whether rats on average leave patches earlier due to drug treatment, and ii) how closely rats’ decisions correlate with the value of leaving patches, which is given by the slope of the curve. A decrease in the slope of the curve (i.e. shallower curve) indicates that rats’ decisions are less closely correlated with the value of leaving the patch, and thus are more noisy. AGT predicts that stimulating LC tonic firing should favor disengagement from the task (and thereby earlier leaving) by increasing decision noise, as indexed by a decrease in slope (Figure 4.3).

To examine whether rats on average left patches earlier, I tested pairwise contrasts for each drug treatment (saline, .3 mg/kg CNO, and 1 mg/kg CNO) within group, as well as the change between treatments across groups. hM3Dq rats were more likely to leave patches on average due to both doses of CNO (saline vs. .3 mg/kg: p < .001; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .046). CNO had no effect on mCherry rats (saline vs. .3 mg/kg: p = .972; saline vs. 1 mg/kg: p = .801; .3 mg/kg vs. 1 mg/kg: p = .891), and the effects of CNO were significant between hM3Dq and mCherry rats (saline vs. .3 mg/kg: p = .002; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .080). To examine changes in slope, I tested the same contrasts as described above, but focusing on their interaction with the value of leaving the patch. Both
doses of CNO significantly decreased the slope (i.e. increased decision noise) in hM3Dq rats (saline vs. .3 mg/kg: p = .021; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .327), but not mCherry rats (saline vs. .3 mg/kg: p = .530; saline vs. 1 mg/kg: p = .112; .3 mg/kg vs. 1 mg/kg: p = .611). The interaction between drug treatment and slope of the curve was significant.

**Figure 4.3.** DREADD stimulation of LC increased patch leaving in the foraging task. Top) hM3Dq-HA rats stayed in patches for fewer trials following CNO administration than following saline administration. CNO administration had no effect on behavior in mCherry animals. Bottom) Likelihood of leaving patches as a function of the value of leaving the patch. Points and error bars are the average likelihood of leaving the patch within each value of leaving bin ± standard error. Lines are the average GLMM predicted probability of leaving the patch for each trial within each value of leaving bin. There is an increased likelihood of leaving the patch due to CNO administration in hM3Dq rats but not mCherry rats. CNO administration increased decision noise, indicated by a shallower slope in hM3Dq rats but not mCherry rats.
between hM3Dq and mCherry rats (saline vs. .3 mg/kg: p = .013; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .107).

There were a few potential concerns with the simple form of the model described above. i) The model predicts that the likelihood of leaving the patch should continue to increase with the value of leaving, whereas the likelihood that rats left patches plateaued around 50%. Since rats often left the patch before the reward rate within the patch dropped substantially below the average reward rate, there are far fewer observations in this range. However, rats’ reduced likelihood of leaving relative to model predictions could still skew curve fits. ii) As rats’ behavior (e.g. number of trials, response times, omission rates) changes over the course of the session, the likelihood of leaving and its correlation with the value of leaving may also change, which is not accommodated by the model. iii) A related concern is that the logistic model assumes that all foraging decisions are independent and identically distributed, but there may be violations of this assumption, such as variation in the patch-leaving threshold between patches or an increasing bias towards staying in the patch with more trials spent in the patch (i.e. an increasing perseveration bias). To ensure that these concerns were not driving effects reported above, I controlled for them by i) fitting the same model, but omitting observations beyond the indifference point (GLM2), ii & iii) adding a random effect of time and random intercept for each patch, allowing this threshold to vary across patches (GLM3). Our results were robust to controlling for all three concerns. In GLM2, when omitting observations beyond the indifference point, hM3Dq rats were still more likely to leave patches on average and exhibited greater decision noise due to CNO administration, with no changes observed in mCherry rats (effect of treatment between hM3Dq and mCherry– saline vs. .3 mg/kg: p < .001; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .115; treatment x slope between hM3Dq and mCherry– saline vs. .3 mg/kg: p < .001; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .359). The same was true for GLM3, which included all data and all parameters in GLM1, plus a random effect of time and a random intercept for each patch (effect of treatment between hM3Dq and mCherry– saline vs. .3 mg/kg: p < .003; saline vs. 1 mg/kg: p < .001; .3 mg/kg vs. 1 mg/kg: p = .043; treatment x slope between hM3Dq and mCherry– saline vs. .3 mg/kg: p < .032; saline vs. 1 mg/kg: p = .001; .3 mg/kg vs. 1 mg/kg: p = .140). As GLM2 included different data than GLM1, I could not directly compare the likelihood of each model to assess model fit. However, I used AIC and BIC to examine whether
the additional parameters in GLM3 improved model fit compared to GLM1. It was unclear whether adding effects of time and a random intercept for each patch (GLM3) improved model fit, as AIC was lower (AIC\textsubscript{GLM1} = 31004.95, AIC\textsubscript{GLM3} = 30632.37) but BIC was greater (BIC\textsubscript{GLM1} = 31291.83, BIC\textsubscript{GLM3} = 31527.77). At the least, these analyses suggest that the factors listed above do not provide a compelling alternative to the interpretation that CNO increased decision noise.

**Discussion**

I used a patch foraging task to examine the influence of elevated LC tonic activity on decisions to exploit a depleting reward source within a patch vs. search for a new patch. I confirmed two predictions of AGT (Aston-Jones & Cohen 2005): (i) tonic LC stimulation impaired performance within the task, including reduced participation and increased omission rates, and (ii) tonic LC stimulation caused disengagement from patch exploitation, exhibited by earlier patch leaving associated with increased decision noise.

Our findings that LC tonic stimulation impaired performance within the task and caused disengagement from patch exploitation are consistent with predictions of AGT (Aston-Jones & Cohen 2005). To date, evidence in support of AGT has come from physiological recordings from LC neurons in non-human species (Aston-Jones et al. 1994, 1997; Clayton et al. 2004), and from studies in humans using pupillometry as an indirect measure of LC activity (Aston-Jones & Cohen 2005, Joshi et al. 2016, Varazzani et al. 2015). The latter have tested predictions of AGT in a variety of behavioral tasks designed to probe exploratory behavior, and have shown that baseline pupil diameter (a proxy for LC tonic firing) tracks utility in a foraging-like task (Gilzenrat et al. 2010), predicts decisions to explore an unknown reward vs. exploit known rewards (Jepma & Nieuwenhuis 2011), and tracks neural gain (Eldar et al. 2013, Warren et al. 2016), the mechanism by which NE release is proposed to impact processing noise. All of these findings have been correlational. Our findings are not only among the first to directly examine the relationship between LC function and decision-making but, perhaps most importantly, establish a causal relationship of LC-NE function in such behavior. Our findings indicate that increasing LC tonic firing increases the likelihood of changing behavior, and does so by increasing decision noise as predicted by AGT. These findings are consistent with the time course for effects of DREADD
stimulation on LC-NE activity and neural activity in other regions (starting ~10 min. following i.p. injection; Smith et al. 2016, Vazey & Aston-Jones 2014). Additionally, our findings corroborate previous work showing that manipulation of LC-NE activity specifically within the ACC regulates behavioral variability (Tervo et al. 2014).

A few notes of caution are warranted in interpreting the effects of chronic DREADD stimulation, as well as the task impairments observed. Although our DREADD manipulation was selective to LC-NE neurons, behavioral and neural effects of chronically stimulating LC tonic activity via DREADDs are not well characterized. In anesthetized animals, DREADD stimulation increases the firing rate of LC-NE neurons within the physiological range, to ~5Hz (Vazey & Aston-Jones 2014), but its impact on firing rates or NE release in awake behaving animals is unknown. Behavior similar to the immobility I observed has been reported using optogenetic stimulation in mice, for which the intensity of the stimulation is known. Photostimulating LC tonic firing for 5 or more minutes increased measures of anxiety (McCall et al. 2015), and high frequency photostimulation (5-20 Hz) induced behavioral arrest (Carter et al. 2010), a sustained period of immobility. One possible explanation for the immobility I observed is depletion of NE caused by long-term, high-frequency stimulation. Carter et al. (2010) found that 10 minutes of 10 Hz LC stimulation reduced NE levels in prefrontal cortex, and that treatment with NE reuptake inhibitors increased latency to behavioral arrest and reduced the duration of behavioral arrest due to high frequency stimulation. Additionally, the LC-NE system has broad efferent projections, and DREADDD stimulation affects NE release not only in regions involved in foraging decisions, but throughout the nervous system (Aston-Jones & Cohen 2005). It is possible that increased NE release into sensorimotor systems caused impairment in task performance and immobility. However, stimulation of LC-NE neurons using the same DREADD approach found no change of locomotor activity in a novel environment (Vazey & Aston-Jones 2014). To avoid potential confounds, I restricted our analyses to the period prior to the emergence of bouts of inactivity, to focus on task-related consequences of elevated LC activity. The cause of these bouts, their relationship to normal LC function, and any prodromal effects they may have had on the behaviors I observed all remain important questions for future research.

I also addressed several potential concerns that may have influenced the modeling of decisions in the foraging task: i) that the value of likelihood of leaving the patch plateaued at 50%,
despite an increasing value of leaving the patch, ii) factors influencing the rats’ behavior may have changed over the course of the session, and iii) violation of the assumption that all decisions are independent and identically distributed. Analyses that controlled for these factors increased confidence in our results. Controlling for these factors did not diminish the significance of the predicted results. One additional concern that is not accounted for in the analysis described above is how incorporating hyperbolic discounting, which was found in Chapter 3 to explain overharvesting, would influence the modeling results here. AGT would predict that stimulating LC would not influence the discount rate, but still would only influence decision noise.

Despite these limitations, our findings represent an important step in examining the behavioral function of LC tonic firing, providing causal evidence that increased LC tonic activity causes disengagement from current behavioral goals, and increases the likelihood of leaving a patch in a foraging task. Future studies will be necessary to test additional predictions of AGT, as well as predictions of related theories of LC-NE function (Bouret & Sara 2005, Sara & Bouret 2012, Yu & Dayan 2005), including whether LC tonic activity promotes exploration of options with unknown values. The involvement of the LC-NE system in a variety of fundamental processes – including arousal, attention, memory, and decision execution – and its dysfunction in a variety of psychiatric disorders – including anxiety, depression, and cognitive disorders – indicate the importance of studying the function of the LC-NE system. As LC stimulation impaired rats ability to focus on important task variables, our findings suggest an important role for LC in attentional deficits, such as those seen in attention deficit hyperactivity disorder. Future studies will provide an important foundation for characterizing the neural mechanisms underlying many forms of adaptive and maladaptive behavior.
Chapter 5: The Anterior Cingulate Cortex Monitors Decision Variables

Introduction

Study of the neural mechanisms of foraging behavior have elucidated the importance of the anterior cingulate cortex (ACC). ACC is a hub of cognitive control, and has been hypothesized to act as a monitor of conflict in information processing, activating upon sensing the pressure to act on two conflicting inputs, and signaling the need to exert cognitive control to reduce such conflict (Botvinick 2007; Botvinick et al. 2001, 2004). More recently, ACC has been proposed to compute the value of exerting cognitive control, taking into account both the benefits (reward) and costs (effort expended) for exerting control (Shenhav et al. 2013, 2016a, 2017). Under both of these theories, ACC is hypothesized to signal the need for increased cognitive control as reward in a patch depletes: as the value of staying becomes closer to the value of leaving, the decision about which course of action to takes gets more difficult and, as a consequence, is increasingly likely to engage control. This hypothesis is supported by recordings of single units in monkeys as well as fMRI of humans performing foraging tasks (Blanchard & Hayden 2014, Hayden et al. 2011, Kolling et al. 2012, Meder et al. 2016, Shenhav et al. 2014, 2016a). Activity of ACC neurons increases as monkeys spent more time in patches (Hayden et al. 2011). In a foraging task in which human participants decided to accept an offer of reward or reject it in favor of waiting for an alternative offer, ACC activity was greatest when the current offer was similar in value to the perceived average value of alternative options (i.e. when participants were equally likely to accept the offer or to reject it; Shenhav et al. 2014, 2016a). Furthermore, single units in ACC have been shown to encode multiple task variables. In foraging tasks, they encode information pertaining to opportunity costs – the value of rejected rewards or the delay (the cost in terms of time) to receive an accepted reward (Blanchard & Hayden 2014) – and they signal reward value across a variety of contexts (Heilbronner & Hayden 2016, Kennerley et al. 2011, Seo & Lee 2009). These results, among others, have led to alternative interpretations regarding the function of ACC: that it may be signaling the value of changing from a default to an alternative course of action (i.e. from accepting an offer or staying in a patch to rejecting the offer or leaving the patch; Kolling et al. 2012, 2016; Meder et al. 2016).
To directly test the function of the ACC in foraging, it is necessary to manipulate ACC activity in behaving animals. Lesions and inactivations of ACC in monkeys indicate that ACC is critical for adapting to dynamic environments, ones in which reward contingencies change over time (Kennerley et al. 2006, Shima & Tanji 1998), but causal manipulations of ACC have not been performed in foraging tasks in which the environment is static. Rodent models provide a number of tools to perform selective manipulations to neural circuits, but whether rodent ACC serves as a good model for human and primate ACC is under debate (Heilbronner & Hayden 2016). First, to confirm that rat ACC performs a similar function as human and monkey ACC in foraging, I recorded single and multi-unit activity in rats performing a patch foraging task. Next, I used chemogenetic manipulations to test the causal role of the ACC.

5.1 Rat ACC activity increases as rewards within a patch deplete

Methods

**Animals**

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

**Foraging task**

Rats were trained and tested in the foraging task as described in Chapter 4. Rats were tested on a wide range of patch types within each behavioral session, starting with 30-150 µL of reward and depleting at the same rate. Only one travel time was tested: 10 s. Rats were trained until reaching a criterion of stable behavior, then underwent surgery to implant an electrode array. After recovery from surgery, rats were retrained to perform the task while tethered to the recording system, and then tested for one session while neural activity was collected.
**Surgery**

Anesthesia was induced and maintained using isoflurane (5% and 2% respectively). Rats were aligned in a stereotaxic frame such that bregma was 2 mm below lambda (~15° head angle). Rats were implanted with a 32 channel wire array (either 32 single wires or 8 tetrodes, 50 or 25 μm stainless steel wire respectively) at the following coordinates from bregma: +2.0 mm AP, ± 0.5 mm ML, and -1.4 mm DV from brain surface. After surgery, rats received carprofen (.1 mg/kg i.p.) for post-operative analgesia, and were given 7 days to recover before further testing.

**Electrophysiological recording**

Signals were acquired using a 32 channel digitizing headstage, passed through a commutator to allow the animal free movement, and sent to a Plexon Omniplex Recording System. Signals were filtered between .6-6 kHz and sampled at 40 kHz. Spikes were detected at a threshold of 5 times the standard deviation of the signal, and manually sorted in Plexon Offline Sorter software using a combination of principal components and other waveform features (peak, peak to valley ratio, etc.). 107 single and multi-units recorded from 12 rats were considered: all isolated units with <5% ISI violations (ISI < 2 ms) that maintained stable firing throughout the session were included for analysis. All analyses included all of the 107 recorded units, and did not distinguish between single and multi-units.

**Histology**

Following testing, rats were anesthetized with 2% isoflurane and current (20 μA for 10 s) was passed through the most anterior and most posterior recording wires to create a lesion at the tip of the electrodes to mark recording sites. After 24 hr to allow for gliosis to develop, rats were transcardially perfused with saline and 4% paraformaldehyde. Brains were post-fixed in 4% paraformaldehyde overnight, then stored in 30% sucrose in phosphate buffered saline (PBS) for at least 48 hr before freezing and cryosectioning. 40 μm sections were taken throughout the ACC. To identify electrode locations, a 1:12 series was stained with an antibody against GFAP: tissue was first blocked for 1 hr in normal donkey serum in PBS with .1% Triton-X, then incubated in primary antibody (Rb x GFAP) overnight. Sections were washed 3x5 min in PBS, then
incubated in secondary antibody (Dk x Rb conjugated to Alexa-fluor 488). Sections were washed again for 3x5 min in PBS, mounted on slides and cover slipped.

**Data Analysis**

All analyses were conducted using R. Analysis was focused on the time around decisions to stay in vs. leave a patch (before and after lever presses and nose pokes). Peri-event time histograms (PETHs) of time from 5 s before to 2 s after lever press to stay in the patch and from 5 s before to 12 s after nose poke to leave the patch were created using a moving average with 250 ms window that moved in 50 ms bins. To compare activity across units, firing rates were z-scored: baseline activity was measured in 1 s bins throughout the entire task, firing rates within each bin were z-scored according the mean and standard deviation of the baseline activity.

To examine the effect of relevant task variables – reward for staying in the patch, the number of trials until leaving the patch, the value of leaving as defined by MVT (the difference between the cumulative rate across patches and reward rate offered on the current trial), and response time (RT) as a proxy for choice difficulty – on activity, poisson regressions were fit to each unit on the time before and time after lever presses, separately. There were 9 regressors total: time, the four task variables above, and four interaction terms (each of the four variables x time).

First, heterogeneity among the activity of units was assessed by looking at the correlation of regression coefficients before and after lever press for each unit. Next, to assess to what degree ACC units were encoding each of the task variables, the coefficient of partial determination (CPD), or partial $R^2$ was calculated for each variable (Cai et al. 2011, Kennerley et al. 2011, Miller et al. 2018). To account for correlations among regressors, CPD provides a measure of how much variance is explained by the regressor of interest that cannot be accounted for by other variables. CPD was calculated for each unit by fitting a regression model containing all regressors, $X_{ult}$, and another without the regressor(s) of interest, $X_i$: 

$$ CPD = 1 - \frac{\text{variance explained by } X \text{ in model } X_{ult}}{\text{variance explained by } X_i \text{ in model } X_{ult}} $$
\[ \text{CPD}(X_i) = \frac{SSE(X_{-i}) - SSE(X_{all})}{SSE(X_{-1})}, \]

where SSE is the sum of squared error. CPD was calculated for each unit and aggregated across time before and after lever press. Population CPD was aggregated across units, \( u \):

\[ \text{CPD}_{\text{pop}}(X_i) = \frac{\sum_u [SSE(X_{-i}) - SSE(X_{all})]}{\sum_u SSE(X_{-1})} \]

To test for statistical significance, null distributions of regression coefficients and CPD values were created by re-fitting regression models on data with spike counts shuffled across time bins and across trials for each unit (100 shuffles per unit). Regression coefficients and CPD values were deemed significant if they exceeded 99% of the values in the null distribution (regression coefficients in the top or bottom .5% of null distribution, CPD values in the top 1%). Distributions of regression coefficients and CPD values for each unit were tested against one another using Wilcoxon signed-rank tests and the Holm method to correct for multiple comparisons.

**Results**

*Decisions to stay vs. leave become more difficult as patches deplete*

Rats were tested in the foraging task with many patch types, with starting reward ranging from 30-150 \( \mu \text{L} \) (as in Chapter 4). In patch foraging tasks, animals typically receive large rewards upon entering a new patch and, over time, these rewards deplete to the threshold at which animals decide to leave (the average reward rate across all patches per MVT). Thus, the value of staying in a patch is always much higher than the value of leaving a patch upon entry and, over time, the two values converge as the value of staying depletes. Per MVT, animals leave a patch when the value of staying is equal to the value of leaving, such that the maximum value of leaving is also the point of maximal choice difficulty, and animals are rarely tested on decisions in which the value of leaving is much greater than the value of staying (i.e. decisions in which animals should have an easy decision to leave the patch). A wide range in patch types, particularly patches that start with very low reward volume (30-45 \( \mu \text{L} \)), were chosen such that rats...
occasionally encountered patches for which the starting reward was below their average threshold for leaving a patch. In these patches, the value of leaving was greater than the value of staying, and decisions to leave these patches should be easier than decisions to leave patches that start with greater reward and deplete to their leaving threshold.

To characterize rat foraging three quantities were assessed: time spent in patches, the reward rate at which rats decided to leave patches, and response times – a measure of choice difficulty, as greater response times indicate rats spent more time deliberating about the decision. As in Chapter 4, rats stayed for more trials in patches that started with greater rewards ($\beta = 59.275, \text{SE} = 1.738, p < .001$; Figure 5.1A). As predicted by MVT, the average threshold at which rats decided to leave patches was consistent across medium (60-105 $\mu$L) and high (120-150 $\mu$L) starting reward patches, but as dictated by task design, rats left low starting reward patches (30-45 $\mu$L) at a lower threshold than both medium ($\chi^2(1) = 163.144, p < .001$) and high reward patches ($\chi^2(1) = 130.297, p < .001$). As patches depleted and rats became closer to leaving, response times increased ($\beta = .283, \text{SE} = .030, p < .001$), indicating that decisions became more difficult as patches depleted (Fig 5.1C). As rats left low starting reward patches at a lower reward

![Figure 5.1](image)

**Figure 5.1.** A) The number of trials spent in patches as a function of the starting reward. B) The reward rate threshold at which rats decided to leave patches (reward rate on the last decision to stay in each patch). C) Response time as rats became closer to leaving a patch. Zero trials remaining is the decision to leave, -1 the last lever press, and so on. Points and error bars represent mean ± standard error.
rate threshold than their average leaving threshold (i.e. the relative value of leaving was greater),
I hypothesized that decisions to leave these patches would be less difficult, and that response
times on these decisions would be faster, than decisions to leave medium or high starting reward
patch types. However, rats response times on leave trials were not different between low vs.
medium vs. high starting reward patches ($\beta = .970, SE = .782, p = .239$), indicating that decisions
to leave low starting reward patches were not less difficult.

**ACC neurons are sensitive to patch depletion**

To examine activity at the time of decisions, peri-event time histograms (PETHs) were
created around the time of decisions to stay (lever presses) and decisions to leave (nosepokes;
Figure 5.2). Qualitatively, ACC activity, on average, increased leading up to decisions to stay, and
was inhibited (indicated by z-score < 0) after decisions when rats received reward. As patches
depleted and response times were greater (e.g. on the last decision to harvest from the patch),
ACC activity was elevated throughout the period of deliberation. Following decisions to leave,
ACC activity was not inhibited as it was when receiving reward, but its activity only slightly
decayed throughout the travel time to the next patch.

Further analysis was focused on the time around decisions to stay in the patch. To
determine the variables that influenced ACC activity throughout this time, separate regression
models were fit to assess the effects of i) reward for staying in the patch, ii) the number of trials
until leaving, iii) the relative value of leaving the patch, iv) response time and v-ix) the effect of
each of these four variables on the slope – the change in activity over this time period. The
number of trials in the patch represents an important variable that provides information about the
state of the decision process, beyond what can be inferred from the current reward magnitude or
value of leaving the patch. The relative value of leaving was defined by MVT as the difference
between the cumulative reward rate across patches and reward rate offered on the current trial
(as it was calculated in Chapter 4). Response time served as a proxy for choice difficulty on the
current trial. Separate regressions were fit to each unit on the 5 s leading up to the lever press
and on the 2 s after the lever press, during which time rats received reward. In the 5 s prior to
lever press, 52% of units exhibited a significant change in activity over this time period
irrespective of task variables, with 32% increasing and 21% decreasing in activity over this time.
period. Although many units changed their activity over time, the average slope across all units was not significant (mean $\beta = .016$, $p = .87$). Similarly, a large proportion of units significantly changed activity in response to each task variables (i.e. significant regression coefficient for each variable; 53% to reward, 65% to trials, 49% to value of leaving, 64% to response time), but the average regression coefficients across units were not significant for any variables (mean $\beta_{\text{reward}} = .025$, $p = .715$; $\beta_{\text{trials}} = .031$, $p = 1$; $\beta_{\text{value}} = .03$, $p = 1$; $\beta_{\text{RT}} = .001$, $p = 1$). Furthermore, 48% of units significantly changed activity in the 2s following lever press, but again, the average change over this period was not significant (14% increase, 34% decrease, mean $\beta = -.171$, $p = .489$).

**Figure 5.2.** A) Average PETH of ACC activity around the time of decisions to stay in the patch, split by low reward patch types (left), medium reward patch types (middle) and high reward patch types (right). B) Average PETH of ACC activity around decisions to leave the patch. Activity was calculated via a moving average with a window of 250 ms in 50 ms bins. Lines represent the mean and ribbon ± the standard error across all units.
Similar to the dynamics before lever press, many units were significantly affected by all task variables (42% to reward, 54% to trials until leaving, 50% to value of leaving, and 55% to response time), but the average effect of these variables was not significant (mean $\beta_{\text{reward}} = .597$, $p = .715$; $\beta_{\text{trials}} = .038$, $p = 1$; $\beta_{\text{value}} = .068$, $p = 1$; $\beta_{\text{RT}} = .032$, $p = .843$).

To further assess activity patterns of individual units, I looked at correlations between the change in activity over time (the slope), the effect of trials until leaving, and the time-by-trials interaction before and after decisions (Figure 5.3). There were four significant correlations: i) units that increased (or decreased) in firing before lever press tend to do the opposite after lever press (pre:time vs. post:time, $r = -.328$, $p = .014$); ii) units that increase average firing before decisions as rats become closer to leaving (i.e. negative effect of trials until leaving) maintain their increased average firing after decisions (pre:trials vs. post:trials, $r = .46$, $p < .001$); iii) for units that increase average firing before decisions as rats become closer to leaving, their slope decreases as rats become closer to leaving – units that increase average firing rise slower over the course of trials (pre:trials vs. pre:time-trials $r = .7$, $p < .001$); and iv) similarly, units that exhibit greater activity after decisions as rats are closer to leaving a patch tend to have more positive changes in activity in the time after lever press (i.e. they increase more or decrease less over the course of trials (pre:time trials vs. post:time trials $r = .7$, $p < .001$).

![Figure 5.3](image)

**Figure 5.3.** Correlation matrix for the change in ACC activity over time (slope), the effect of trials until leaving, and time x trials interaction both before and after decisions to stay in the patch. Color indicates strength of the correlation (pearson’s r) and X’s indicate insignificant correlations. Threshold for significance was $p = .05$, corrected for multiple comparisons using the Holm method.
time on these trials; post:trials vs. post:time-trials $r = -.765$, $p < .001$); and similarly, iv) for units that increase in average firing after decisions as rats are closer to leaving a patch, their slopes increase over the course of trials – units that typically exhibit a rapid decrease in firing will decrease more slowly after decisions over the course of trials in a patch (post:trials vs. post:time-trials $r = -.765$, $p < .001$). Altogether, these correlations describe dynamics consistent with those seen in Figure 5.2 – as patches deplete, ACC units tend to exhibit greater firing rates. Rather than ramping up faster, greater activity is maintained throughout an extended period of time before lever press, and is further maintained after the decision during reward consumption.

**ACC units encode multiple task relevant variables**

As many important variables in the foraging are correlated with one another – the value of leaving a patch and response times both increase as reward depletes and the number of trials until leaving decreases – the coefficient or partial determination (CPD, or partial $R^2$) was calculated to assess encoding of specific variables by ACC neurons. CPD is a measure of the additional amount of variance explained by a variable once all other variables have been accounted for. CPD was calculated for the four variables identified above: reward, trials until leaving, the value of leaving, and response times, all with their interaction terms. CPD for each unit was aggregated across the two regression models before and after decisions. Specific encoding of all variables was quite weak (mean CPD$_{\text{reward}} = .0030$, CPD$_{\text{trials}} = .0055$, CPD$_{\text{value}} = .0036$, CPD$_{\text{RT}} = .0036$), but all ACC units significantly encoded at least one predicting variable, and the vast majority of units exhibited significant encoding for all four predictors (1% encode 1 variable, 6% encode 2, 13% encode 3, and 80% encode all four; Figure 5.4). Furthermore, there were significant correlations among all variables regrading the strength of encoding – units that exhibit stronger encoding for any given variable tend to exhibit stronger encoding for the other three (Figure 5.5A). To test whether certain variables are encoded more strongly than others across units, pairwise wilcoxon signed-rank tests were performed on the log of the distribution of CPDs across units for each variable. There was significantly stronger encoding for the trials until leaving the patch compared to reward magnitude ($p < .001$) and compared to the value of foraging ($p = .008$). Encoding for trials in the patch was slightly, but not significantly stronger than response time ($p = .125$). There were no significant differences in the strength of encoding
between any other pairs. Lastly, I examined the strength of encoding across the population by aggregating CPD across all units. Consistent with previous analyses, there was slightly stronger encoding for the number of trials until leaving the patch compared to all other variables both before and after decisions to stay in the patch (Figure 5.5B). Overall, these analyses indicate that ACC units are engaged in the task, and suggest they may encode an important decision variable – time left in the patch – more strongly than reward and the value of leaving, but we cannot conclude that encoding is stronger for the time left in the patch vs. response times, a proxy for decision difficulty.

Figure 5.4. Distributions of CPD for each of the four task variables against one another. Each point represents the CPD for one unit, color coded for statistical significance. P values in the corner of the each plot are from wilcoxon signed rank test between the two distributions in the plot, the p value is plotted in the corner with the greater mean.
5.2 Chemogenetic manipulation of ACC activity does not influence foraging decisions

Recordings revealed that, similar to monkey and human ACC (Hayden et al. 2011; Kolling et al. 2012; Shenhav et al. 2014), the activity of ACC neurons in rats increased as rats became closer to leaving a patch. Next, I used chemogenetic stimulation and inhibition of rat ACC neurons to determine the causal role of the ACC in foraging decisions.

Methods

Animals

Adult Long-Evans rats were used (N = 24, Charles River, Kingston, NY). Rats were housed on a reverse 12 h/12 h light/dark cycle. All behavioral testing was conducted during the dark period. Rats were food restricted to maintain a weight of 85-90% ad-lib feeding weight, and were given ad-lib access to water. All procedures were approved by the Rutgers University Institutional Animal Care and Use Committee.

Figure 5.5. A) Correlation between CPDs for each variable among all units. Color indicates the strength of the correlation. Correlations among all variables were statistically significant. B) Population CPD (aggregated across all units) for each of the variables before and after decisions to stay in the patch. Black line represents 99% confidence interval around null distribution (i.e. significance level).
**Virus injection surgery**

Prior to behavioral training and testing, rats underwent surgery to inject virus encoding DREADD receptors of GFP as a control. Anesthesia was induced and maintained using isoflurane (5% and 2% respectively). Rats were aligned in a stereotaxic frame such that bregma was 2 mm below lambda (~15° head angle). Rats received bilateral injections of an AAV8-CamKII α-hM3Dq-mCherry (n = 9), AAV8-CamKIIα-hM4Di-mCherry (n = 9), or AAV8-CamKIIα-EGFP (n = 6) into the ACC. Virus injection was targeted to the anterior portion of Cg1, similar to the location of recordings. Two small injections (100-200 nL) were made bilaterally at the following coordinates from bregma: +2.0 and +1.5 mm AP, ± 0.5 mm ML, and -1.4 mm DV from brain surface. After surgery, rats received carprofen (.1 mg/kg i.p.) for post-operative analgesia, and were given 7 days to recover before further testing.

**Foraging task**

Rats were trained and tested in the foraging task as described in Chapters 4 and 5.1. Rats were tested on a wide range of patch types within each behavioral session, starting with 30-150 µL of reward and depleting at the same rate. Only one travel time was tested: 10 s. Once rats exhibited stable behavior in the foraging task, pharmacological testing commenced. Rats received an i.p. injection of either saline or CNO (.3 mg/kg) every other day (one day with no injection between each injection day). Rats were tested for three sessions with each treatment.

**Histology**

Animals were perfused and tissue was processed as in Chapter 4. To confirm expression of hM3Dq and hM4Di in ACC neurons, mCherry fluorescence was amplified by staining with a primary antibody against mCherry protein (Ms x mCherry) and secondary antibody conjugated to Alexa-fluor 594 (Dk x Ms), using the same protocol as described in Chapter 4. All animals exhibited at least unilateral expression. Confirmation of changes in activity due to CNO administration are forthcoming.
Data Analysis

Analyses of foraging behavior were conducted in R. Mixed effects models were used to test for changes in time spent in patches and response times, with random effects for drug treatment and starting reward of patches or trials spent in the patch, respectively. Post-hoc tests were conducted using the phia package (de Rosario-Martinez 2015). As in Chapter 4, mixed effects logistic regression was used to assess the likelihood that rats decided to leave a patch as a function of the value of leaving (per MVT).

Results

Chemogenetic manipulations do not affect foraging decisions

To assess the effects of stimulating and inhibiting ACC neurons on foraging behavior, rats expressing the stimulating hM3Dq DREADD (n = 9), the inhibiting hM4Di DREADD (n = 9), or EGFP (n = 6) as a control were tested in the foraging task. Across all groups and treatments, rats stayed longer in patches that started with greater rewards ($\beta = 74.541$, SE = 4.493, $p < .001$). Administration of CNO had no affect on foraging decisions in either the hM3Dq, hM4Di or EGFP rats (Figure 5.6A): within groups, CNO did not significantly affect the number of trials rats spent in patches (EGFP: $\chi^2(1) = 2.468$, $p = .349$; hM3Dq: $\chi^2(1) = .019$, $p = 1$; hM4Di: $\chi^2(1) = .064$, $p = 1$), and across groups, the effect of CNO did not differ (EGFP-hM3Dq: $\chi^2(1) = 1.694$, $p = .508$; EGFP-hM4Di: $\chi^2(1) = 1.888$, $p = .508$; hM3Dq-hM4Di: $\chi^2(1) = .007$, $p = .934$).

Response times were analyzed both as a function of the number of trials until leaving the current patch and of the number of trials rats have spent in the current patch. Across groups and treatments, decision times increased as rats spent more time in patches ($\beta = .046$, SE = .013, $p = .003$) and as rats became closer to leaving a patch ($\beta = .035$, SE = .016, $p = .038$). CNO administration had no affect on response times based on trials spent in patches nor to the number of trials until leaving (Figure 5.6B), either within or across groups (trials in patch: EGFP: $\chi^2(1) = .105$, $p = 1$; hM3Dq: $\chi^2(1) = .326$, $p = 1$; hM4Di: $\chi^2(1) = .399$, $p = 1$; trials until leave: EGFP: $\chi^2(1) = .012$, $p = 1$; hM3Dq: $\chi^2(1) = .226$, $p = 1$; hM4Di: $\chi^2(1) = .752$, $p = 1$).
Discussion

Multiple studies have implicated an important role for the anterior cingulate cortex in foraging. In the present study, I provide evidence that rats can serve as a useful model to investigate role of the ACC in foraging decisions. Via single and multi-unit recordings, I show that activity of rat ACC is strikingly similar to that of humans and monkeys performing similar foraging tasks. Like in monkeys (Hayden et al. 2011), rat ACC neurons increase at the time of decisions, and they exhibit greater firing rates as patches deplete. In the present study, the time course for the increase in ACC activity did not perfectly match that of Hayden et al. (2011). In that study, ACC neurons exhibited large phasic responses that preceded decisions, and the phasic responses grew in magnitude as patches depleted. In rats, ACC activity increases earlier and greater firing rates are maintained during the period of deliberation. One reason for this discrepancy could be time intervals in the task: in the monkey patch foraging task developed by Hayden et al. (2011), there were very short delays in between decisions (~1-2 s), whereas in our
rat foraging task, these delays span 7-12 s, such that ACC activity has more time to increase or remain elevated.

Human fMRI studies have found that ACC activity is greater when the current offer of reward is similar to the value of searching, and that ACC encodes the value of searching or choice difficulty, particularly in the period after decisions (Kolling et al. 2012, Meder et al. 2016, Shenhav et al. 2014, 2016). This post-decision encoding of the value of unchosen vs. chosen options has been corroborated by single-unit recordings in monkeys (Blanchard & Hayden 2014).

In the present study, rat ACC neurons not only exhibited greater average firing rates as patches depleted, but they also encoded the value of leaving a patch and choice difficulty (as assessed via response times), before and after decisions. Across the population of ACC neurons, encoding of all of these variables was slightly stronger after decisions. Overall, these results suggest that rat ACC exhibits similar responses to humans and monkeys in foraging tasks.

Despite numerous studies in humans, monkeys, and now rats that have shown that ACC neurons encode important task variables, chemogenetic manipulation of ACC activity produced no observable change in rat foraging behavior. Prior studies investigating the effect of inactivation or lesions to ACC in monkeys have found that ACC is critical for switching behavioral strategies upon a change in reward contingency (Kennerley et al. 2006, Shima & Tanji 1998). However, due to changes in reward contingency, these tasks may be better described as exploit-explore tasks, in which the optimal behavior requires exploration or sampling of alternative options to learn their value when changes in the environment are detected. Foraging tasks, on the other hand, assume complete knowledge of the reward contingencies, and predictions of optimal foraging theory no longer apply when this assumption is broken (Stephens & Krebs 1986). ACC’s lack of influence in basic foraging decisions doesn’t preclude a role for ACC in guiding selection of behavioral strategies, but it provides an important challenge to this theory. One potential explanation is that ACC may play an indirect role, providing information to other systems that guide action selection, such as the locus coeruleus (LC, Aston-Jones & Cohen 2005). Under this hypothesis, LC would use information from ACC to decide to exploit vs. explore. Consistent with lesion studies (Kennerley et al. 2006, Shima & Tanji 1998), this hypothesis would not necessarily predict a role for ACC in foraging, but this signal may be used to guide decisions to explore – to switch to an alternative behavioral strategy. The hypothesis that ACC encodes choice difficulty may provide a
simpler explanation: stimulating ACC would signals the need to apply cognitive control to stay vs. leave decisions, and inhibiting ACC eliminates this signal to increase control. However, foraging decisions may not require increased cognitive control. Thus, animals may continue make adaptive foraging decisions with or without a functioning ACC.

The results presented in this chapter establish that rats can serve as model for studying the role of the ACC in decision-making and provide an important initial test of the influence of the ACC in foraging decisions. ACC’s lack of influence on foraging decisions provides important information into its function, and emphasizes the need for future examination into the causal role of the ACC in decision-making.
Chapter 6: General Discussion

Animals, including humans, frequently encounter stay vs. leave problems in many domains and over many timescales – in deciding to stick with one’s current employment, to stick with one’s romantic partner, or when foraging for resources. Given the ubiquity of stay vs. leave problems, making adaptive decisions is critical for animal health and well-being (Pearson et al. 2014). The work in this thesis provides important insights into how animals make foraging decisions, and advances knowledge of two neural circuits hypothesized to play a role in foraging. Ultimately, this work lays the foundation for rigorous examination into the neural mechanisms of foraging decisions.

In Chapter 2, a novel foraging task for rats is presented. Using this task, I show that rat foraging behavior in a controlled laboratory task is similar to foraging behavior of other humans and monkeys in similar lab foraging tasks (Barack et al. 2017, Constantino & Daw 2015, Hayden et al. 2011). As it is performed in operant conditioning chambers, this task is amenable to application of techniques to record and manipulate neural activity, representing a powerful new paradigm to study the neural mechanisms of foraging decisions.

In Chapter 3, rat behavior in the foraging task was thoroughly characterized. I identified the cognitive biases rats use in making foraging decisions, and examine a variety of computational models to formally test what biases can explain behavior across many foraging environments. I found that rats, like many other species, exhibit a bias towards accepting smaller rewards within a patch rather waiting for larger rewards received after traveling to a new patch. This bias was best explained by hyperbolic discounting – the value of future rewards (i.e. rewards in future patches) are discounted as a consequence of being available later, with rewards discounted more steeply in the near future than over longer-terms. This was evident in rats behavior: they were less likely to stay in a patch if the reward for staying was received after a short delay, indicating the value of the reward was reduced due to a delay to receive it, but they also were less likely to stay in a patch if the delay after the reward but before the next decision was longer, indicating that they consider future rewards to some degree. The same hyperbolic discounting model was able to explain rats behavior in a different kind of intertemporal choice task: one for which options are presented simultaneously, rather than in series in the foraging
task. This result contradicts the results of a few recent studies which have suggested that temporal discounting does a poor job of explaining animal foraging behavior (Blanchard & Hayden 2015, Carter & Redish 2016, Carter et al. 2015). These studies have suggested that animals are more patient or less impulsive in foraging tasks than simultaneous choice tasks – they exhibit less of a bias towards accepting more immediate rewards when more profitable rewards are available in the future. Our results question this interpretation, and provide evidence for similar valuation systems across different task structures. Future studies examining hyperbolic discounting across foraging and intertemporal choice tasks in other species, such as humans and monkeys, could provide an important test of this hypothesis.

Next, I began to investigate the neural mechanisms of foraging behavior in Chapters 4 and 5, starting with the locus coeruleus (LC). The Adaptive Gain Theory (AGT) hypothesized that LC regulates the balance between exploiting known rewards vs. exploring – sampling options for the purpose of learning – for better opportunities. According to AGT, increased LC activity increases noise in downstream circuits, which impairs their ability to narrow focus to a few relevant task variables and makes these circuits more responsive to “distracting” information. Thus, neural circuits are relatively less likely to respond selectively to task-relevant stimuli and more likely to respond to alternative stimuli and pursue other behaviors. I tested this idea in the foraging task by performing chemogenetic stimulation of LC neurons. Stimulating LC initially produced a behavioral phenotype consistent with AGT: rats left patches earlier not due to a change in perception of reward value within patches, but due to increased decision noise. However, over a longer timeframe (30-60 minutes), LC stimulation produced a general behavioral impairment, including a lack of participation in the task, a large increase in response times, and it induced extended bouts of immobility. Impaired performance within a go/no-go task has been observed previously during periods elevated LC activity (Aston-Jones et al. 1994, Usher et al. 1999), although I observed a greater degree of such impairment, with our manipulation of LC. This study provides evidence that LC activity regulates task performance as hypothesized by Aston-Jones & Cohen (2005) – increased LC activity initially impaired rats’ ability to selectively attend to task-relevant information, and after prolonged stimulation, it completely impaired animals ability to perform the task. These results don’t confirm that LC tonic activity naturally increases, driving decisions to leave a patch – testing this prediction would require
electrophysiological recording and temporally specific inhibition of LC neurons as rats perform the foraging task. Nevertheless, this study was an important first step in showing the ability of LC to influence decision – the mechanism by which AGT predicts it drives exploratory decisions.

Lastly, I investigated the function of the anterior cingulate (ACC) in foraging decisions. The ACC has received considerable attention in foraging tasks, as some of the first papers investigating the neural mechanisms of foraging decisions identified the ACC as playing a critical role. These studies have found that ACC activity tends to increase as the value of a current offer of reward decreases relative to the value of alternative options (Blanchard & Hayden 2014, Hayden et al. 2011, Kolling et al. 2012, Meder et al. 2016). Furthermore, these studies have shown that ACC encodes a number of task-relevant variables, including the value of searching for a better reward, opportunity costs for the chosen option, and the difficulty in choosing to stay vs. leave (Blanchard & Hayden 2014, Kolling et al. 2012, Shenhav et al. 2014, 2016). Existing findings have been interpreted as evidence that ACC signals the value of switching to an alternative course of action (leaving a patch vs. the default option of staying; Kolling et al. 2012, 2016; Rushworth et al. 2012). Others have argued that ACC activity reflects the difficulty in choosing to stay vs. leave, and that this signal represents the need to exert cognitive effort to make an adaptive decision (Shenhav et al. 2013, 2014, 2016a,b). To test the function of the ACC in foraging, I first recorded ACC neurons in rats performing the foraging task to confirm that rat ACC exhibits responses similar to that of humans and monkeys. I replicated many of the results described above: rat ACC activity increased as patches depleted and rats became closer to leaving a patch, and that rat ACC neurons encoded many task-relevant variables, including the number of trials until leaving the current patch, the value of leaving the current patch, and response times (a proxy measure for the difficulty in choosing to stay vs. leave). I next tested the causal role of ACC using chemogenetic stimulation and inhibition (Alexander et al. 2009, Armbruster et al. 2007). Surprisingly, neither stimulation nor inhibition influenced rat foraging behavior. This result poses an important challenge to the ‘value of switching to an alternative behavior’ account. If ACC is hypothesized to drive this change in behavioral strategy (from stay to leave), then stimulating this signal should drive earlier switching and inhibiting it should drive later switching. The ‘value of exerting effort’ account would predict that increased ACC activity would encourage increased commitment to following MVT, whereas inhibiting this signal eliminates the
drive to increase commitment to following MVT. Ultimately, if rats are not under any pressure to deviate from MVT (i.e. there is no other tasks worth exploring), the ‘value of effort’ account could predict that rats would continue to follow MVT with or without a functioning ACC. To confirm this prediction, it would be necessary to examine the causal role of ACC in an exploit-explore task in which animals could choose to exploit a foraging environment vs. explore for alternative, potentially more rewarding tasks.

Altogether, the work presented in this thesis advances our understanding of how animals make foraging decisions and to the role of two neural circuits in foraging decisions. The behavioral studies and modeling results presented in Chapter 3 – that rats exhibit time preferences in foraging tasks that can be explained by hyperbolic discounting – suggests that regions implicated in temporal discounting, such as limbic vs. prefrontal and parietal regions (McClure et al. 2004, 2007), may play important roles in foraging behavior. Furthermore, studies of the function of LC and ACC in Chapters 4 and 5 provide a foundation for future studies to examine the neural circuits important for optimizing decision-making. The finding that LC influences decision noise makes interesting predictions regarding the role of LC across many domains – most importantly, that it may regulate exploration. Confirmation that rat ACC exhibits similar activity as monkeys and humans further validates using rodents as a model to study the function of the ACC. Our findings suggest that ACC may not have a major influence on decisions in stationary environments, but this knowledge combined with existing findings suggests that ACC, possibly through previously hypothesized interactions with LC (Aston-Jones & Cohen 2005), may play a role in optimizing decision-making in dynamic environments.
References


criteria: reward rate ultimately beats accuracy. . 640–57
Assessment. Front. Integr. Neurosci. 5:
Journal of Statistical Software. 67(1):1–48
Bateson M, Kacelnik a. 1996. Rate currencies and the foraging starling: The fallacy of the
Blanchard TC, Hayden BY. 2014. Neurons in Dorsal Anterior Cingulate Cortex Signal
Blanchard TC, Hayden BY. 2015. Monkeys Are More Patient in a Foraging Task than in a
Standard Intertemporal Choice Task. PLOS ONE. 10(2):e0117057
Blanchard TC, Pearson JM, Hayden BY. 2013. Postreward delays and systematic biases in
measures of animal temporal discounting. Proceedings of the National Academy of
Sciences. 110(38):15491–96
Boorman ED, Rushworth MF, Behrens TE. 2013. Ventromedial Prefrontal and Anterior Cingulate
Cortex Adopt Choice and Default Reference Frames during Sequential Multi-Alternative
Botvinick MM. 2007. Conflict monitoring and decision making: reconciling two perspectives on
Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. 2001. Conflict monitoring and
cognitive control. Psychological review. 108(3):624–52


Carter EC, Pedersen EJ, McCullough ME. 2015. Reassessing intertemporal choice: human decision-making is more optimal in a foraging task than in a self-control task. Frontiers in Psychology. 6:


Hayden BY, Platt ML. 2006. Fool me once, shame on me—fool me twice, blame the ACC. Nature Neuroscience. 9(7):857–59


Pearson JM, Hayden BY, Platt ML. 2010. Explicit Information Reduces Discounting Behavior in Monkeys. Frontiers in Psychology. 1:


R Core Team. 2015. R: A Language and Environment for Statistical Computing

Rushworth MF, Kolling N, Sallet J, Mars RB. 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Current Opinion in Neurobiology. 1–10


Shenhav A, Straccia MA, Botvinick MM, Cohen JD. 2016b. Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. Cognitive, Affective, & Behavioral Neuroscience. 16:1127–39


