Decision-Making and Optimal Foraging:

Norepinephrine and the Exploration-Exploitation Tradeoff

Abstract

The decision to exploit a resource or explore the environment presents a common economic tradeoff. The decision-making process of this tradeoff, however, is not well understood. Recent neurobiological findings show that Norepinephrine may regulate the transition between exploitation and exploration behaviors through altering levels of arousal. Using foraging theory models, I developed a mouse experiment to test Norepinephrine's role in the exploit-explore paradigm. The experiment requires the mice to receive smaller rewards that arrive predictably and reliably, or receive larger rewards that arrive unpredictably. Compared to normal mice, mice with deficient Norepinephrine function show a tendency towards exploitation behaviors rather than exploration. This demonstrates that proper Norepinephrine functioning is essential for the evaluation of the exploit-explore tradeoff.

Matthew Pease[1]

Mentor: Dr. Rachel Kranton and Dr. John Pearson

[1] I am currently working as a research assistant in the laboratory of Dr. Michael Chee at the Duke-NUS Graduate Medical School in Singapore. I am researching the cause of Alzheimer's disease. In addition, I am applying to medical school and hope to attend medical school next year.

Acknowledgements

I am especially grateful to both of my thesis advisors, Dr. Rachel Kranton and Dr. John Pearson. Dr. Kranton was invaluable for helping me draft and write this thesis. Dr. Pearson helped me understand the science, develop the experiment, and analyze the data. I am also incredibly thankful for the help throughout the research and writing process from my Econ 199S professor Dr. Michelle Connolly. Lastly, I would like to thank my fellow classmates for their helpful comments during presentations and early drafts.

I. Introduction

The choice to exploit or explore poses a dilemma: individuals can either receive immediate rewards or learn new information that may lead to a larger future reward. While both humans and animals frequently face this exploit-explore tradeoff, the neural mechanisms of the decision processes are poorly understood. Recent work shows that separate brain regions control the implementation of exploitation and exploration actions (Daw et al., 2006). The mechanism for alternating between these two actions, however, remains unclear. A neuromodulatory protein called Norepinephrine (NE) may be responsible by regulating attention and arousal. During heightened arousal, individuals are unable to execute a task and, instead, explore the environment. Conversely, moderate levels of arousal allow an individual to focus on a particular task with high accuracy.[2] To test the effects of NE, I developed an exploit-explore task that mimics natural situations. In my experiment, I compared the performance of a breed of genetically altered mice with impaired NE functioning, called Norepinephrine Transporter knock outs (NET's), to a group of genetically normal mice. The NET mice performed exploitation for 80.5% of exploit-explore actions during baseline conditions, compared to 70.4% for the normal mice.[3] Overall, the NET mice exhibited a tendency to exploit a resource rather than explore the environment.

The tradeoff between exploitation and exploration is present in many real world situations. For example, a common day laborer experiences an exploit-explore tradeoff when deciding where to work each day. The laborer can choose to work a construction job

---

[2] In the opposite extreme, low arousal corresponds to drowsiness and sleep.
[3] Baseline conditions refers to times when the Variable Arrival patch is inactive. This will be explained in Section V.

that pays a steady, but low wage, or at a nearby strawberry farm to pick valuable strawberries. Unfortunately for the laborer, strawberries are only occasionally and unpredictably ripe. The laborer wastes a day of work if he travels to the faraway strawberry farm and finds the strawberries are unripe. The laborer can choose to exploit a known resource (the construction job) or explore for the potentially more valuable job (strawberry picking). Likewise, a traveler faces an exploit-explore decision when entering a new town in search of food. The traveler can choose to eat at a chain restaurant with a familiar menu and food quality, or he or she can eat at an unknown, local venue. The local venue represents exploration, as the food quality is unknown, while the chain restaurant represents exploiting known information.

For years, behavioral ecologists have studied the exploit-explore tradeoff described above through optimal foraging theory (Hamelin 2006; Stephens and Krebs, 1986; Stevens, 2002). This theory provides a model for investigating animal behavior as animals search for resources such as food and hosts that are unevenly distributed in an environment. The resources are typically located in discrete patches that vary in quality (Charnov 1976). Just as the aforementioned day laborer can work at the steady construction job, some patches offer consistent, albeit less valuable, rewards. At any point, the animal can choose to search for a more valuable patch at another location. Moreover, some patches only offer rewards during certain periods that occur unpredictably. Exploring the environment and increasing knowledge of patch rewards allows the animal to maximize high reward actions and minimize low reward actions. The decision-maker must weigh the expected improvements in performance from information gathered while exploring with the lost opportunity to harvest resources (Wai-Tat Fu 2006; Stephens and Krebs, 1986).

Several common economic and game theory problems have similar tradeoffs. The single-armed bandit problem models how an individual explores the environment for information (Audibert et al., 2009; Whittle, 1980; Whittle, 1988; Gittins 1979; Bolton and Harris 1999). In the classical setup, a gambler walks into a room with several slot machines with variable, unknown payoff rates. Hoping to maximize his or her payoff, the individual must strategically probe each slot machine for information about the payoff rate. Gittins and Jones (1974) solved this problem by proposing that each slot machine is assigned an index to calculate the projected value of using a particular slot machine. The index incorporated the expected value of the slot machine and the expected increase in information about the slot machine payoff rate. To maximize reward, an individual simply chooses the slot machine with the highest index.

While theorists can find optimal solutions to the single-armed bandit and other similar foraging problems, these solutions frequently require complex calculations that are unrealistic for a person to perform given cognitive and time constraints (Wai-Tat Fu, 2006; Kahnman, 2002). Indeed, when researchers studied animals in the wild, some animals failed to closely follow Charnov's predictions. For example, some species of sea bass, moose, shrews, insects and wasps followed sub-optimal foraging strategies (Kamil 1983, Anderson 1984; Barnard and Brown 1981; Zimmerman 1981; Pyke 1977; Waage 1979; Krebs et al. 1974; Outreman et al. 2005). Despite this failure, Charnov's theorem provides a good framework that is capable of approximating animal behavior. Several animals, such as hummingbirds and the parasite nemeritis, closely approximate Charnov's theorem (Hubbard and Cook 1978). Apparently, animal species attempt to generate similar results to what Charnov predicts, but may not perform the complex calculations required for his

theory. Humans and animals have neural systems that execute other, cognitively feasible, processes that produce these behavioral outcomes.

New fields such as neuroeconomics explore how neural mechanisms affect decision-making and behavior. Classical economic theory builds models where humans are represented as rational agents. These agents spend time carefully planning all decisions (Mullainathan and Thaler, 2000; Davidoff, 1965). In reality, people can act irrationally and make suboptimal decisions. For example, stock market investors frequently hold losing stocks longer than a rational agent would (Koszegi 2008). Neuroeconomics combines methods from neuroscience, economics, and psychology to offer alternate models for the underlying processes of decision-making.

Additionally, neuroeconomics studies indicate that certain individuals deviate from standard behavior in systematic ways. Recent neurobiological findings show that various genetic and environmental factors can change behavior in certain categories of individuals. Parkinson patients taking particular types of medications, for example,[4] show an inclination towards a gambling addiction (Dodd et al., 2005; Driver-Dunckley et al., 2003). Likewise, the gene 5-HTTLPR increases the likelihood of depression when an individual experiences stressful life events (Pezawas et al., 2005; Hairiri et al., 2002; Caspi et al., 2003). In both of these cases, affected individuals will systematically deviate from optimal behavior. Similar interactions between genetics and the environment could have a large effect on individuals performing tasks such as foraging. Across the entire human population, large differences probably exist in the levels of neurochemicals present in each

[4] Specifically, the individuals are taking dopamine agonists. Dopamine is a neurotransmitter that regulates variety of functions including movement, pleasure, and attention. A dopamine agonist is a compound that activates dopamine receptors while dopamine is absent, mimicking the actions of dopamine in the brain.

person's brain. These differences, such as altered NE functioning, could account for systematic variations in behavior. Modern genetic techniques allow researchers to analyze these differences in groups of a population, and create a more comprehensive model of decision-making.

A new understanding of the neuromodulator Norepinephrine (NE) gives insight into the exploit-explore decision-making process. Mentioned previously, NE is a neurochemical that regulates arousal and attention behavior (Aston-Jones and Cohen, 2005; Berridge and Waterhouse, 2003; Jouvet, 1969; Robinson and Berridge, 1993; Wise and Rompre, 1989). Regulation of NE may cause individuals to either perform a task more efficiently or disengage from a task and explore the environment. This idea is motivated by findings showing that rat and monkey brain cells release NE when presented with arousing stimuli that normally elicit behavioral responses (Aston-Jones and Bloom, 1981; Brun et al., 1993). Further work showed that these brain cells have direct connections with brain areas associated with attention processing and motor response (Morrison et al., 1982; Foote and Morrison, 1987). Taken together, these findings led to a theory of NE function stating that NE may produce behavioral adjustments in attention level that optimizes performance while completing an exploit-explore task.

Investigating this theory will give economists a greater knowledge of exploit-explore decision-making. Economists can use this information to build more accurate models that more accurately depict human behavior and account for systematic deviations due to genetic factors. To assess NE's role in the exploit-explore tradeoff, I conducted an experiment where two cohorts of mice completed an exploit-explore task. Each night, the mice were individually placed in a small box with two portholes into which a mouse can

"nosepoke," an action whereby a mouse sticks his nose into a porthole to gain a reward. The portholes represent a foraging patch. One patch, called the Fixed Interval (FI) patch, offers a constant, low reward value. The other patch is called the Variable Arrival (VA) patch. This patch is either active for a defined period and offers a high reward, or inactive and offers no reward. The mouse chooses how much to alternate between exploiting the FI patch and exploring the VA patch to discover when the more valuable active VA patch is available for exploitation.

This experiment provides an opportunity to assess relative levels of exploitation and exploration in different groups of mice. Compared to normal mice, the NET mice showed a generally tendency towards exploitation. While the VA patch was inactive, the NET mice nosepoked the more valuable FI patch instead of exploring VA patch to determine if it was active. After the VA patch activated, the NET mice adjusted nosepoking behavior to a larger extent than normal mice to successfully exploit the active, highly valuable VA patch. This task demonstrates that NE helps regulate the exploit-explore tradeoff.

The rest of this paper is divided into seven sections. Section II examines the relevant economic literature and explains how foraging theory is useful for studying decision-making paradigms. Section III summarizes our current understanding of the neural mechanisms of the exploit-explore tradeoff. Section IV discusses NE and its role in the explore-exploit tradeoff. In Section V, the experiment is described in more detail. Section VI presents the theoretical framework for the experiment. Section VII discusses the analysis and results of the experiment. Section VIII concludes the paper.

II. Economic Literature Review

Optimal foraging theory originated from studying animal food-gathering strategies in natural habitats (Krebs, 1973). Foraging theorists developed models with four basic features: (1) how long an animal searches for patches; (2) which patch types the animal visits; (3) when an animal leaves a patch; (4) which type of food the animal consumes at a patch (Zimmerman, 1981). Overall, foraging theorists discovered that animals appear to choose strategies to maximize resource intake by balancing the resources gained from exploiting a discovered patch and the cost associated with searching for a more valuable patch. In this section, I describe an optimal foraging model and then compare this to a similar economics problem, the single-armed bandit problem. This problem describes the tradeoff present in my experiment. The remainer of my thesis will examine features (1) and (2) from above.[5]

2.1 Eric Charnov and the Optimal Foraging Problem

Eric Charnov developed the first mathematical model and optimal solution to foraging theory in 1974. Charnov proposed a patch leaving strategy that allowed an animal to gather resources at an average rate γ, the average resource capture rate for an environment (Hamelin et al., YEAR). The model set-up is similar to the foraging environment previously described in the introduction, with three distinct features (Pleasents and Zimmerman, 1979; Weins, 1976):

(A) A lone forager encounters resources arranged in nonrandom, discrete patches.

(B) Each patch exhibits diminishing returns to resource accumulation rate.

[5] While question three appears to be relevant to this thesis, the question actually requires a different mathematical approach.

(C) Other foragers are absent, leaving the lone forager to search without competition.

The forager's goal is to maximize cumulative resource intake. To do this, the forager faces a choice: exploit a known patch or explore the environment for a new patch. The forager chooses when to leave a patch, called the "patch leaving time," and then explores the environment until a new patch is found. The forager maximizes resource intake by relating the expected time exploring for a patch to the reward from exploiting a known patch. According to Charnov, an animal should exploit a known patch until the intake rate drops below γ. At this point, the animal should leave the patch to explore for new patches. Thus, an animal should search for a new patch when the marginal capture rate in a patch is below the average capture rate for an environment (Stephens and Krebs, 1986).

2.2 The Gittins Index and Slot Machines

While Charnov's paper led to the creation of optimal foraging theory, he mainly addressed feature (3): when does an animal decide to leave a patch. To understand the two features of foraging models that my thesis addresses, we must look at game theory and the single-armed bandit problem.[6] Introduced earlier, the single-armed bandit problem describes the strategies available to a gambler in a room with several slot machines with variable, unknown payoff rates (Whittle 1988). Each period t, the gambler uses slot machine i with a mean payoff rate $x_i(t)$, and gains a reward $g_i(x_i(t),t)$. The slot machines do not have diminishing returns as Charnov's patches did, but rather fluctuate payoff rates from period $t \rightarrow t+1$ via a stochastic process. As a result, the single armed bandit problem

[6] Once again, the first two features of foraging models are: (1) how long an animal searches for patches and (2) which patch types the animal visit.

models how to choose a slot machine that will maximize reward over an infinite future instead of modeling when an agent should leave a patch.

Since all of the slot machines have differing payoff rates, the gambler should occasionally probe each slot machine to gain information about the machine's payoff rate. Even though this may lead to a lower short-term expected payoff, the gambler gains information about payoff rates through exploration that will maximize the long-term payoff. Information now has a quantifiable value, as it can help the gambler choose the slot machine with a higher current payoff rate. Gittins showed that each slot machine should be assigned an index $v_i(x_i)$ that estimates the payoff rate from previous uses and the informational value from increasing the knowledge of the slot machine's payoff state (Gittins 1974). Each trial, an optimal gambler will choose the slot machine with the highest index.

The single-armed bandit problem and the Gittins index closely mimics the situation of many foraging animals. Some valuable resource patches are only available occasionally, such as ripe strawberries. Animals must devote resources and time towards discovering if these resources are available to maximize reward intake.

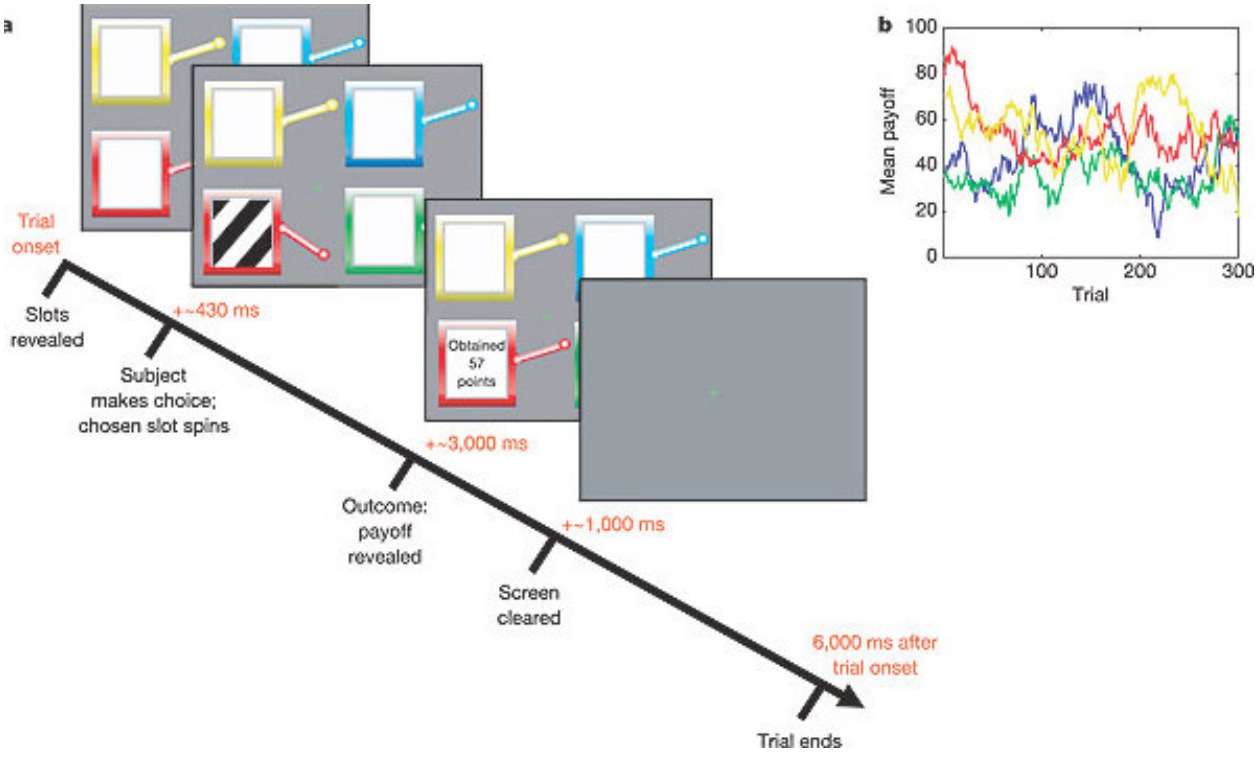
III. Neuroeconomic Findings in Foraging Theory

Neuroeconomics can expand Gittins' findings through studying the neural mechanisms of decision-making. Beyond discovering which slot machine an optimal agent chooses, neuroeconomics attempts to elucidate how humans make choices and ascertains why deviations occur from optimal behavior. Neuroeconomists predominately study how humans evaluate and obtain rewards, as well as create strategies to maximize reward intake (Doherty 2004). The human brain has an organized reward representation circuit to

estimate the value of a reward, predict future rewards, and use this information to guide behavior (Hyman 2006). While this system is not entirely understood, it utilizes several brain regions to constantly update and reevaluate reward representations based on current information (Samejima 2005).

In a foraging task, reward representations help agents evaluate and choose between exploitation and exploration. A recent finding by Daw et al. (2006) has greatly enhanced our understanding of the exploit-explore tradeoff by elucidating the neural mechanisms of these two actions. Daw et al. uses a functional Magnetic Resonance Imaging device to observe brain activation while subjects participate in a single-armed bandit task. Numerous brain regions involved in the reward representation circuit were activated during the task (Figure 1).

Figure 1: Task Design*



* From Daw et al., 2006

The experiment mimics the single-armed bandit problem described above. Initially, the subject chooses between four slot machines. Each slot machines awards points to the subject, which can later be redeemed for money. The slot machines pay off noisily around randomly changing means.

After the subject completed the task, Daw et al. used a modified version of the Gittins index to categorize each trial as either exploitation or exploration. The subject performed an exploitation action when he or she chose the slot machine with the highest perceived reward; the subject performed an exploration action when he or she chose a slot machine with a high informational value, but a lower expected reward. Then, Daw et al. examined differences in brain activation during exploitation and exploration. They found that several brain regions, each involved in the reward representation system, were active during exploration and not exploitation. Apparently, these brain regions suppressed a

natural tendency to exploit a resource, and led to exploration instead. This finding showed that the brain uses different regions to perform exploitation and exploration tasks. The mechanism for activating the different brain regions involved in exploitation and exploration, however, is unknown. Recent models suggest that NE may regulate the propensity to explore the brain through altering the functionality of the regions involved in exploitation and exploration.

IV. NE's Role in the Exploit-Explore Paradigm

NE is part of a class of brain chemicals called neuromodulators. These chemicals regulate the functionality of various brain regions.[7] For example, a neuromodulator can make particular brain regions more or less active during a given task.[8] The change in activation can increase task performance or inhibit actions. As a consequence, NE indirectly controls behavior through altering the effectiveness of different parts of the brain.

NE was traditionally thought to regulate arousal and attention (Aston-Jones and Cohen 2005). Neuroscientists posited that NE had simple, basic functions such as regulating alertness due to its broad and general connections to multiple brain regions. Indeed, neuronal recordings show that neurons release NE at high rates during walking, low rates during drowsiness, and virtually no NE during sleep (Aston-Jones and Bloom 1981). In contrast to these early hypotheses, recent findings show that NE may have a larger role regulating behavior.

---

[7] Actually, neuromodulators affect the functionality of neurotransmittors.
[8] Increased activity of a brain region generally corresponds to a greater role for that region in performing a task.

Through modifying alertness and arousal, NE helps to optimize behavior by increasing or decreasing the attention given to a task. Arousal is difficult to characterize with neurobiological mechanisms, but easy to define informally. Simply, arousal is alertness or the ability to pay attention to a task. Arousal is essential for performing even simple tasks. At low levels of arousal, individuals have difficulty functioning. Dampened arousal leads to drowsiness or, at the extreme, sleep. In the opposite side of the spectrum, heightened arousal can lead to distractibility. If an individual is interested in every loud noise or other stimulus, performing a task can be quite difficult. Individuals perform optimally at a happy medium between heightened and dampened arousal.

With connections to the reward representation circuit, the NE system's regulation of attention and arousal can affect reward related tasks. In the exploitation-exploration paradigm, NE may regulate whether an individual devotes attention towards exploiting a resource or abandons the resource and explores the environment. Low levels of arousal lead to torpor and poor task performance; medium levels of arousal correspond to exploitation; and high levels of arousal lead to distractibility and eventually exploration of the environment. Hence, the NE system provides a neural mechanism for switching between exploitation and exploration behaviors through regulating arousal.

Figure 2: Attention and Task Performance

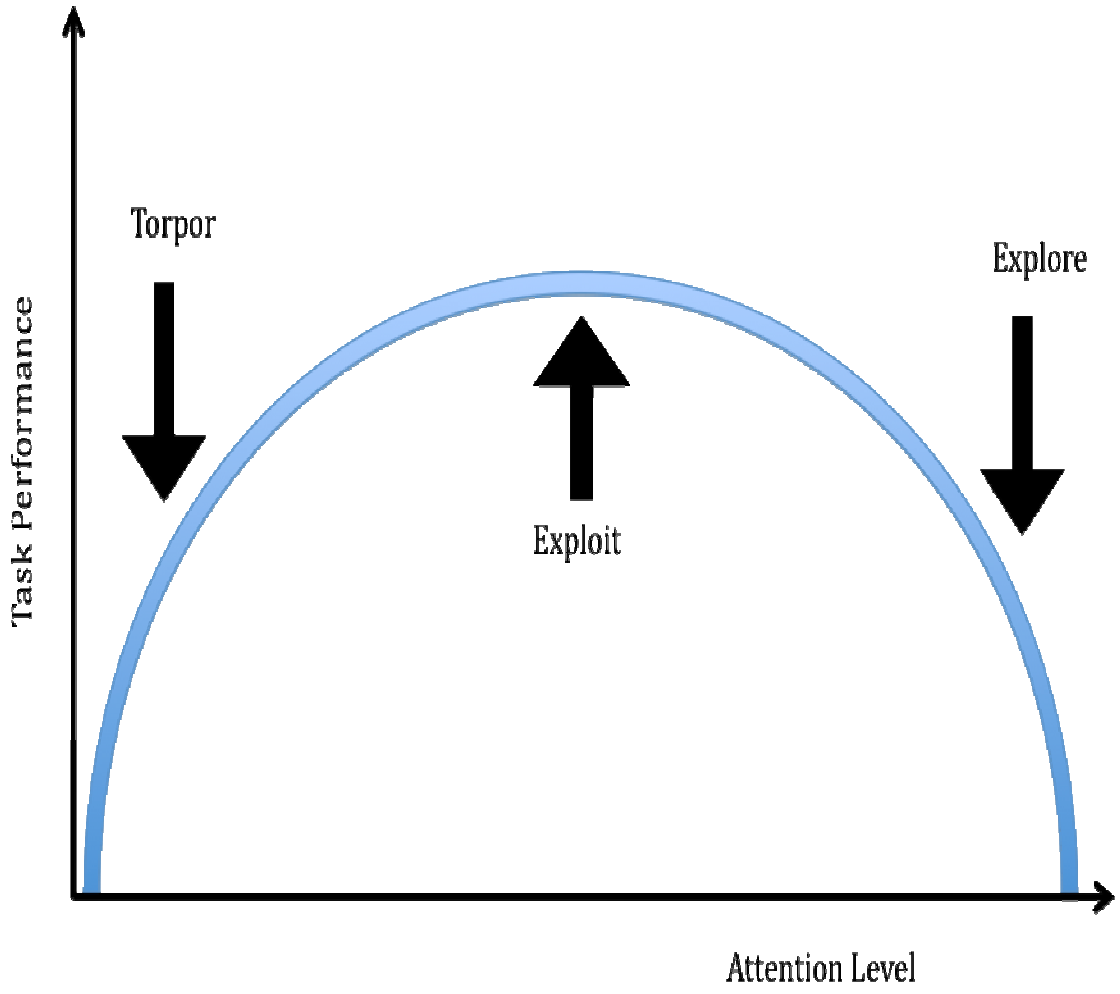


Figure 2: Adapted from Aston-Jones and Cohen (2005)

V. Methods and Details from the Experiment

Thus far, I have presented a model of NE functioning that may regulate the transition between exploitation and exploration. This model, however, is unconfirmed experimentally. Additionally, the current battery of mouse experiments lacks tests for the exploit-explore tradeoff. In this section, I describe an experiment created to investigate this tradeoff. Section VI then demonstrates that the mice can adequately perform this

exploit-explore task, and shows that the behavior of the NET mice deviates from the behavior of the normal mice.

5.1 The Principle Actors

I use three groups of mice in this experiment:

(1) A pilot group of normal, genetically identical mice[9,10]

(2) A second group of genetically identical mice, age-matched to mice in group (3)[11]

(3) A group of genetically altered NET (Norepinephrine Transporter knock out) mice

The mice in the first and third groups have normal gene expression and are referred to as wild type (WT) mice. Primarily, I used the first group of mice as a pilot group to develop the exploit-explore experiment. These mice participated in numerous unsuccessful exploit-explore experiments in addition to the final version of the experiment. Subsequently, groups (2) and (3) participated in the experiment. I then compared the two groups' results to determine how the NET mice deviate from WT mice in the exploit-explore task.[12,13]

5.2 Experiment Details

The WT and NET cohorts participated in a foraging experiment that emulates a natural foraging experience. As described earlier, the mice were individually placed each

---

[9] These are C-57 black mice.
[10] The mice in the first group are older than the mice in the other two groups. While age should not affect performance in this task, older mice do behave differently in some experiments.
[11] This eliminates any difference age may have on task performance.
[12] Norepinephrine transporter is a protein responsible for recycling NE after use (Xu et al. 2000; Hall et al., 2009; Perona et al., 2009). NET mice are genetically altered and lack this transporter. After NE is used to send a signal from one neuron to another, the neuron is slow to recoup lost NE efficiently.
[13] Knockout mice, like the NET mice, are born with a genetic deficiency. As the mice develop, alternate mechanisms develop to compensate for this deficiency. This makes extrapolating results obtained from knockout mice difficult since alternate mechanisms may cause odd results.

night in a small box with two portholes.[14] Each porthole, which represents a foraging patch, released liquid rewards to the mice. Since a mouse did not have free access to food or water while in the box, it obtained liquid through nosepoking into the portholes.[15]

The box was approximately 13 cm by 10 cm with one porthole on each the left and right end. This box size is large enough for a mouse to comfortably explore, but not too large that traversing the box is a hindrance. The portholes were approximately 2 cm by 2 cm boxes that protrude from the side of the boxes. At the end of the box, a liquid dispenser released small amounts of a liquid reward. Each porthole box has a laser motion detector that records when the mouse nosepokes into the porthole. Upon nosepoke, the liquid dispenser released the liquid reward for the mouse to collect. The liquid reward was a mixture of water and sweet'n'low artificial flavoring. See Figure 3 for a visual representation of the box and portholes.

[14] Each mouse spent twelve hours per day in the boxes. The two groups of mice lived under different light-dark cycles. When the lights were on for one group of mice (day), the lights were off for the other group (night). Mice are most active at night. This allows both groups of mice to spend the night period in the experimental box.

[15] A program called Med PC collected data on the mouse's nosepoke behavior from the box to analyze.

Figure 3: Experiment Box

Each porthole offers different reward rates during different time periods. This gives the mouse two different patches from which to forage. One patch, called the Fixed Interval (FI) patch offers a low reward value, *r*, at a constant rate. After the mouse nosepokes at the FI patch and receives a reward, the mouse must wait a constant delay period of Δ (5) seconds before receiving another reward for a nosepoke. In simpler terms, the FI patch offers a maximum reward rate of r reward every Δ seconds. The other patch is called the Variable Arrival (VA) patch. This patch is either active or inactive. When inactive, the patch offers zero reward per nosepoke. When active, the VA patch offers a large reward, *R*, when nosepoked, with $R > r$. There is no waiting time in-between nosepokes while the VA patch is active. Essentially, the active VA patch offers continuous, large rewards.[16] The VA patch becomes active via a Poisson process with an arrival rate λ. After the patch is active, the patch remains active for *S* (90) seconds before inactivating.[17]

[16] The mouse is constrained by the physical limitations of a maximum nosepoke rate. This rate is roughly two nosepokes per second.
[17] Note, the variables *R*, *r*, *S*, λ and Δ are constant and exogenous in the experiment.

Table 1: Summary of Exogenous Constant Variables and Terms

Variable or Term	Purpose
VA	Variable Arrival patch
FI	Fixed Interval Patch
R	Large Reward from VA patch
r	Small Reward from FI patch
Δ (5 seconds)	Delay period between rewards for FI nosepokes
λ	Poisson arrival rate for VA patch
S (90 seconds)	Duration of active VA patch

VI. Theoretical Section

To analyze this experiment, I first create a model of how a constrained-optimal mouse will behave. This is not an optimal agent, but rather a cognitively and physically limited mouse.[18] For example, a mouse is unable to continuously nosepoke, and is constrained physically by the maximum nosepoke rate of about two nosepokes per second. The mouse still performs optimally given reasonable constraints.

6.1 The Constrained-Optimal Mouse

Two situations exist for the optimally-constrained mouse: (1) the mouse does not know if the VA patch is active and (2) the mouse knows the VA patch is active.[19] For each of these situations:

(6.1A) The mouse can alter its overall nosepoke rate. If the mouse does not know if the VA patch is active, it nosepokes at some rate ν (nosepokes / second). If the mouse knows the VA patch is active, it nosepokes at a rate ν^* (nosepokes / second).

[18] For example, we will assume that our agent does not condition his nosepoking probabilities on information regarding patch turnoff time. An ideal agent would delay nosepoking immediately after the VA patch turns off.

[19] The mouse does not know if the VA patch is active until the mouse nosepokes at the VA patch and receives a reward. Likewise, the mouse knows the VA patch is active until the mouse nosepokes at the VA patch and is unrewarded.

(6.1B) The mouse nosepokes at the VA patch instead of the FI patch with some probability p, $0 < p < 1$. The probability p occurs when the mouse is unaware of the VA patch state; p^* corresponds to the mouse knowing the VA patch is active. Consequently, the mouse nosepokes at a rate $p\,\nu$ at the VA patch while its status is unknown, while poking at a rate $(1 - p)\,\nu$ at the FI patch.

(6.1C) The mouse nosepokes at a constant rate determined by a Poisson process for each of the rates ν and ν^*. These nosepoke rates are dependent on the leisure preferences for the mouse.[20]

Each situation offers different reward opportunities for the mouse. When the VA patch status is unknown, the mouse can choose to nosepoke at the FI patch and receive a constant, smaller reward. This represents an exploit behavior, as the mouse receives a reward from exploiting a known reward rate. While this would maximize present reward, the mouse can occasionally nosepoke the VA patch to gain information about the status of the VA patch. This information can lead to large future rewards if the VA patch is found active.[21] Nosepoking the VA patch represents an explore activity. It presents a direct cost to the mouse since the mouse receives no reward when the patch is inactive. Instead, the mouse could perform other activities, such as grooming, sleeping, or nosepoking the FI patch. In the second situation, the mouse knows that the VA patch is active. When active, the mouse can continuously nosepoke at the VA patch until it becomes inactive. The active VA patch offers a much higher reward than the FI patch offers, and without a delay period.

Since the mouse faces two different situations, a constrained-optimal mouse would vary its nosepoke rate according to its knowledge of the VA patch. While the VA patch status is unknown, the expected value of each nosepoke is low. A constrained-optimal

---

[20] Leisure activities are anything the mouse does beside nosepoking while in the experiment. For the mouse, the marginal nosepoke reward equals the marginal leisure reward.

[21] Recall, the VA patch offers a large reward without a delay, while the FI patch offers a smaller reward with a five second delay (Δ).

mouse would increase other leisure activities during his time, and nosepoke at a lower rate (ν), as nosepoking is less valuable. In contrast, while VA patch is active, each nosepoke has a high expected value. The mouse can maximize reward intake by nosepoking at a fast rate (ν^*) and reduce leisure activities.

Hypothesis 6.1 The mouse will increase nosepoke rate while the VA patch is active compared to inactive, $\nu^* > \nu$.

Similarly, a mouse will adjust p and p^* to maximize reward while the VA patch is active. The constrained-optimal mouse would exclusively nosepoke at the high value VA patch while it is active, and at a lower probability while the status is unknown.

Hypothesis 6.2 The mouse will increase the probability of nosepoking at the VA patch while the VA patch is active compared to inactive, $p^* > p$.

Completing these two behaviors show that the mouse recognizes the tradeoffs present in the experiment. Hypothesis 6.1 and 6.2 will be used later to check if the experimental WT mice can successfully complete the task. Neither of these hypotheses tests whether NET mice show more or less exploratory behavior. To do this, we need to examine how the constrained-optimal mouse maximizes reward intake.

6.2 The Exploit-Explore Tradeoff for the Constrained-Optimal Mouse

A constrained-optimal mouse creates a behavioral strategy where the reward benefit from exploiting the active VA patch equals the cost from exploring the VA patch while its status is unknown. The mouse adjusts the probability p to balance these rewards and costs. Specifically, this probability is dependent on the expected value of the VA patch and the FI patch. When the mouse nosepokes the VA patch with an unknown status, the expected value of the reward is the probability the patch is active multiplied by the overall reward, or:

$$(5.1) \quad EV[VA_{\text{unknown status}}] = R\,S / (S + 1 / \lambda).$$

For the FI patch, the expected value depends on the nosepoke rate. If the mouse nosepokes at a rate faster than once per Δ seconds, the mouse would receive a maximum reward of r / Δ. The delay period Δ seconds encourages the mouse to nosepoke at a slower rate. The constrained-optimal mouse would nosepoke at a rate equal to or slower than $1 / \Delta$, and get a reward r for every nosepoke.

Hypothesis 6.3 A mouse will attempt to nosepoke at the FI patch at a rate equal to or slower than $1 / \Delta$ nosepokes per second.

Although the mouse is nosepoking the FI patch at a slower rate, the expected value of the FI patch is larger than equation 5.1.[22] A mouse exclusively nosepoking at the FI patch would receive the expected reward per nosepoke:

$$(5.2) \quad EV[FI] = r\,\nu.$$

Although Equation 5.2 is larger than equation 5.1, nosepoking the VA patch can lead to a larger future reward. To compute this additional reward, I compare the loss from allocating nosepoking to the inactive VA patch with the gain in profit from nosepoking the active VA patch. The loss from nosepoking the inactive VA patch is the VA nosepoke rate multiplied by the lost FI reward and the average time nosepoking the VA side before the mouse discovers an active VA patch. This equates to:

$$(5.3) \quad p\,r\,\nu / (2\lambda).$$

The gain in profit from nosepoking the active VA patch is the total reward from nosepoking the VA patch minus the alternative, or nosepoking the FI patch. Both total reward values depend on the average amount of time spent nosepoking the active VA side once it is

[22] The experiment sets the parameters S, λ, R, and r to ensure this is true.

discovered active. Since the mouse randomly nosepokes, the amount of time left in an active VA patch after discovery is $S / 2$.[23] Using this result, the total gain from an active VA patch is:

$$(5.4) \quad (S / 2) (\nu^*) R$$

The overall gain in profit is the gain from the active VA patch minus the expected value of the FI patch nosepoked over the same time duration and nosepoke rate, or:

$$(5.5) \quad (S / 2) * (\nu^* R - r \nu)$$

For an optimally performing mouse, equation 5.5 and 5.3 should be equal. Solving for p:

$$(5.6) \quad p = \lambda S / (\nu r) * (\nu^* R - r \nu) = \lambda S * [(\nu^* / \nu) * (R / r) - 1)]$$

Equation 5.6 describes the behavior for a constrained-optimal mouse, and leads to several conclusions. First, the probability *p* for devoting nosepokes to the VA patch is proportional to the ratios (R / r), (ν^* / ν), and $S / (1 / \lambda)$. The reward ratio (R / r) and ratio of the time duration of the active VA patch to inactive VA patch [$S / (1 / \lambda)$] affect the nosepoke rates by altering the respective values of the VA and FI patches. Both of these ratios are determined by the conditions of the experiment, and are independent of the mouse's actions.

Second, the value of the ratio ν^* / ν and *p* show whether a mouse is engaging in more exploit or explore behaviors. When comparing the NET and WT mice, a larger increase in ν^* / ν corresponds to more exploit behavior. Altering ν^* / ν indicates the mouse is more efficient at exploiting the valuable active VA patch. Likewise, a low value of

[23] EV[time left | VA active] = S / 2 since the mouse will, on average, discover the active VA patch in the middle of the period S.

p indicates high exploitation behavior while the VA patch is inactive, as the mouse nosepokes at the FI patch more frequently.

Hypothesis 6.4 A high ν* / ν and a low *p* correspond to exploitation behavior, while a high *p* corresponds to exploration behavior.

Lastly, equation 5.6 shows that the ratio of mouse nosepoking rates (ν* / ν) is positively related to probability of nosepoking the VA patch *p*. Mice with a high ν* / ν are better able to exploit an active VA patch, and have a larger expected future reward from discovering it. Therefore, mice with a higher ν* / ν should spend more time exploring the VA patch (*p*) while the status is unknown to reap this large reward.

Hypothesis 6.5 The ability to efficiently exploit the active VA patch (ν* / ν) should lead to more exploring activity while the VA patch status is unknown (high *p*). A mouse with a low ν* / ν should have a lower *p*.

In summary, Hypothesis 6.1 and 6.2 confirm that the mice understand the task. Hypotheses 6.3, 6.4 and 6.5 compare the exploit and explore behaviors of the NET and WT mice.

VII. Results

Recall that three groups were used in this experiment.[24] The first group was the older WT mice. There are 10 of these mice, and each participated in the task for 16 days.[25] I used this data to show that mice are capable of understanding the exploit-explore task. In addition, the mice underwent training prior to the experiment. The training acclimated the mice to the experiment chambers, trained the mice to nosepoke the portholes for a liquid

[24] All the mice are numbered. See Appendix B for the mouse numbers, groups, and genotypes.

[25] Typically, the mice nosepoke about 1000 times per night. This is a very large amount of data for a mouse experiment.

reward, and taught the mice that each porthole offers a different reward. The second and third groups of mice are age-matched WT and NET mice, respectively. I compared these two groups to find differences in exploit and exploration behaviors. There are four mice in each of the two groups, and the mice participated in the experiment for 6 days after training.[26]

7.1 Do the Mice Understand the Task?

This section shows that mice can understand and complete the experiment. I examine the data for each of the ten older WT mice over 16 days. I loot at 6.1A, B, and C. Also, I show that the mice follow Hypotheses 6.1 and 6.2. In addition, I check if the mice show learning over the course of the experiment. This verifies that the training was adequate for the experiment. Section 7.1 is divided into three sections that each address one of the conditions mentioned above:

(1) Do the mice nosepoke at a constant rate determined by a Poisson process? This addresses 6.1C.

(2) Are the ratios ν^* / ν and p^* / p positive and greater than 1? This shows that the mice successfully exploit the active VA patch, satisfying 6.1A and 6.1B and also Hypotheses 6.1 and 6.2.

(3) Do the mice exhibit learning behavior across days?

[26] The younger mice nosepoke at a slightly lower rate: about 600 – 700 nosepokes per night. As mentioned earlier, younger mice perform slightly differently in certain experiments than older mice. Younger mice are more timid in experiments that older ones. The age of the mice should have little affect on the decision to exploit or explore. Most of my results are either percentages or ratios, making the absolute number of nosepokes inconsequential.

After addressing all three questions, I determine if each individual mouse can complete the experiment. Mice that fail to complete the experiment are removed from the data set. This is a common scientific practice. Since all of the mice are genetically identical, performance differences arise from a failure to comprehend the task, rather than differences in cognitive abilities.

Question 1

To determine if a Poisson process determines the mouse nosepoke rate, I calculated the time interval between nosepokes and compared these to an exponential distribution.[27] I then performed a goodness-of-fit χ^2 calculation. For nosepoking when the VA patch is inactive and active (ν and ν^*), the inter-nosepoke interval fails to follow an exponential distribution ($p = 0.999$ for both). This violates 6.1C. Each nosepoke rate, however, has a distinct peak in the inter-nosepoke interval histogram that deviated from the exponential distribution. These peaks occur for different reasons related to the task parameters, and help show that the mouse understands the task. See below for sample distributions from one mouse (Figure 4 and 5).

[27] An exponential distribution will describe the time intervals between two events for a Poisson process.

Figure 4

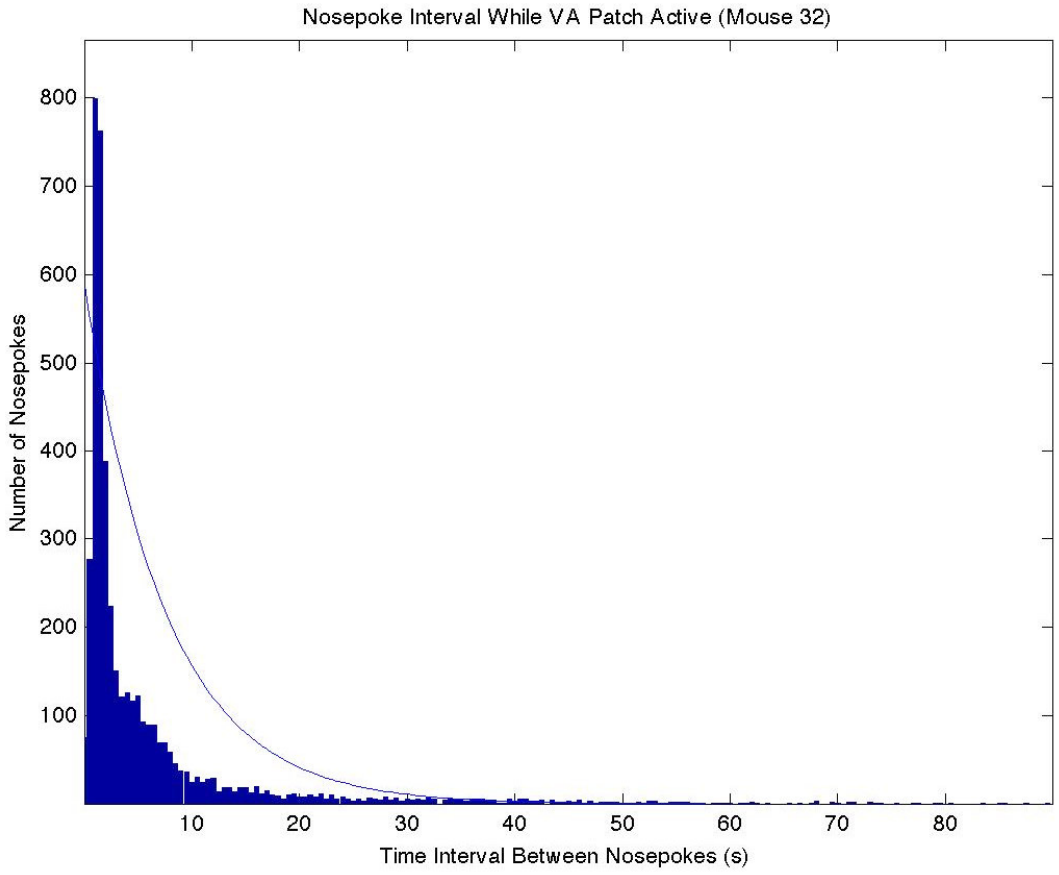


Figure 4: The dark blue bars represent nosepoke intervals for WT mouse 32, while the light blue line is an exponential distribution. All nosepoke intervals greater than 90 seconds were discarded. Since the mouse spent roughly twelve hours each night in the experiment box, the mouse occasionally fell asleep or ignored the nosepoke boxes for extended period of times. These large times do not show exploit-explore preferences.[28]

Visually, the mouse behaves significantly different than a Poisson-determined nosepoke rate would suggest while the VA patch is active. A sharp peak occurs close to the one second inter-nosepoke interval. When examining the task, the mouse has an incentive to nosepoke quickly at the VA patch while it is active. This will cause a short inter-

[28] This mouse was chosen as an example because it shows the most pronounced effects.

nosepoke interval, explaining this deviation. The peak shows that the mouse understands the task because the mouse nosepokes as quickly as possible when it recognizes that the VA patch is active.

Figure 5

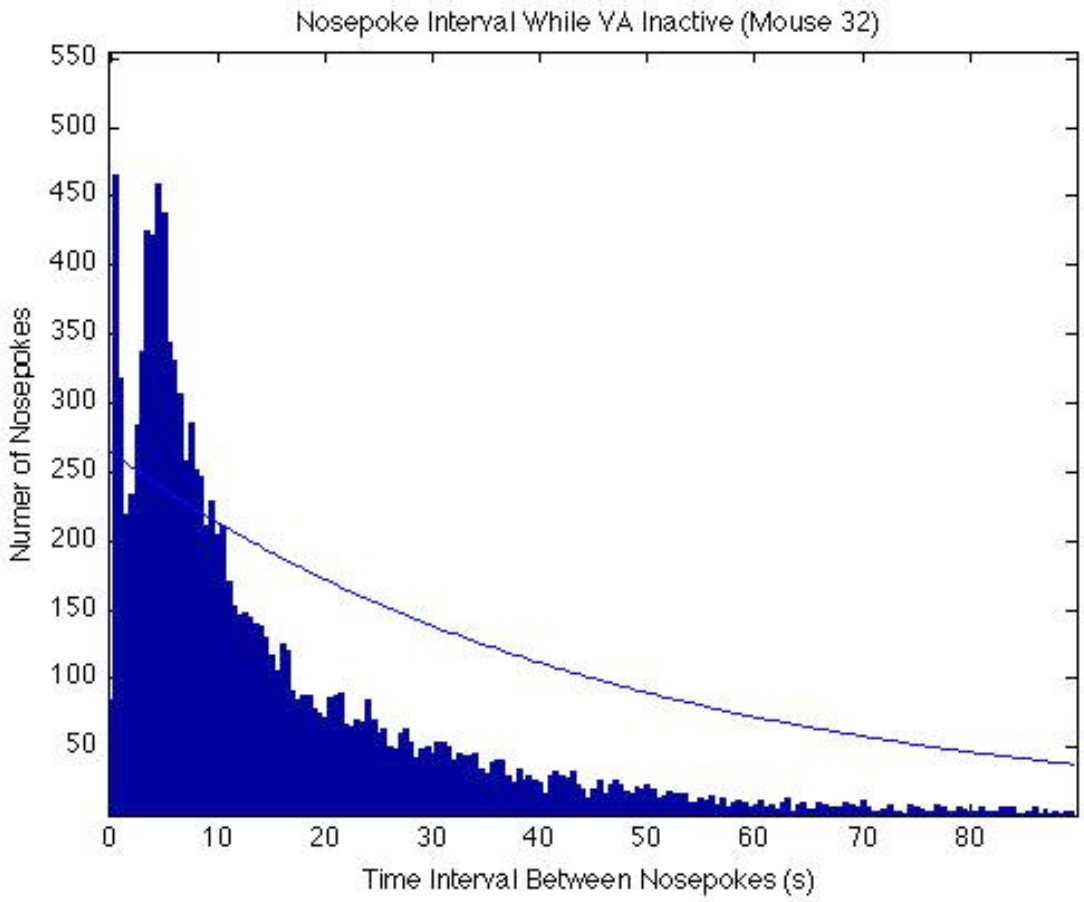


Figure 5: See legend of Figure 4 for a description.

In Figure 5, a large peak occurs around the five second nosepoke interval period. Recall that the FI patch has a five second delay period, Δ. Since the peak occurs during this five second interval, the mice learn to time their nosepokes at the FI side to obtain a maximum reward rate (Hypothesis 6.3). Note, this second peak disappears while the VA

patch is active (Figure 4). The mouse only times nosepokes when the VA patch is inactive, and the mouse is nosepoking predominately at the FI side.[29]

While the mouse violates the assumptions of a Poisson distribution for the nosepokes, the tails of the nosepoke time intervals appear to follow a Poisson distribution. Neither quick nosepokes in succession or timing nosepokes five seconds apart should affect the distribution of nosepoke intervals from the ten second period onwards. Figure 6 shows this nosepoke data.

[29] The peak close to the one second interval period still exists. This occurs because mice have a tendency to nosepoke in quick succession. The peak while the VA patch is active is much larger than the peak near one second while the VA patch is inactive. Considering the VA patch is inactive for the majority of the night, this shows that the peak during the VA active period is from the mouse adjusting its nosepoking strategy rather than just nosepoking in quick succession.

Figure 6

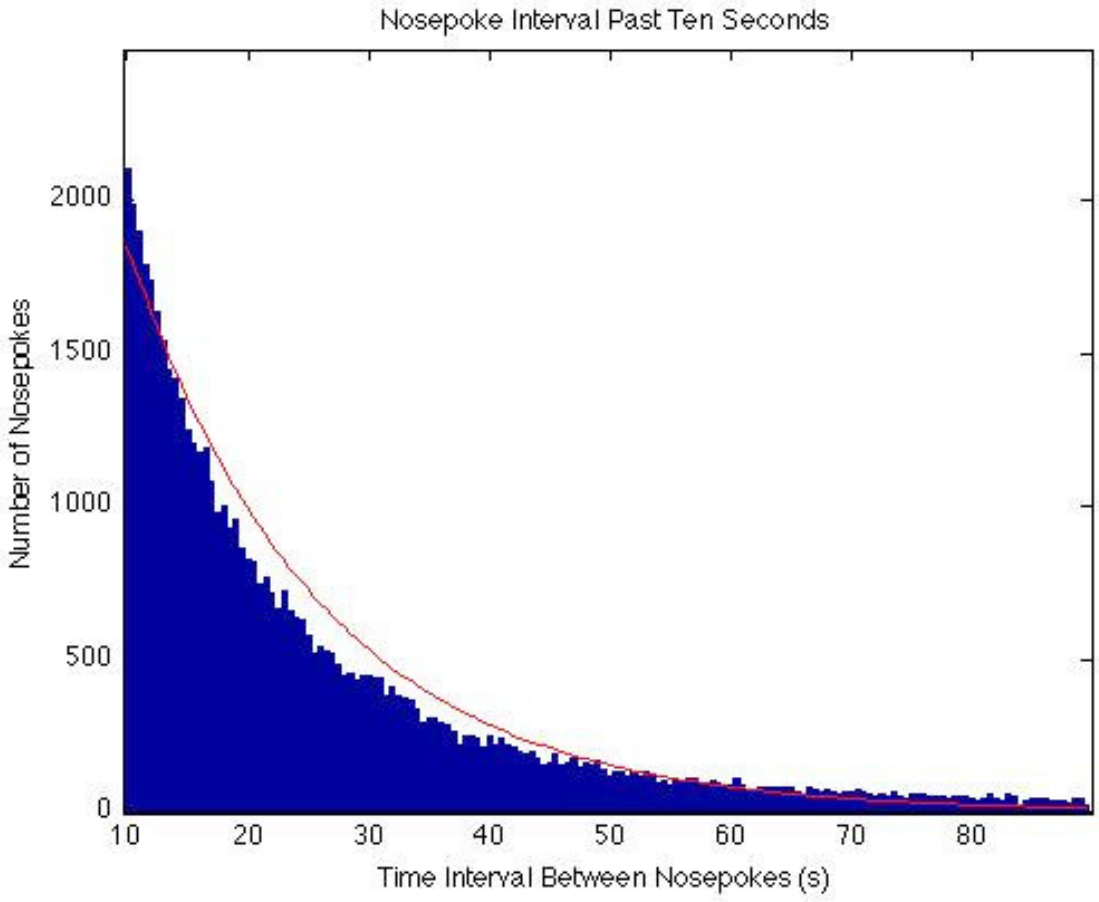


Figure 6: I removed all the nosepokes intervals from 0 to ten seconds. Then, I recalculated lambda for the exponential distribution and plotted it.

The data is visually much closer to an exponential distribution. Still, the goodness-of-fit p-value is large and insignificant. The small inter-nosepoke interval bins used in the histogram introduce a large variance, and could explain this failure.

Even though the mice fail to nosepoke according to a Poisson process, the other hypothesis and assumptions remain valid. The tradeoffs presented in the model proposed in the previous section may be altered, but the intuitions about mouse behavior still hold.

A new, more accurate model should create a new method for determining mouse nosepoke behavior.

Question 2

Even though 6.1C failed to hold, the mice show an ability to perform the task well. The mice nosepoke at rapidly after the VA patch is turned on, and time nosepokes to maximize the reward rate at the FI patch. To quantitatively show that the mice understand the task, I show that the mice alter nosepoke rates (ν) and the probability of nosepoking the VA patch (p) when the status of the VA patch changes. Likewise, Hypotheses 6.1 and 6.2 predict that, if the mouse understands the task, the mouse will have a ν^* / ν and p^* / p ratios greater than one.

Table 2 records the nosepoke rates for when the VA patch is active (ν) and inactive (ν^*).

Table 2: Wild Type Mouse Nosepoke Rates

Genotype	Mouse Number	Number of Observations	Nosepoke Rate While VA Patch Active (nosepokes / second)	Nosepoke Rate While VA Patch Inactive (nosepokes / second)	P-value
WT	All	10	0.0837	0.0193**	0.0019
WT	26	16	0.0799	0.0303**	0.0173
WT	27	16	0.0811	0.0137**	0.0011
WT	28	16	0.0648	0.0212**	0.0006
WT	29	16	0.0990	0.0100**	0.0013
WT	30	16	0.0639	0.0123**	0.0019
WT	31	16	0.0761	0.0166**	0.0004
WT	32	16	0.0986	0.0214**	0.0004
WT	33	16	0.1075	0.0306**	0.0007
WT	34	16	0.0823	0.0191**	0.0013
WT	35	16	0.0838	0.0183**	0.0006

Table 2: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance. I performed a Wilcoxon signed-rank test to determine the p-value in the table. The Wilcoxon signed-rank test is a non-parametric hypothesis test for repeated measurements on a single sample.[30] To generate the data, I created an average nosepoke rate while the VA patch is on and off for each mouse on each of the 16 experimental days. Then, I ran the Wilcoxon signed-rank test for each mouse individually. For all the mice, I created an overall nosepoke average across all days. I used the Wilcoxon signed-rank test again to compare all of the mice nosepoke rates and record a p-value in the "All" row.

The table shows that all mice significantly increased the nosepoke rate while the VA patch was active.

While results from the mice are significant, the data fails to account for times when the mice sleep or are otherwise inactive. The mice spend nearly twelve hours in the experiment boxes. During this time, the mice spend long periods sleeping or performing other leisure activities instead of nosepoking. I dropped all time periods longer than five minutes without a nosepoke. The inactive periods give no information about exploitation

[30] The Wilcoxon signed rank test is the non-parametric version of the paired student t-test.

and exploration preferences. Table 3 shows the data with sleeping periods removed. For all data reported from this point forward, sleeping periods are removed.

Table 3: WT Mouse Nosepoke Rates without Sleeping Times

Genotype	Mouse	Number of Observations	Nosepoke Rate While VA Patch Active (nosepokes / second)	Nosepoke Rate While VA Patch Inactive (nosepokes / second)	P-value
WT	All	10	0.1341	0.0531	0.002***
WT	26	16	0.1255	0.0563	0.000***
WT	27	16	0.1422	0.0381	0.001***
WT	28	16	0.1035	0.0592	0.001***
WT	29	16	0.1620	0.0436	0.000***
WT	30	16	0.1478	0.0501	0.000***
WT	31	16	0.1310	0.0523	0.000***
WT	32	16	0.1447	0.0552	0.000***
WT	33	16	0.1469	0.0833	0.000***
WT	34	16	0.1165	0.0510	0.000***
WT	35	16	0.1217	0.0423	0.000***

Table 3: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance. I followed the same methods as described in the Table 2 legend.

Table 3 confirms that Hypothesis 6.1 is correct. The mice successfully alter their nosepoke rates when the VA patch is active.

Next, I tested Hypothesis 6.2 by comparing the probability of nosepoking at the VA patch while active (p) and inactive (p^*). The table shows that six of the ten mice can alter nosepoking rates.

Table 4: Probability of Nosepoking at the Active VA Patch

Genotype	Mouse	Number of Observations	Nosepoke Probability While VA Patch Active (p^*)	Nosepoke Probability While VA Patch Inactive (p)	P-value
WT	All	10	0.5902	0.4387	0.0028**
WT	26	16	0.6420	0.3840	0.0386**
WT	27	16	0.5216	0.6987	0.0979*
WT	28	16	0.6596	0.4607	0.0879*
WT	29	16	0.6502	0.3566	0.0261**
WT	30	16	0.5737	0.2596	0.0494**
WT	31	16	0.5509	0.4898	0.3519
WT	32	16	0.5625	0.4225	0.4691
WT	33	16	0.6515	0.3626	0.0071***
WT	34	16	0.5292	0.4643	0.7173
WT	35	16	0.5617	0.4886	0.4691

Table 4: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance.

Overall, the mice performed well in the exploit-explore task. All ten of the mice changed the nosepoke rate according the status of the VA patch, while six out of ten mice altered nosepoking probabilities at the VA patch. The four mice that failed to alter nosepoke probabilities at the VA patch would be removed in future data sets. The attrition of four mice is higher than most mouse tasks, but acceptable considering the complexity of this task compared to other mouse tasks.[31]

Question 3

The last question concerns whether the mice show learning behavior over the course of the experiment. In other words, I am checking if the session effect is significant. Throughout the sixteen experimental days, the mice can show a session effect through

[31] In most mouse tasks, about one or two mice out of thirty are removed from the data set. The exploit-explore task, however, is significantly more complicated than the average mouse task. In other, comparably difficult tasks, similar attrition rates are common.

either changing nosepoking rates (ν and ν*) or changing the probability of nosepoking at the VA patch (*p* or *p**). The mice showed no indication of a session effect by changing nosepoking rates while the VA patch was active or inactive (Table 5). The mice did, however, show a session effect through changing the probability of nosepoking the VA patch while active and inactive.

Table 5: Learning Across Days

Type of Learning	Coefficient of Correlation	P-value	Number of Statistically Significant Mice
Nosepoking while VA Patch Inactive (ν)	0.0457	0.5660	0
Nosepoking while VA Patch Active (ν*)	0.0320	0.5436	0
Probability of Nosepoking the VA Patch While Inactive (*p*)	-0.1470	0.0637*	3
Probability of Nosepoking the VA Patch While Inactive (*p**)	0.2152	0.0063***	3

Table 5: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance

While the mice as a group showed indications of a session effect across trials, most of the data from individual mice are statistically insignificant. Mouse 35 was the only mouse that showed a session effect across experiment days for both nosepoke probabilities (*p* and *p**).[32] Despite this, the session effect had a small affect on the data, and can reasonably be ignored. If the session effect has any affect, it would skew the data towards

[32] Mouse 35 failed to nosepoke the VA patch with different probabilities (Table 4), and is discarded from the data set.

showing that the mouse failed to complete the task. The training sessions generally succeeded.[33]

7.2 The Norepinephrine Transporter Knockouts and Wild Type Mice

Section 7.1 established that mice are capable of performing the exploit-explore experiment. The mice alter behavior to successfully exploit the VA patch, and the session effect is small and generally insignificant. After establishing that the experiment is viable, I performed the experiment again with age-matched NET and WT groups of mice. This is a partial data set, as I am continuing to collect data. For this data set, eight total mice performed the experiment for six days. I explored all of the questions from Section 7.1, and the results are summarized in Table 7.[34] Since the mice ran for a shorter number of days, many of the mice show statistical trends rather than statistical significance.

[33] When the mice did show learning behavior, the learning showed improvements in task performance. The mice decreased the probability of nosepoking at the VA patch while inactive, indicating that the mice exploited the FI patch. Likewise, the mice increased the probability of nosepoking at the VA patch while active, indicating that the mice exploited the active VA patch. Future experimenters should either increase training, or use the first few experimental days as training to eliminate the session effect.

[34] Appendix C shows the results.

Table 7: NET and WT Mice

Genotype	Mouse	Alters Nosepoke Rate (yes / no)[35]	Alters Probability of Nosepoking the VA Patch (yes / no)[36]	Understands the Task (yes / no)
NET	2575	**Yes**	No	No
NET	2554	No	**Yes**	No
NET	2553	**Yes**	No	No
NET	2552	**Yes**	**Yes**	**Yes**
WT	2577	No	**Yes**	No
WT	2547	**Yes**	**Yes**	**Yes**
WT	2574	No	No	No
WT	2573	**Yes**	No	No

Table 7: Bold indicates a yes answer. Only two mice met both criterion: mouse 2552 (NET) and 2547 (WT). Mouse 2554 (NET) and 2577 (WT) were close, and will be included in some analyses.

7.3 Results from Experiment

The NET and WT mice showed differences in exploitation and exploration behaviors. Compared to the WT mice, the NET mice demonstrated an increased tendency for exploitation, and diminished amounts of exploration. The mice demonstrated this tendency in two ways:

(1) The NET mice had a lower probability of nosepoking the VA patch while it was inactive than the WT mice did (p-value: 0.075). While the VA patch is inactive, only the FI side offers a reward. Nosepoking at the FI side at a high rate indicates more exploitation, and an unwillingness to explore the VA patch (Hypothesis 6.4) (Figure 7).

[35] Hypothesis 6.1
[36] Hypothesis 6.2

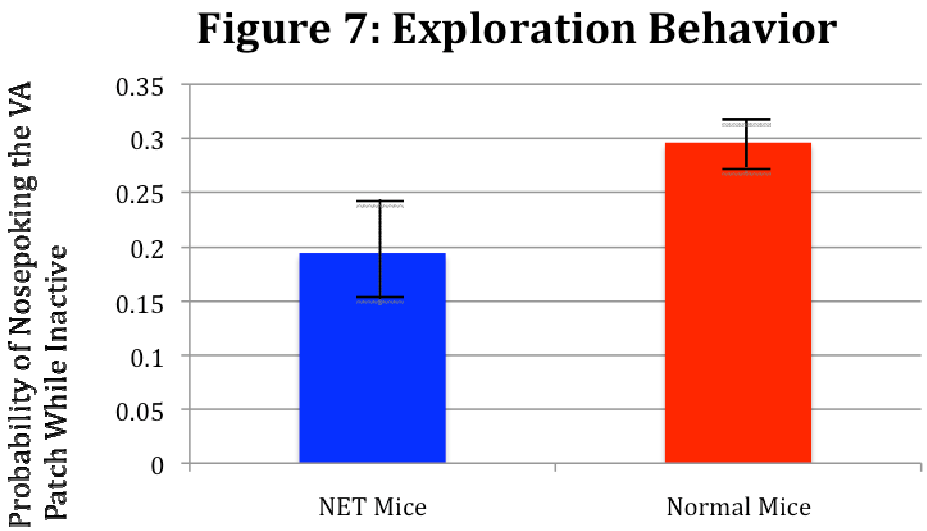


Figure 7: This is the average probability of nosepoking the VA patch while inactive. Only the mice that successfully completed the task were used in this graph: 2552 (NET), 2554 (NET), 2547 (WT), and 2577 (WT). This is significant to 10% (p-value: 0.075).

(1) Compared to the WT mice, the NET mice increased the difference of the nosepoking rate of the active VA patch and the inactive VA patch (p-value: 0.094[37]). Increasing the difference ($\nu^* - \nu$) demonstrates that the NET mice were more successful at adjusting behavior to exploit the active VA patch (Hypothesis 6.4) (Figure 8).

[37] For this p-value, the mouse nosepoke rates were normalized to account for differences in the absolute nosepoke rates.

Figure 8: Behavioral Differences Between NET and WT Mice

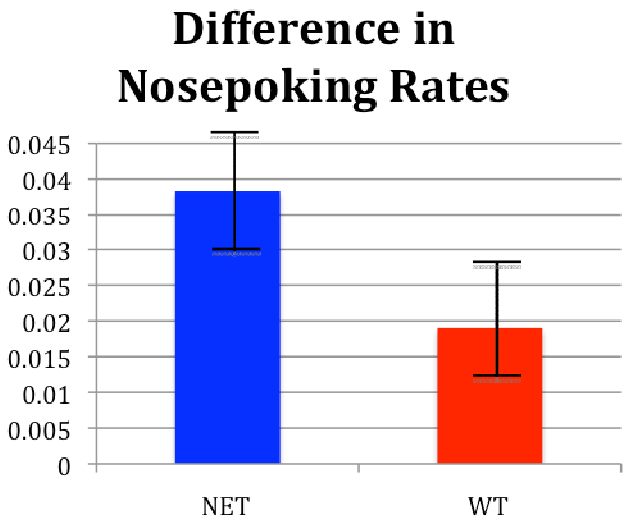


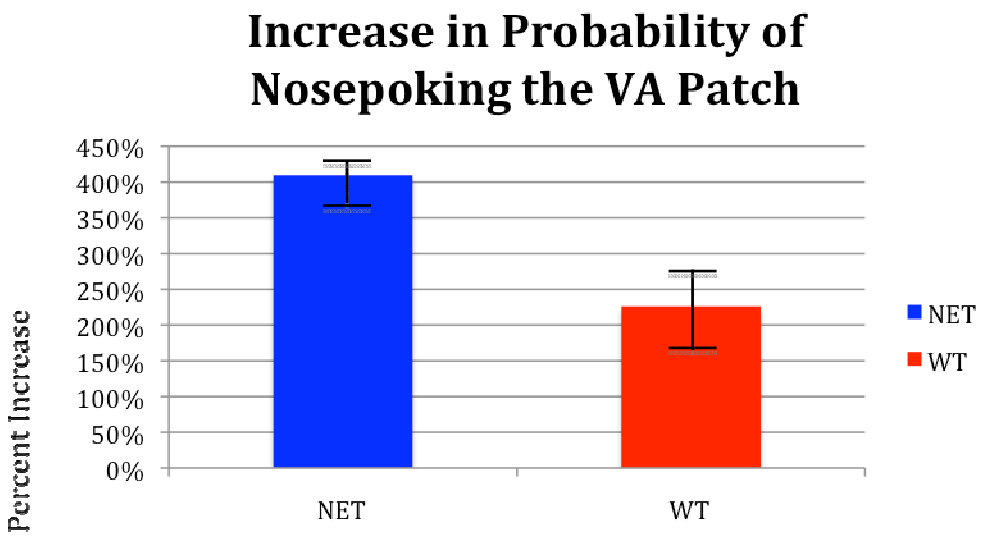
Figure 7: This is the difference in nosepoking rates from the active VA patch to the inactive VA patch. Only the mice that successfully completed the task were used in this graph: 2552 (NET), 2553 (NET), 2575(NET), 2547 (WT), and 2577 (WT). This is statistically insignificant (p-value: .400), mainly because the nosepoke rates are not normalized. When the nosepoke rates are normalized, the values are significant to 10% (p-value: 0.094).

While Figure 7 and 8 indicate that NET mice exhibited more exploitation behaviors than the WT mice, the WT mice had a larger probability of nosepoking the active VA patch than the NET mice (WT: 71.4%, NET: 66.4%; p-value: 0.667). Changing the probability of nosepoking the VA patch appears to be a more difficult task for the mice.[38] The probability of nosepoking the VA patch while active may depend on the baseline probability of nosepoking the inactive VA patch. In other words, p and p^* may be related. To test this, I compared the percent increase of the probability of nosepoking the VA patch for the NET and WT mice (Figure 9).[39] The NET mice increased the relative probability of nosepoking

[38] Indeed, recall that four of ten WT mice from group 1 failed to change nosepoke probabilities, while all ten changed nosepoke rates from the VA patch active to inactive.
[39] This is $(p^* - p) / p$.

the VA patch more when compared with WT mice (p-value: 0.049). This suggests that the NET mice are indeed better at altering their probability of nosepoking at the VA patch.

Figure 9: Percent Increase in the Probability of Nosepoking the VA Patch



In this section, I showed that the NET mice exhibit a tendency towards exploitation over exploration. The NET mice increased nosepoking rates significantly while the VA patch is active, nosepoked predominately at the FI patch while the VA patch is inactive, and increased the relative probability of nosepoking the VA patch. This shows that NE has an effect regulating the exploit-explore tradeoff.

VII. Conclusion

In my thesis, I investigated the role of NE in the exploit-explore tradeoff. Previous research in optimal foraging theory provided an exploit-explore model for animal behavior. This model, however, failed to properly describe animal and human behavior. The models required agents to make complex calculations that are unfeasible given both time and cognitive constraints. New fields such as neuroeconomics have reinvestigated the exploit-explore tradeoff by examining the neural mechanisms of decision-making. Through

regulating arousal and attention, NE provides a model for transitioning between exploitation and exploration.

In my thesis, I completed two tasks. First, I developed an exploit-explore task that mice can successfully complete. The mice can choose to nosepoke at either the FI or the VA patch. The FI patch offers a constant, but small reward, while the VA patch offers an unpredictable, but high reward. Mice successfully increase the nosepoke rate when the VA patch is active, and increase the probability of nosepoking at the VA patch while active. Both behaviors indicate that the mice can alter behavior to exploit the valuable VA patch. The task also provides an opportunity to measure the relative amounts of exploitation and exploration between two different groups of mice. A high ratio of ν^* / ν and a low p value indicate that the mice exhibit a tendency towards exploitation, while the reverse corresponds to exploration. Future researchers can use this task with other groups of genetically altered mice to examine the exploit-explore tradeoff.

Second, I determined that NE helps regulate the exploit-explore tradeoff. Mice with deficiencies in NE functioning predominately performed exploitation rather than exploration. Recall from Section III that two different brain regions are responsible for exploitation and exploration behavior. The region that controls exploration appears to suppress a natural tendency for exploitation. In the experiment, the NET mice may be unable to properly activate these two brain regions, and thus are unable to transition from exploitation to exploration. NE may affect this transition by changing the level of arousal. Increasing arousal leads to distractibility, and causes an increase in exploration. From this, I hypothesize that mice with deficient NE functioning are unable to properly increase

arousal during the exploit-explore task and engage in exploration. As a result, the NET mice effectively remain in an exploitation mode.

Although this experiment determined that NE is involved in the exploit-explore tradeoff, I am unable to make a definitive conclusion about the mechanism about NE regulation. Many mice experiments similar to mine can only provide broad statements about the involvement of neurobiological systems in a task. Future research should focus on discovering the mechanism for exploitation and exploration at the cellular level. This will give larger insights into the decision-making mechanism.

Still, my experiment has important implications for economists. The explore-exploit tradeoff is found in numerous economic problems and real world situations. For example, investors face an exploit-explore tradeoff when deciding whether to invest in a well-known company or a newer company with an unknown performance profile. A greater understanding of the mechanisms of the exploit-explore decision will allow economists to create more accurate models. Additionally, my results will allow economists to explain systematic deviations from optimal behavior due to genetic differences between people. Certain individuals may have lower levels of NE functioning, and may deviate from optimal behavior in a systematic way. Future research should extend my findings to human populations, and incorporate the effects of altered NE function in economic models.

Appendix A: Basic Introduction to Neuroscience

This section serves as a basic introduction to the necessary neuroscience to understand the concepts in this paper. Readers familiar with basic neuroscience may feel free to skip this appendix. The neuron, the basic cell found in the brain, has three major parts: the cell body, the axon, and the dendrite. The cell body performs normal cellular functions necessary to maintain the cell. The axon and dendrite are long wire-like projections from the cell that give and receive, respectably, information from other cells. The information transmitted is electrical impulses. Neurons interconnect in vast networks to process information. This is analogous to a computer, and allows the brain to perform complex functions.

Remember above when we discussed brain regions. While we only generally mentioned brain regions, a brain region is a collection of neurons. These neurons are connected to other brain regions that process and receive other information. For example, when the region of the brain that processes visual information locates a piece of food, it sends that information to the reward representation region.[40] This region then integrates the information and, if the person is hungry, decides to eat the food. The reward representation region then sends this decision to the motor region, which then performs the action.

Now that we understand the basics of brain functioning, we will learn how the brain initiates the electrical impulses that send information. The axon of one neuron sends information to the dendrite from another neuron. Then, the information is propagated

[40] This is a hypothetical and very simplified example. The example does, however, get across the major points necessary to understand how the brain works.

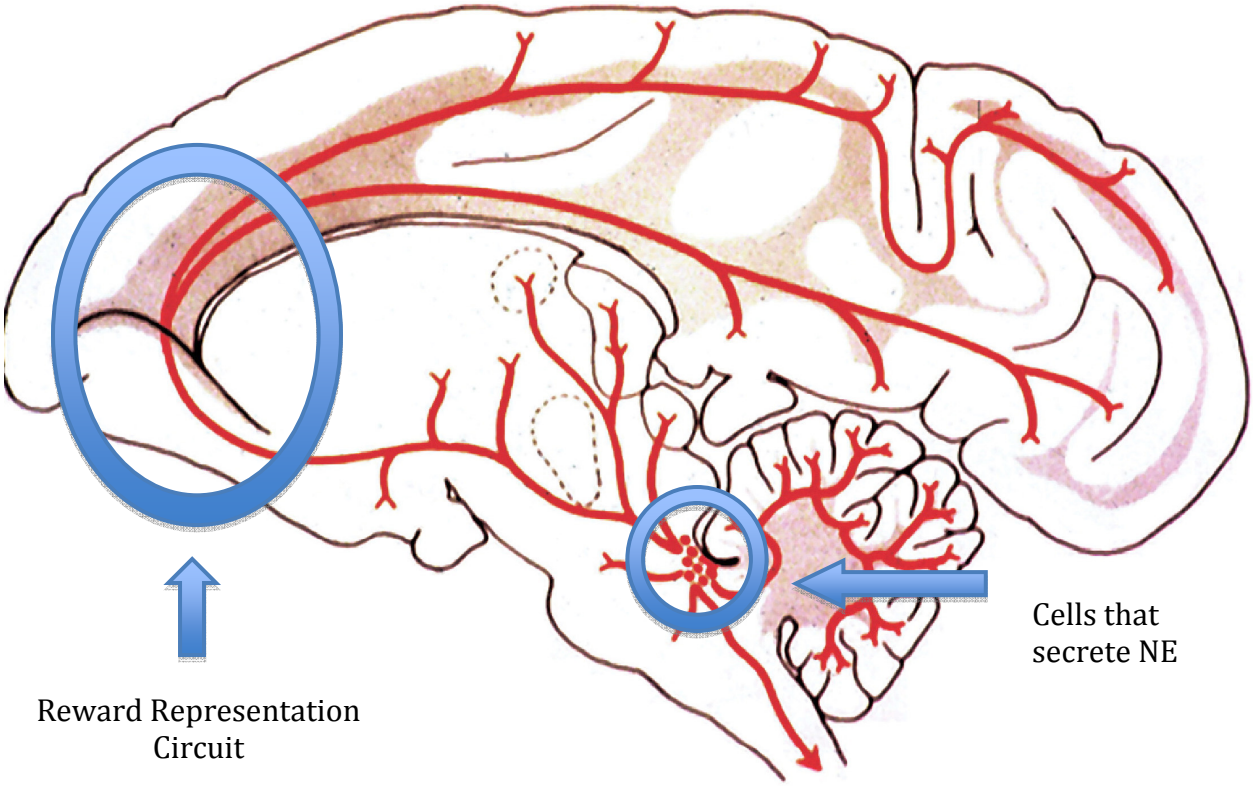
down the neuron body and to the axon to repeat this process. Similar to an electrical wire, neurons connect in long chains to transfer information across the brain. In-between each axon and dendrite combination is a small empty space called the synaptic cleft. This is responsible for regulating, initiating, and ending electrical impulses. In the synaptic cleft, neurons release chemicals called neurotransmitters that initiate electrical impulses in the next neuron. Once released, these neurotransmitters are recycled and returned to the original neuron cell. The neurotransmitters can be released again and again to initiate the electrical impulses, which are called action potentials. NE, a nueromodulator, affects neurotransmitters and their ability to elicit electrical signals. While this process remains unclear, NE could raise or lower the ability of neurotransmitters to send information via electrical impulses to other neurons. In our model described previously, high levels of arousal may correspond to a high ability for neurotransmitters to initiate electrical signals. Low levels of arousal may lead to NE inhibiting neurotransmitters from propagating electrical signals.

Lastly, action potentials (electrical impulses) are an all-or-none phenomenon. An electrical threshold exists where, above this threshold, a neuron will initiate an action potential when stimulated. Below the threshold, the neuron will remain inactive. Each neuron receives numerous inputs from other neurons. These impulses are additive, and can combine to generate a strong electrical stimulation above the threshold level in the neuron receiving the inputs. This will elicit an action potential in the neuron, and propagate an electrical signal. The action potentials, though, always have a constant electrical amplitude for each neuron. Basically, action potentials are constant when they occur, not graded. To illuminate this point, imagine a single neuron (A) that is weakly

connected to four other neurons (B, C, D, E). Stimulation from a single neuron, B, is under the threshold value and A remains inactive. When all four neurons B, C, D, E activate, the sums of their electrical signals is greater than the threshold value, and A becomes active and creates an action potential.

There are three basic points to take away from this discussion. First, the brain sends information via electrical impulses called action potentials. Second, the brain is interconnected with different regions working together to perform an action. Third, neuromodulators regulate the effectiveness of neurotransmitters.

Figure A.1: Projections from Norepinephrine Neurons



From Aston-Jones and Cohen 2005
Note: The above image is of a monkey brain. The connections in a human brain are similar.
This image shows the connections between the brain region that secretes NE and the regions involved in evaluating rewards. The red lines represent connections between brain regions.

Appendix B

Table B.1: List of Mice and Genotypes

Mouse Number	Genotype	Group
26	WT	1
27	WT	1
28	WT	1
29	WT	1
30	WT	1
31	WT	1
32	WT	1
33	WT	1
34	WT	1
35	WT	1
2577	WT	2
2547	WT	2
2574	WT	2
2573	WT	2
2575	NET	3
2554	NET	3
2553	NET	3
2552	NET	3

Appendix C

Table C.1: Nosepoke Rates for WT and NET Mice

Genotype	Mouse	Nosepoke Rate While VA Patch Active (nosepokes / second)	Nosepoke Rate While VA Patch Inactive (nosepokes / second)	P-value
NET	2575	0.043021	0.016085	.0625*
NET	2554	0.026165	0.034220	1
NET	2553	0.070420	0.026123	0.031**
NET	2552	0.070921	0.027225	0.031**
WT	2577	0.0589756	0.0251112	0.312
WT	2547	0.0433460	0.0325298	0.125
WT	2574	0.0397241	0.0289258	0.437
WT	2573	0.0595987	0.0323970	0.156

Table E.1: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance. There are six observations per mouse. See Table 2 for an explanation of methods. Due to the lack of experiment days, mice 2575, 2553, 2552, 2547, and 2573 pass Hypothesis 6.1.

Table E.2: Probability of Nosepoking the VA Patch for WT and NET Mice

Genotype	Mouse	Nosepoke Rate While VA Patch Active (nosepokes / second)	Nosepoke Rate While VA Patch Inactive (nosepokes / second)	P-value
NET	2575	0.6314	0.5461	.812
NET	2554	0.6368	0.2077	.125
NET	2553	.59309	0.57427	1
NET	2552	0.69141	0.18178	.0310**
WT	2577	0.6528012	0.2604670	.0625*
WT	2547	0.7755587	0.3314550	.0625*
WT	2574	0.5815374	0.4248023	0.437
WT	2573	0.6914196	0.1817849	0.312

Table E.2: * values indicate 10% significance, ** values indicate 5% significance, and *** values indicate 1% significance. There are six observations per mouse. See Table 2 for an explanation of methods. Due to the lack of experiment days, mice 2554, 2552, 2577, and 2547 pass Hypothesis 6.2.

Works Cited

Daw D, O'Doherty J, Dayan P, Seymour B, Dolan R (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441: 876 – 879.

Pemberton, S. (2003). Hotel heartbreak. *Interactions, 10,* p. 64.

Pirolli, Peter. (2007). *Information Foraging Theory: Adaptive Interaction with Information*. Oxford UP, Oxford.