**Face Page**

Personnel:

Co-PI: Jean-Marc Fellous (UA Psychology), Co-PI Robert Wilson (UA Psychology)  
Consultant: Dr. Siyu Wang (NIH)  
Undergraduate students: Julia Wieland (Fellous lab),

Grant Mechanism: Faculty Pilot Grant Award

Project performance sites: Fellous (rodent) and Wilson (human) laboratories

Abstract (30 lines):

Humans and animals have to balance the need for exploring new options and exploiting known options that yield good outcomes. This tradeoff is known as the explore-exploit dilemma. To better understand the neural mechanisms underlying how humans and animals solve the explore-exploit dilemma, a good animal behavioral model is critical. Most previous rodents explore-exploit studies used ethologically unrealistic operant boxes and reversal learning paradigms in which the decision to abandon a bad option is confounded by the need for exploring a novel option for information collection, making it difficult to separate different drives and heuristics for exploration. In addition, these paradigms do not allow for observing model-based exploration behaviors, such as utilizing prior information and adaptation to the volatility of the environment. In this proposed study, we will investigate how rodents make explore-exploit decisions using a spatial navigation Horizon Task (Wilson, Geana, White, Ludvig, & Cohen, 2014) adapted to rats to address the above limitations. We will also record from distal hippocampal neurons known to be selective conjunctively to reward and spatial location (Xiao, Lin & Fellous 2020). Importantly, we will compare the rats behavioral performance to that of humans using identical measures. We will build on preliminary data showing that rats use directed exploration like humans, but the extent to which they explore has the opposite dependence on time horizon than humans. Moreover, we also attempt to explain why free choices and guided choices have fundamentally different influences on exploration in rodents. Given the similarities and important disparities observed between humans and rats, we will reveal a more complex explore-exploit behavior than previously thought, amenable to further studies at the behavioral and neural levels.

**Research Plan**

**I Specific aims**

It is well established that when humans are making repeated decisions between choices that have uncertain outcomes, past experiences with these choices matters. Recent work has also shown that knowing how many future decisions are permitted also matters (Cite Wilson). This effect is known as the Horizon effect and has been well documented in humans. The neural mechanisms involved are largely unknown and a rodent model of this phenomenon that matches the experiments conducted in human is lacking. Recent work in the Fellous laboratory has shown that a population of reward cells can be found in the hippocampus with conjunctive reward and location firing characteristics (Cite Fellous)

Rats can detect the cue of horizon, as their behaviors are different for different horizons. In the most successful pilot experiments (2 rats, but over XXX experiments), we showed that rats can associate sound cues to horizon conditions. The horizon conditions were randomly interleaved with the sound cue, and in this version, the 2 rats showed significantly different behavior in short vs long horizons. We did not have enough rats to develop these conclusions in our submitted behavioral paper.

Rats visit the unguided feeder less in long horizons, contrary to what we intended to train them.

Specific Aim1: Behavioral experiments and modeling

Specific Aim 2: Electrophysiological experiments

**II Research Strategy (4 pages)**

**General Methods**

## Subjects: Animals

A total of 14-16 Brown Norway rats will be used in these experiments. All rats will be male and female between 6 and 7 months old at the start of the experiment. We do not expect sex differences in these experiments, but data will be analyzed separately to confirm this hypothesis. All rats will be housed under reverse 12:12 light cycles. All animal procedures have been approved by the IACUC of the University of Arizona and follow NIH guidelines.

Figure 1:

## Subjects: Human participants

Male and female undergraduate subjects 18 years old or above from the University of Arizona will participate in this study. Participants who will not perform significantly above chance will be excluded. All participants will be from the undergraduate psychology subject pool and will earn academic credits for their participation in the study. The human experiments have been approved by the University of Arizona Institutional Review Board.

**Specific Aim1. Behavioral experiments**

Rats: The rodent experiments will be run in an open field maze that consists of a circular area (1.5 m diameter) with 8 equidistant feeders at its periphery (B. Jones, Bukoski, Nadel, & Fellous, 2012; B. J. Jones, Pest, Vargas, Glisky, & Fellous, 2015). Each feeder delivers sugar water (0.2g/ml) in the form of computer-controlled drops. A blinking LED is attached to each feeder and acts as a cue when desired. During each session, only 3 feeders are activated (Fig 1A, yellow light bulbs). One feeder is the home base; the two others, equidistant from the home base, are the reward feeders**.** The home base will never be rewarded, but animals have to reach it to trigger/activate the 2 rewarded feeders. The home base is flanked by two Lego blocks, forcing the animal to start its navigation to the 2 reward feeders without directional bias (Fig 1A, blue rectangles). Each feeder will initially deliver a predetermined amount of reward unknown to the rat between 0 and 5 drops (e.g. 2 vs 4 drops), and the rat will choose to explore or exploit its choices at each trial. For most of the times (1-H), the next reward generated by each feeder will be the same as the last, but for a small H percent of times, the next reward will be randomly drawn from 0 – 5 drops. If H is low, then the rewards of the two feeders shall remain unchanged for a long time. If H is high, then the rewards of the two feeders will be reset more often. After H trials, the size of the reward will change independently at each feeder (e.g. 0 Vs 3), and another set of H trials starts. H is termed the Hazard rate and is inversely directly related to the notion of Horizon in previous work (refs). Each sessions will feature two different hazard rates (Long horizon, low Hazard rate: H=5+-1, Short horizon, high hazard rate H=10+-1) in blocks of 20 sets of identical Hs. Behavioral and electrophysiological analyses will be done within blocks. Our preliminary data show that a rat can run 300-400 choices per day, which is enough for at least one high H and one low H block per day.

Humans:

In this experiment, participants will be sitting in a booth, in front of a computer screen. They will be asked to choose between two slots machines (also referred to as bandits, Fig 1B) that give out a fixed number of reward points uniformly drawn from 0 to 5. Participants will be instructed to maximize the total number of points. The height of the boxes indicate the number of choices allowed in the current game (i.e. the horizon condition, H=2 in Fig 1B) and each row represents a trial. Before participants make their own choices, in the very first trial, they will be guided to pick one of the bandits (Trial 1 cue, nG=1, Fig 1B). The option available will be cued with a green background color. Participants indicate their choices by pressing an arrow key on the keyboard. Their response will be followed by an indication of how many rewards they obtained, the reward of the unchosen option will not be shown and will showed up as ‘XX’ (Trial 1 response, Fig 1B). From the 2nd trial, both bandits will be available and participants will be free to make their own choices. There will be four horizon conditions (H=1, 2, 5, 10 free choices), and games with different horizons will be pseudo-randomly interleaved. Fourty human participants would complete a total of 6080 games (33440 trials).

Caveats and alternative plans for the rodent experiments:

**Specific Aim 2: Electrophysiology**

General methods:

Electrophysiology: Male and female Brown Norway (6-8 months old) rats will be food restricted and pre-trained as previously done in the Fellous laboratory (refs). Animals will be implanted with a 18-tetrodes device targeted at the dorsal CA1 area of the hippocampus (-3.8 mm posterior to bregma, 2.1-2.6 mm lateral to midline). This targeting will be made possible by 3D printed implant exit tips produced in-house. Data will be collected with a second-generation Digital FreeLynx wireless system connected to a Halo-18 tetrode drive. Both spike and LFP data (e.g. theta, sharpwaves ripples) will be collected. Single neurons will be isolated offline and the quality of the isolation will be assessed (refs). The data will be analyzed with custom written Matlab code (refs). Animals will be tracked using an overhead video camera at 25-30 Hz. Place fields are computed making sure to exclude sharpwave-elicited activity (refs). The position of the tetrode tips will be determined in all animals by electrolytic lesions and standard Nissl staining.

Experiment 1: Conjunctive reward-location cells

In recent publications, we and others have found the presence of a continuum of conjunctive reward and spatial location cells in the distal portion of area CA1 of the hippocampus {Xiao, 2020 #3088;Gauthier, 2018 #3148}. A subset of these cells transition from reward-dominated firing to spatial-dominated firing within the same day, for reasons yet unknown. Some reward cells (15-20% per session) were predictive of the upcoming rewards, others are indicative of the reward just obtained. We will implant a subset of rats and record from the distal CA1 in the task described in SA1. We will analyze the firing of the cells in blocks of High or Low hazard rate. In each block, the population of recorded cells will be separated in 3 categories: Reward dominant, Spatial dominant and Conjunctive, using place and reward scores computed as in {Xiao, 2020 #3088}. We expect that the number (or firing rate, or both) of reward-dominant cells will differ between high hazard blocks (short horizons) and low hazard blocks (long horizon). A larger amount (or higher firing rate) of reward cells in one condition versus the other would suggest a finer coding of reward values.

Experiment 2: Analyses of Vicarious Trial and Error

During spatial decision making, rats are known to stop temporarily and head scan for a few seconds before resuming their movement. Evidence suggest that this pause, termed Vicarious Trial and Error (VTE) occur when reward contingency change or when decision making became difficult and was associated with a change in firing characteristics of hippocampal place cells {Redish, 2016 #3214}. This phenomenon was seen in spatial tasks in humans and noted to be predictive of spatial performance {Santos-Pata, 2018 #3215}.

We will extract VTE epochs in rats using established methods {Johnson, 2007 #1798}. <more>

**III Plan for Submission of External Grant Application**

We plan to submit an RO1 application to NIMH or NINDS. By that time, we expect the behavioral data obtained so far to be published. The application will include preliminary results from the electrophysiological experiments collected as part of this pilot grant (the collection of these data cannot be achieved otherwise in the Fellous lab as of summer 2022)

**IV Timeline**

Months 1-12: behavioral data collection (10 rats). Electrophysiology data collection (4 rats).  
Months 5 -16: rodent data analyses, figure production for inclusion in grant application.  
Months 17-18: submission of grant application (estimated summer 2024)