Proyecto de tesis doctoral para optar al grado de Doctor en Neurociencia en la Pontificia Universidad Católica de Chile: Orexin and uncertainty effects on food-seeking behavior

Luis Nicolás Luarte Rodríguez - Claudio Peréz-Leighton

Introduction

Mammals evolved in an environment of food scarcity; thus, maintaining fat reserves and overall caloric intake is extremely important. Most animals accomplish this by increasing food-seeking behavior when food access is limited, resulting in extended foraging bouts. However, foraging and larger fat reserves can increase predation risk by increasing exposure and reducing mobility, respectively. Therefore, increasing foraging bouts should only occur when the risk of starvation outweighs the risk of predation. One environmental clue that animals use to make this decision is uncertainty in food access: higher levels of uncertainty in food access predicts future food scarcity, triggering food-seeking behaviors. This prediction allows animals to act upon proximal cues without knowing the complete state of the environment.

Relating uncertainty to food scarcity is a successful adaptive strategy when food is limited, but modern environments are plentiful. Thus increased food-seeking behavior leads to excessive intake, higher risk of overweight and obesity. We propose that the prediction error, between expected and actual intake from foraging bouts, translates into a measure of food-access uncertainty. Increasing uncertainty generates unreliable expectations, therefore, more prediction errors. Leading to an orexin mediated increase in food-seeking behavior.

The neurobiological mechanisms that regulate food-seeking behavior and foraging must use information about a nutrient deficit to increase alertness and physical activity to forage successfully. The neuropeptide orexin has the potential to be a mediator of foraging behavior. Fasting and intake respectively increase and decrease the activity of orexin neurons, and activation of orexin receptors in different brain sites increases physical activity and food intake. Additionally, orexin promotes foraging behavior by increasing locomotor activity before food intake in response to a nutrient deficit.

Together, these data led us to hypothesize that orexin promotes food-seeking-related behaviors when facing uncertainty related to food access. Our overall aim is to determine how uncertainty in food-access increases food-seeking behavior, and how orexin mediates uncertainty-driven increased food-seeking behavior.

General research strategy

To determine how uncertainty in food-access increases food-seeking behavior we will use two experimental strategies: (1) behavioral modeling using the reinforcement learning framework to assess changes in reward processing due to uncertainty, and (2) operant tasks to manipulate uncertainty levels in food acquisition. To determine how orexin mediates uncertainty-driven food-seeking behaviors we will (1) assess differential gene expression in conditions with and without uncertainty using real time qPCR, obtaining correlational-level data on orexin as a mediator of uncertainty-driven food-seeking behavior; (2) determine if orexin is necessary to drive increased food-seeking behavior in uncertain conditions with the use of orexin antagonists, and gain of function with orexin agonist, (3) We will test functional connectivity between hypothalamic orexin neurons and paraventricular nucleus/ventral tegmental area (VTA), to determine if orexin mediation of uncertainty-driven food-seeking behavior can be supported as a modulation of reward-related systems.

Theoretical and empirical framework

The natural setting for food-seeking behavior

Foraging encompases all behaviors related to obtaining food, including food-seeking, feeding and hoarding (Kramer, 2001). Food-seeking behavior is the re-orientation of attention and locomotor activity with the goal of acquiring food. A successful food-seeking behavior minimizes energy expenditure while reducing exposure to potential predators to find food. Because the environment is constantly changing, food seeking strategies must also be constantly adapting (Bartumeus et al., 2016; Kölzsch et al., 2015). However, animals are never completely aware of when or how an environment changes, so they always have incomplete knowledge of food location, quality and the probability of obtaining food (G. H. Pyke, 1984). Thus, the main challenge for the foraging animal is how to determine an optimal strategy to procure food while having only partial knowledge of its food environment.

The necessary information to determine a food-seeking strategy are the overall statistical properties of the environment: how dense are the food targets in a given area, how variable is the quality of food (i.e. calories, organoleptic properties), and what is the expected outcome of foraging in a given area. Given that an animal must consider all these potential characteristics of the environment to search and find food while minimizing energy expenditure and predator exposure begs the question: ***what kind of strategy could accomplish this while having only partial or incomplete knowledge about the food environment?***

The movement data of several species (including human) during food-seeking behavior can be described using Lévy-walk patterns (Garg & Kello, 2021; Kölzsch et al., 2015; A. Reynolds et al., 2018; Viswanathan et al., 1996). Lévy-walks are random walks with a Lévy distribution, which produces heavy-tails and describing multiple concentrated movements with sharp turning angles followed by few ballistic displacements. These patterns produce optimal searches in environments with a patchy distribution of resources (clumped resources distant to one another) (**97UESCC6?**). Although how animals produce Lévy-walks patterns is unknown (A. Reynolds et al., 2018), this mechanism might be partially independent of sensory information regarding environmental feed-back (Humphries & Sims, 2014; Sims et al., 2019). Wait and move foraging patterns, across multiple species, retain the same Lévy-like distribution in natural, simple captive environments (Wearmouth et al., 2014) and when devoided from environmental feed-back (Maye et al., 2007). Showing that important features of foraging such as deciding when to wait for prey, and when to move on to catch prey are independent of sensory information. Moreover, the Lévy-walk foraging pattern is present in multiple species which inhabit diverse environments and have different sensory abilities (Humphries et al., 2010). Evidence of Lévy-walks being either completely or partially independent of sensory information are thought to reflect a strategy selected in early evolution to optimize food searching with partial or complete lack of knowledge of food location (**97UESCC6?**).

While Lévy-walks can describe a search strategy to find food with only partial knowledge of the environment (Kölzsch et al., 2015; Sims et al., 2019), sensory information can inform when to switch the strategy. It has been observed that food encounters create deviations from Lévy-walks making movement more focused within the area where food was found (de Jager et al., 2011). The intuition here is simple, if an animal finds food in a particular area, then there is a good chance that more food will be nearby. This strategy switch is thought to reflect an adaptation to environments where food is usually clumped together (Nauta et al., 2020; A. M. Reynolds & Frye, 2007). While

this strategy switch is informed by sensory information, the main feature is that it relies on simple strategies that are in tune with the usual distribution of food in the environment (Humphries & Sims, 2014). Thus, limiting the need to use sensory information.

Starting from the base that animals have only partial knowledge about food location, we discussed that Lévy-walks are a food seeking strategy partially independent of perceptual information, and we proposed that this strategy could have emerged to deal with environmental uncertainty. In the next section, we will explore how considering uncertainty, within the food-seeking strategy, can be useful when perceptual abilities are limited, and how uncertainty itself can modify food-seeking behavior (Anselme et al., 2017; Anselme & Güntürkün, 2019; Anselme & Robinson, 2019).

Foraging and uncertainty

As discussed previously, animals adopted food-seeking strategies to procure food within uncertain environments (Humphries & Sims, 2014; Sims et al., 2019). Here, we will discuss how environmental uncertainty can regulate food-seeking behavior and the neuronal representation of uncertainty.

To illustrate how environmental uncertainty can influence food-seeking behavior, we can consider two scenarios, one where food-seeking behavior strategies does not consider uncertainty and one where it is considered. In the first scenario, an animal will randomly sample a section of its environment for food sources and pick the one that, at that time, delivered the greatest amount or quality of food. This sampling and choice strategy is called greedy (Tokic & Palm, 2011) and implies that the animal will continue to choose the option that yielded the most value in an initial random sampling, missing out on potentially better alternatives that did not deliver good rewards or sticking with its original choice even when its value or quality has dropped due to food resources being consumed. This is problematic as it makes food-seeking strategies insensitive to reward variation (Sutton & Barto, 2018). An improvement over acting greedily is to store in memory the mean value (i.e. quantity or quality) of the food obtained over time, to compare the selected alternative against the overall quality of the environment. This idea is implemented in the now classical model by Charnov (1976), which proposes that animals should decide whether to stay or leave a certain location based on the mean value of its past choices. However, this model does not provide a good fit to behavioral data, showing a clear bias towards longer than predicted staying times (Hayden et al., 2011; Le Heron et al., 2020; Nonacs, 2001; G. H. Pyke, 2010).

For the second scenario, there are several options to describe uncertainty and we will discuss two: By considering the spread (i.e. standard deviation) over possible outcomes (Rothwell & Stock, 1988) (i.e. a food source has food consistently but the food obtained in quality/quantity vary over time) or by considering a change in the contingency between an action and its outcome (i.e. a food source that had high quantity and quality food over time suddenly has no food). When considering the spread as a measure of uncertainty, it is necessary to calculate an average outcome from which the spread is observed. The anterior cingulate cortex (ACC) is able to represent the magnitude of the spread over possible outcomes in decision making tasks (Christopoulos et al., 2009) with high precision, for every alternative, different cues and contexts (van Holstein & Floresco, 2020). The second option, the change in the outcome of an action, is tracked by increased ACC activity (Behrens et al., 2007). Thus, ACC activity can signal whether an outcome is within the expected values or the value obtained in a particular occasion is very different from expected. Being able to distinguish between these two cases can increase food-seeking success. Being able to compute these uncertainty features allows animals to stay in an area that provides food, despite occasionally

not finding food (change in action-outcome but obtained value is within expected values based on tracked average), but leave this area if consistently fails to find food (change in action-outcome and obtained value has large variation relative to tracked average). The orbitofrontal cortex (OFC) has a role in this regulation. Stolyarova & Izquierdo (2017) showed that rats could choose the option with largest mean value, that is, the option that delivered sucrose pellets with the least amount of delay, despite high variability in delay times. However, a lesion to the orbitofrontal cortex (OFC) impaired this ability, making rats unable to change their option when the mean value was decreased (increased delay time). Together, these data show that different measures of uncertainty are tracked by brain structures, allowing one to choose the best option while filtering out the variation around the expected value.

The intuition that an animal should behave differently when the outcome is within expected values or has changed drastically is formalized in the expected and unexpected uncertainty model (A. J. Yu & Dayan, 2005). This model describes expected uncertainty as the uncertainty regarding outcomes when contingencies remain stable, but are subject to some noise (variation around the expected value). On the other hand, unexpected uncertainty represents a drastic change in the contingencies, likely due to a structural change in the environment. Using these concepts, the expected/unexpected uncertainty model proposes that if the obtained rewards fluctuate near its expected value (i.e. a food source has food consistently but the food obtained varies in quality/quantity over time) the animal should not modify the learned actions that lead to rewards in this environment (i.e. the animal should ignore cues that could point to new potential food sources and continue to choose this food source), so top-down control (where value representations are) should dominate over bottom-up input (sensory information). On the other hand, if the variation of the obtained rewards increases, the balance should switch to prioritize bottom-up input to increase learning about new actions, that could lead to increase or maintain the expected value of the rewards (i.e. the animal should pay attention and follow-up on cues that indicate new food sources) (Soltani & Izquierdo, 2019; A. J. Yu & Dayan, 2005).

As discussed, animals can determine whether changes in the outcome are expected or unexpected. If the action-outcome contingencies never changes, the environment is described as stationary, whereas if contingencies change or the mean value goes up or down over time, then the environment is defined as non-stationary (Raj & Kalyani, 2017). Animals tend to behave as if the environment is non-stationary (Wu et al., 2018), increasing exploratory behavior (Ryali et al., 2016), as predicted by the expected/unexpected model (Cohen et al., 2007; A. J. Yu & Dayan, 2005). While the functional reasons behind this behavior are not clear, assuming a non-stationary environment does not decrease performance in stationary environments significantly (Ryali et al., 2016), and can provide near optimal performance in natural environments (Reverdy et al., 2014).

However, excessive exploration can be detrimental to obtaining food, so exploiting current knowledge about food location is also necessary. This leads to the exploration-exploitation dilemma (Sutton & Barto, 2018). Exploiting means that behavior should be consistent with previously learned reward contingencies, choosing the option with the highest expected value. On the other hand, exploration implies re-sampling the environment to improve or re-learn contingencies. Uncertainty is a key variable to the resolution of this dilemma. Unexpected uncertainty increases the exploratory behavior, boosting learning of new contingencies, whereas expected uncertainty bias behavior towards exploitation (Cohen et al., 2007; G. C. Harris et al., 2005). This idea implies that animals should stay in areas where foraging is successful, without paying much attention to small fluctuations in the results. However, if obtaining food suddenly becomes more uncertain,

the animal should start exploring for other options, because the expected value of that location might have changed, and other locations might now be better.

A key assumption behind the usefulness of uncertainty is that animal perceptual abilities are limited. We argue against the idea of perceptual abilities being dominant over the search strategies already discussed (LaScala-Gruenewald et al., 2019), because sensory perception is always limited in some aspect, introducing uncertainty of what lies beyond its limits (Bartumeus et al., 2016). All sensory organs have a receptive field, which defines the range where information can be sensed and thus used to orient food-seeking behavior (Fletcher et al., 2013). If an animal must know what is beyond such range, the animal must move to a new area and explore it. However, the decision of where, when and how to move requires input from memory, integrating perceptual information or other cognitive processes (Ranc et al., 2021), because by definition the animal has no sensory information about the new area, and somehow must determine if moving to the next area is good or bad. However, what is remembered about a certain location loses validity over time due to the environment being in constant change. Thus, the animal must integrate information about how food availability changes over time, if the food source is always present then there's certatinty of its availability, on the other hand, if the food source comes and goes then there's high uncertainty over its availability (Fagan et al., 2017).

Integrating information of location and availability over time can inform food-seeking behavior. However, obtaining good quality information requires the animal to explore its environment. Exploring requires a balance between how fast it should be done and how exhaustive the animal should be, this balance is represented by the speed-perception tradeoff, which determines that perceptual abilities are degraded as speed is increased (Campos et al., 2013). For example, fast speeds are required to capture a moving prey, as the prey can move or be consumed by another animal, rendering the information useless. However, moving fast prevents animals from making an exhaustive exploration of its environment, as detection accuracy drops by moving faster (Fagan et al., 2017). A second problem is the intensive-extensive tradeoff, which points out how finding food-resources nearby impairs finding resources far-away. If the animal performs an exhaustive search in its nearby areas, it must be done slowly to be accurate. However, being slow means that far away areas are left unattended longer, increasing uncertainty about their food resources (Bartumeus et al., 2016; Raposo et al., 2011). Both tradeoffs imply that obtaining knowledge about food resources leads to an inevitable perceptual uncertainty; moving fast makes information about food location harder to obtain, but at the same time moving fast is necessary to actually obtain food. Additionally, appropriately exploring the environment forces the animal to focus on one area and to grow increasingly uncertain of other areas' food resources. Together, these tradeoffs illustrate how, even when perceptual abilities can inform food location, there is an inescapable uncertainty that animals must deal with.

The limitation in perceptual abilities makes uncertainty inescapable. Even when an experienced animal can remember optimal foraging paths, random searches with distinct cycles of exploration/exploitation phases persists (Kembro et al., 2019). Animals still explore, even when having knowledge of food location, because introducing stochasticity in food-seeking behavior improves success, by making strategies more resilient to cognitive errors derived from perception (Campos et al., 2013). The persistence of strategies that balance exploitation/exploration (Campos et al., 2013), even when they are not technically needed (Ryali et al., 2016) shows how relevant uncertainty is for food-seeking behavior.

In this section we discussed the relevance of uncertainty and the limitations of sensory information

for food-seeking behavior. In the next section, we will examine models that consider the case of foraging in uncertain environments to inform about the underlying processes in food-seeking behavior.

Foraging models and underlying processes in food-seekings behavior

Having discussed the importance of uncertainty for food seeking, we now move to provide a more formal framework to relate uncertainty with the exploration/exploitation dilemma and perceptual limitations discussed in the previous sections. A model that includes uncertainty into food-seeking behavior should include the following considerations: first, the rules determining the result of the interaction between animal and environment are assumed to be unknown or only partially known due to the stochastic nature of the environment. Secondly, the animal may take any action a within a set of possible actions $a \in A$ for a particular state of the environment s. Any action a causes an stochastic transition from a state s to another state s'. As such, the result of an interaction between animal and environment can be described by its value q, which is a function of both action and current environment state $q(s, a)$ (Fig. 3).

These considerations are included in a Markov decision process (Sutton & Barto, 2018), which captures the intuition of the decision making process where the animal can take action in the environment, but the action outcome is partly random, and dependent on the current state. In this model, all environment dynamics are described by the probabilities $p(s', r|s, a)$, where r is the obtained reward (interaction outcome), and such probabilities are defined for every pair of a and s. A Markov decision process that includes the perceptual uncertainty, which we deemed inherent to food-seeking behavior, can be included by considering that states s are paired with observations o made by the animal to infer state s, because the state cannot be directly observed or there is some sensory noise. Including perceptual uncertainty makes the animal to consider environment states as the conditional probability of any particular observation given a state $p(o|s)$, generating a belief of the current state based on perceptual information (Ma & Jazayeri, 2014).

To model how an animal represents the value of a given option q in an uncertain environment, the value of the option becomes a distribution over possible values that is updated every time an action a is performed. For the simple case where rewards are obtained or not (without any difference in reward magnitude), $q(s, a)$ has a Bernoulli distribution $p(X = reward) = a$ and $(p(X = noreward) = 1 - a)$. Then, these probabilities can be modeled with the Beta distribution, which takes parameters α and β. With $\alpha = 1, \beta = 1$ the Beta distributions produce a uniform distribution over $[0, 1]$ successfully representing the uninformed prior probability for the rewards, representing total uncertainty about option value. To generate the posterior probability, every time an action results in a reward, the parameter α increases by 1, biasing the distribution towards 1. On the other hand, if no reward is obtained, the parameter β increases by 1, biasing the distribution towards 0.

Finally, the mean is defined as

$$\frac{\alpha}{\alpha + \beta}$$

and its variance by

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

With these simple statistical properties of the Beta distribution, we can represent uncertainty over the expected rewards for any given a and s. The previously presented way to model environment uncertainty and q is the general case of Thompson sampling (Thompson, 1933). To select an action, a posterior is built for every action and updated according to the previously stated rules (Fig. 4). Then, for each posterior, a single reward estimate $\hat{r}$ is sampled, resulting in an estimated value for each action. The action selected greedily, so $a = argmax_{a \in A} \hat{r}(a)$ where A is the set of possible actions within an environment (Wang & Zhou, 2020). With this simple algorithm for action selection, the exploration and exploitation of different actions is balanced, actions with high associated value and certainty will likely draw high values in the sampling procedure, actions with low expected value and high certainty will draw low values, and actions with high uncertainty can draw lower or higher values. Because the process must be performed for every state, tractability is limited by the number of states. To avoid the problem with tractability, rewards can be summarised as a weighted average of past rewards, with a step-size parameter $((0, 1])$. The lower the value of this parameter the more weight is given to recent rewards, on the other hand, if it's closer to 1, then all the reward history is equally considered. More complex alternatives to this problem include modeling non-stationarity as Poisson arrival process that modifies the means rewards (Ghatak, 2020), bayesian approaches to modulate past observed rewards (Raj & Kalyani, 2017), and explicitly modeling environment volatility in a bayesian setup (Behrens et al., 2007).

While this general model can work in non-stationary environments, it doesn't explicitly considers the belief of the current state based on the perceptual information received $p(o|s)$. To include this, a probability for every $o \in O$ by state is necessary, where O is the set of all particular observations o. To model state beliefs, the goal is to obtain the function that maps observations o to action a, given an underlying model that relates states with observations. A hidden Markov model (HMM) represents this. HMM generates conditional probability distributions $p(o|s)$. This allows us to explicitly model how an animal infers the current state given perceptual information (Funamizu et al., 2012; Piray & Daw, 2020; Yoon et al., 2018).

In this section, we offered the elementary considerations for a model of food-seeking behavior in non-stationary environments, with uncertainty over action outcomes due to perceptual limitations. Thompson sampling was considered as the base for this due to its simplicity and elegance in balancing exploration/exploitation and including uncertainty in the decision process. The goal here was not to establish or to specify a complete model, but to provide a more formal framework to relate uncertainty with the exploration/exploitation dilemma and perceptual limitations discussed in the previous sections. In the next section we provide evidence on how uncertainty is computed following the framework presented above and introduce the reward prediction error as a simple process that could allow animals to compute uncertainty.

Computing uncertainty

In the previous section we presented HMM to represent what is the belief of the current state of the environment using only partial information. On the other hand, Thompson sampling allowed us to model the uncertainty over q, that is, the uncertainty over the value of an action in a given state. In this section we will discuss how uncertainty can be computed in the animal brain, and how it modifies parameters presented in previously discussed foraging models.

Uncertainty arises from having more than one option, but also when more than one option is attractive. If the probability of choosing any given option has a uniform distribution, then uncertainty increases proportionally with the number of options. Shannon entropy (Shannon, 1948) formalizes this intuition as follows

$$H = -\sum_{i=1}^{n} p_i log_2 p_i$$

This formula expresses uncertainty in bits. So, maximum entropy (one bit) is achieved when all the alternatives have the same probability, such as a coin flip. However, if the coin happens to have two heads, then Shannon entropy is 0 bits. If we apply the Shannon entropy to estimate uncertainty in food seeking, we can start by considering an animal in simple environment with only one environmental state s and one action a which initiates a food-seeking bout. If there are only two possible outcomes: food is found (p) or not found ($q = 1 - p$), then $H = -(plog_2 p + qlog2q)$. If an animal performs multiple food-seeking bouts and all of them have the same result (either successful or unsuccessful), then there is no uncertainty $H = 0$. However, if the probability of a successful food-seeking bout is 0.5, then entropy is maximized $H = 1$.

The computation of entropy has a neural basis. Uncertainty can be determined in decision making tasks, by observing the probability of choosing between two alternatives. The highest uncertainty is achieved when the probability of choosing any given alternative is 1 divided by the number of possible alternatives. Goñi et al. (2011) showed that when choosing between two stimuli with different rewards, the choosing uncertainty for any given pair of alternatives was encoded in the middle cingulate cortex (MCC). Gloy et al. (2020) explored a similar setting by asking participants whether they should bring and umbrella to a forest barbeque by giving them two sources of information (1) whether forecasts and (2) images of the sky, for every tested combination of these sources of informations, decisions were classified as either certain (consistently choosing to bring or not an umbrella) or inconsistent decisions (combination with high decision uncertainty), results showed that higher uncertainty increased activity in MCC, whereas certain decision were encoded within the left supramarginal gyrus. Theses data suggests that uncertainty is encoded as entropy within the MCC, and this process seems to be important when trying to determine the true state of the environment (whether is going to rain or not). Thus, the neural basis of entropy computation seems to dependent on the estimation of the true environment state.

Entropy is not directly available as sensory input, and must be derived from actions and outcomes, which are dependent on the state of the environment. Previously, we discussed how Thompson sampling can encode the variance of an outcome in a posterior distribution and this can be used to model the transitions in a HMM. In Thompson sampling the variance of the posteriors is related to entropy, the more variance the more entropy, as the probability of each possible value is more similar. Nevertheless, a more direct way to compute entropy is through the prediction error. The classical model of Rescorla-Wagner (Rescorla et al., 1972) models how animals could predict the reinforcing value of a given stimuli

$$y_t = y_{t-1} + \alpha\delta_n$$

where the value representation of the stimuli y is obtained by considering the previously estimated value \$y_{t-1} weighted by a learning rate α, and a prediction error δ. $delta$ is the simple differ-

ence between the expected reward and the actual reward

$$\delta_n = r_t - y_t$$

where r_t is the obtained reward. An extension to this model (Sutton & Barto, 2018) modifies the prediction error by giving more weight to current rewards compared to past rewards

$$\delta_n = Rt + \gamma \hat{V_{n+1}} - \hat{V}_n$$

where γ is a discount factor $0 <= \gamma < 1$ for all the history of rewards, and $\hat{V}$ is an estimate for the true value of the reward. Finally, $\alpha : [0, 1]$ is the learning rate which effectively weights the reward prediction error δ so to make small updates $\alpha \approx 0$ or rather large ones $\alpha \approx 1$ to the reward estimation.

The Rescorla-Wagner model allows deriving a prediction error based on experience, where the learning rate can be set lower to simulate unexpected uncertainty or higher to simulate expected uncertainty. However, α in such a model is a hyperparameter, thus is not derived from experience. (**NZFTTQJZ?**) model proposes that α can be controlled by the prediction error magnitude $|\delta|$ so

$$\alpha = \gamma|\delta_{n-q}| + (1 - \gamma)\alpha_{n-1}$$

Higher entropy on reward outcomes increases the probability of error (Feder & Merhav, 1994), thus increasing $|\delta|$ and, consequently, α. The behavioral intuition captured in the Pearce Hall model is that the animals should increase its behavioral vigor and attention towards options with the uncertain outcomes (Diederen & Fletcher, 2021).

In this section, we presented simple models showing how uncertainty can be computed using the reward prediction error without any meaningful computational complexity. In the following section, we will brain structures supporting the computation of the reward prediction error and uncertainty.

Neuronal representation of uncertainty

A good candidate neurotransmitter for representing uncertainty is dopamine (DA). The firing rate of DA VTA neurons has been proposed to encode the prediction error (C. D. Fiorillo, 2003; C. D. Fiorillo, 2011; Gershman, 2019; Khaw et al., 2017; Lak et al., 2014; Nasser et al., 2017; **HLC-SQTJB?**). DA neurons firing rate increases when the number of rewards increases suddenly over the base rate, and decreases its firing rate in response to an omission in reward (Takahashi et al., 2017). Moreover, DA neurons sustained firing has been show to encode reward probability analogously to entropy, that is, displaying a peak of activity when the reward is obtained with a probability of 0.5, and lower levels of sustained activity when the reward probability is either high or low (low entropy) (C. D. Fiorillo, 2003). C. D. Fiorillo (2011) showed how rhesus macaques DA neurons firing rates were higher for uncertain stimuli compared to certain stimuli, while controlling for the expected value by pairing the ceratint stimulus with a fixed amount of reward delivery, and the uncertain stimulus with either no reward or double the reward of the certain stimulus, so the expected values were the same.

Encoding of uncertainty seems to be fairly distributed within brain areas. Hippocampal acitivity has been shown to encode stimulus entropy (Schiffer et al., 2012); substantia nigra encodes unexpected gains and losses within a financial context (Zaghloul et al., 2009); VTA DA neurons activity

has been proposed to be modulated by ACh uncertainty signal (Reverdy et al., 2014), changing the balance between exploration and exploitation (Cinotti et al., 2019). VTA DA neurons seems to respond accordingly to the (**NZFTTQJZ?**) model presented previously, as optogenetic activation of VTA DA neurons during reward prediction errors increased cue-reward learning (Steinberg et al., 2013), and L-DOPA has shown to increase learning for new associations (Chakroun et al., 2020). Thus, general DA activity is sensitive to uncertainty signals carried by reward prediction, showing (**NZFTTQJZ?**) model prediction regarding increasing learning rate, and showing connectivity to uncertainty encoding areas. Moreover, hyperdopaminergic mice show a bias towards exploration even when this implies greater costs, which according to reinforcement learning modeling point could be due to a decoupling between actions and reward history (Beeler et al., 2010). Within the reinforcement learning framework this means that mice give less weight to the expected value of their options during choice, thus increasing action randomness.

The VTA DA neurons computes the reward prediction error (**NLDHLRVN?**), by weighing inputs from multiple brain areas, including the lateral hypothalamus, dorsal and ventral striatum, ventral pallidum, and subthalamic nucleus (Tian et al., 2016). The main intuition here is that VTA DA neurons generate a reward prediction signal using multiple redundant inputs, that compute slightly different reward features. For example, the lateral hypothalamus has been shown to encode reward properties modulated by internal states (**KPIHYUYF?**) and this subjective valuation of reward is passed onto VTA DA neurons (**NLDHLRVN?**).

Acetylcholine (ACh) and norepinephrine (NA) activity has been proposed to signal expected and unexpected uncertainty, respectively (A. J. Yu & Dayan, 2005). Both neurotransmitters are mainly produced in the basal forebrain (Sturgill et al., 2020) and locus coeruleus (LC), showing an important role in the modulation of the reward prediction error signal (Aston-Jones et al., 2010; Sales et al., 2019). ACh antagonists increase sensitivity to random fluctuation in rewards (Marshall et al., 2016), so that they are no longer perceived as expected. On the other hand, normal ACh concentration levels activity signals that random fluctuations are to be expected, and no meaningful changes are present in rewards (A. J. Yu & Dayan, 2005). The opposite role is attributed to LC tonic activity signaling unexpected uncertainty, which promotes the learning of new contingencies, increasing exploratory behavior via DA sensitization to reward fluctuations (Aston-Jones & Cohen, 2005; **9Z525EYW?**).

In this section, we presented how the reward prediction error which allows the computation of uncertainty is supported by DA activity within the VTA modulated by ACh and NA signals of expected or unexpected uncertainty. In the following two sections, we will show empirical evidence on how food-access uncertainty increases food-seeking behavior, and increases the risk of overweight in modern times. Finally, we will propose the orexin neuropeptide as a potential mediator of the uncertainty-driven increase in food-seeking behavior.

An adaptive strategy in modern times

In modern urban environments, high-fat food is of easy access; this, coupled with a food-seeking behavior which seeks to maximize energetic gain when food shortage is predicted due to food-access uncertainty, can lead to overweight, because the mechanism is adapted to low resource environments and not calorically dense one. Thus, in developed countries, where caloric density is extremely high (Drewnowski & Darmon, 2005), increasing food-seeking behavior is likely to result in excessive caloric intake. Here we will present evidence on how food access uncertainty increases food-seeking behavior, and how this can result in a increase risk of obesity.

When food is limited, animals approach cues with the highest associative strength to actual food (Montague et al., 1996). Kacelnik & Bateson (1996) in fact proposed that approaching behavior should be reduced if the cues become unreliable, for example, by delivering rewards only 50% of the time. However, introducing uncertainty in food access reduces the time latency of lever approach and increases lever pressing rate (Anselme et al., 2013). Showing increased food-seeking behavior due to uncertainty in food-access.

Intermittent access protocols allow us to study the effects of introducing uncertainty in food-access. These protocols basically consists either giving a brief access to a palatable or non-palatable food option, randomly or every other day throughout the week. Intermittent access to high-fat diets generates binge-eating behavior, which corresponds to a drastic increase in food intake over a short period of time (Hess et al., 2019; King et al., 2016; Lardeux et al., 2013). Moreover, intermittent access for vegetable shortening in rats increases lever pressing for food upon withdrawal (McGee et al., 2010) and without withdrawal (Wojnicki et al., 2013; Wojnicki et al., 2007, 2015). The increase of food-seeking behavior due to intermittent access to food increases general psychomotor behavior, and attenuates the effects of food devaluation (Hardaway et al., 2016). Furthermore, intermittent access is correlated with an increase DA concentration within the accumbens shell, and increased motativation to obtain food (Rada et al., 2005). Is important to note that intermittent access is typically delivered at regular hours, and this allow animals to predict food arrival (Luby et al., 2012). However, the behavioral effects, except anticipatory behavior, are common disregarding if intermittent access is provided at completely random time or given at fixed times of the day (Muñoz-Escobar et al., 2019). Taken together, this data suggests that introducing uncertainty in food-access robustly increases food-seeking behavior and food intake, while reducing devaluation effects on intake.

Food-seeking behavior is increased to avoid starvation when a food shortage is predicted (by increased uncertainty), and as previously noted, this derives in increased exploration, number foraging bouts, and time expended in foraging (T. R. Harris et al., 2010). In humans, food-access uncertainty is analogous with food insecurity, which is defined as the perception of how secure or certain food access is going to be in the future (Dhurandhar, 2016). Higher levels of food-insecurity correlate with increased food intake (Dhurandhar, 2016) and preference for high-fat alternatives (**EER2TNCJ?**), which is in line with an adaptive behavior to prevent starvation in food-limited environments (Brunstrom & Cheon, 2018). However, this adaptive strategy turns problematic as calorically dense are of easy access and the cost to obtain calorically dense food has dropped in modern countries (Drewnowski & Darmon, 2005), were the effect of food-insecurity on food-intake is more pronounced (Dinour et al., 2007; Moradi et al., 2019; Nettle & Bateson, 2019). Increased food intake due to food-insecurity increases the risk of obesity and overweight, probably due to overeating fats and carbohydrates (which correspond to easy-access food previously mentioned) in periods of high food availability (Stinson et al., 2018).

In this section we showed evidence connecting food-access uncertainty and the risk of overweight and obesity in modern environments, where calorically dense foods are of easy access, and the uncertainty-driven increase on food-seeking behavior results maladaptive. In the following section, we present orexin as a potential mediator of uncertainty-driven foraging because of its pivotal role in both reactive and predictive homeostatic control (Burdakov, 2020), and motivated behavior (**2X7SNKS3?**).

Orexin as a potential mediator of uncertainty-drive foraging

The hypothalamus is a relevant structure in the homeostatic process, being capable of controlling arousal levels (Adamantidis et al., 2010; Kosse & Burdakov, 2014), motivation for food intake (Castro & Berridge, 2017), receiving internal status information of fat deposits via leptin signaling (Meister, 2000; Pandit et al., 2017), and gastrointestinal status via ghrelin signaling (Müller et al., 2015; **W3IUQNUY?**). Hypothalamus capability of sensing internal status and controlling motivation could allow it to be an important contributor to uncertainty-driven foraging, specifically its orexin neurons, as they show important contributions to foraging related adaptations (**LBGV5NJ5?**).

Orexin or hypocretin is a neuropeptide with few neurons producing it, most of them located within the lateral hypothalamus and perifornical area, but with large projections throughout the brain (Peyron et al., 1998 ; Chowdhury et al., 2019). Orexin-A and orexin-B are the two types of orexin coming from the same precursor peptide prepo-orexin (Sakurai et al., 2005), they bind to closely related receptors OX_1 and OX_2, the first being selective to orexin-A, while the latter being non-selective for orexin-A and orexin-B (**KPIHYUYF?**). Orexins A and B are co-located (Chou et al., 2001) and co-released (Li et al., 2014) with dynorphin. Its functions range from regulating sleep/wakefulness states (Chemelli et al., 1999) to energetic balance (**KPIHYUYF?**). Different lines of evidence suggest that orexin neurons can control foraging-related behaviors. Orexin activity promotes locomotor activity but is rapidly inhibited upon contact with food (González et al., 2016), orexin neuronal activity increases upon sucrose predictive cues (Hassani et al., 2016), and the increased spontaneous physical activity is directed towards food sources (Zink et al., 2018). This makes sense as orexin neuronal activity increases during fasting (Almeneessier et al., 2018; Diano et al., 2003; Futatsuki et al., 2018) and is inhibited by glucose and leptin (**KPIHYUYF?**). Thus, orexin-related activity can be interpreted as a food procuring signal. Further support for this interpretation comes from orexin increasing olfactory activity (Prud'homme et al., 2009), enhancing visual attention (Zajo et al., 2016), the impairment of spatial working memory in orexin knockout mice (Dang et al., 2018), among other foraging-related behaviors (**LBGV5NJ5?**).

Orexin might support foraging-related behaviors including uncertainty-driven food-seeking behavior. Orexin connectivity with VTA, LC and basal forebrain (Siegel, 2004), might modulate DA, ACh and NE activity, respectively. Such connectivity hints at a possible role of orexin in promoting food-seeking behavior via prediction-error derived expected and unexpected uncertainty. Orexin antagonist decreases dopamine activity and behavioral motivation to obtain rewards, and direct infusion of orexin into VTA increases dopamine activity and motivation to self-administer drugs (España et al., 2011; **KN9TLADJ?**) and cue-induced reinstatement (Stephen V. Mahler et al., 2013). VTA DA firing activity is modulated by orexin and dynorphin, orexin-A increases VTA increases exctibability of VTA DA neurons, while dynorophin inhibits its activity (Muschamp et al., 2014). Moreover, this modulations seems to be dependents on VTA DA neurons projections in the case of orexin but not for dynorphin; orexin increased nucleus accumbens lateral and medial projecting neurons activity, but not basolateral amygdala projecting ones, whereas dynorphin showed an inhibitory effect regardless of projections (Baimel et al., 2017). Thus hypothalamic peptides can modulate the reward signal via modulation of VTA-DA activity.

Indirect modulation of VTA activity comes from a positive feedback loop between VTA and the lateral hypothalamus. Nucleus accumbens shell exerts inhibitory activity witin the lateral hypothalamus through medium spiny neurons (Perez-Leighton et al., 2017; Qi et al., 2016), which permits rapid inhibition of food intake (O'Connor et al., 2015'Connor_Etal_2015). In turn, inhibition of

nucleus accumbens shell results in a intense feeding response (**VD2436JM?**). VTA projections to nucleus accumbens shell are mainly inhibitors of medium spiny neurons activity, thus modulating overall nucleus accumbens shell output (J. Yu et al., 2019). This loop could allow VTA to increase hypothalamic activity through by releasing nucleus accumbens shell inhibhition on the lateral hypothalamus. Additionally, DA activity, in response to environmental uncertainty, could increase lateral hypothalamus activity via dopaminergic inputs to the supramammillary nucleus, whose activity is correlated, and shows reciprocal inputs with the lateral hypothalamus (Plaisier et al., 2020). This could result in a net increase of food-seeking behavior via uncertainty-driven activity in the VTA.

In addition to directly regulating VTA DA neurons, orexin also has inputs to the LC, another brain region relevant for computing uncertainty through NE and Ach signaling. Orexin signal depolarizes the LC (Hagan et al., 1999), modulates LC response to emotionally salient information (Soya et al., 2017) and central administration of orexin increases LC activity (Soares Naufel et al., 2020). LC activity promotes task disengagement (Kane et al., 2017), altered network representation of tasks (Grella et al., 2019), and updating world models containing action-outcome pairings (Sales et al., 2019), likely via promotion of exploratory behavior related to LC tonic firing (Aston-Jones & Cohen, 2005). Then, orexin could potentially increase LC activity, which in turn can accelerate the learning of action-outcome contingencies (Glennon et al., 2019) triggered by unexpected uncertainty (Sales et al., 2019). Orexin can include internal and external information to motivate food-seeking behavior (Stephen V. Mahler et al., 2014) and its LC connectivity can allow optimal search behavior, by first including a measure of uncertainty via prediction error signal in VTA, which in turn can inform to increase exploration when unexpected uncertainty arises.

The orexin neurons could also promotes uncertainty-driven food-seeking behavior through cholinergic modulation of DA neurons in the VTA. ACh activity provides a valence-free prediction error signal which correlates with DA prediction error signal (Sturgill et al., 2020). This correlated activity can be due to a modulatory role of ACh in DA neurons through the nicotinic ACh receptors (nAChRs). nAChRs modulate DA neurons tonic and phasic activity, tonic firing mode signals overall uncertainty within a task C. D. Fiorillo (2011), whereas phasic bursting signals uncertainty-seeking behavior Jérémie Naudé et al. (2016). Thus, nAChRs modulation increases exploratory behavior when the environment is uncertain, similar to the proposed orexin role. The basal forebrain, which is one of the brain areas containing most of cholinergic neurons (Ballinger et al., 2016), receives important orexigenic inputs (Sakurai et al., 2005), and these inputs are known to be excitatory promoting ACh release (Arrigoni et al., 2010). Orexin modulation of ACh concentrations is present during exploratory behavior Fadel & Burk (2010), and in response to salient stimuli (Villano et al., 2017). Thus, increasing exploration in response to uncertainty.

In this section, we provided a plausible circuit where orexin activity acts as a hub integrating prediction error with unexpected uncertainty and expected uncertainty. This puts orexin as a candidate neuropeptide for modulating uncertainty-driven food-seeking behavior, as it can integrate environmental and internal status information to promote food-seeking behavior when necessary, while considering environmental stochasticity. We derived this plausible orexin function taking theoretical and empirical findings from foraging theory, computational models of reinforcement learning and literature on homeostatic control of food intake, allowing us to propose a functional role for orexin situated in the proper evolutionary and environmental context.

Objectives

General objective

Determine how uncertainty in food-access increases food-seeking behavior, and how orexin mediates uncertainty-driven increased food-seeking behavior.

Specific objectives

1. Determine whether uncertainty in food access required for substance increases motivation for palatable foods, and if this correlates with orexin gene expression.
2. Determine whether uncertainty in obesogenic environments increases food-seeking behavior and assess if increased food-seeking behavior correlates with orexin gene expression.
3. Determine if orexin/dynorphin neurons projecting to the VTA are active during sucrose intake.
4. Determine whether orexin in VTA elicits increased food-seeking behavior towards uncertain options, and orexin agonists inhibits food-seeking behavior towards uncertain options.

Hypothesis

General hypothesis

Food-access uncertainty increases food-seeking behavior, and this increase is modulated by orexin-neurons activity.

Specific hypothesis

1. Food-access uncertainty, in substainance food supply and in obesogenic environments, increases food-seeking behavior.
2. Prolonged exposure to food-access uncertainty environments increases prepro-orexin and orexin receptor expression in the hypothalamus.
3. Orexins and dynorphin neurons projecting to the VTA are active during sucrose intake
4. Orexin injection in VTA increases the choice of uncertain options, while orexin antagonists reduce the choice of uncertain options.

Methods

Specific hypothesis 1, Experiment 1

Animals

16 normal C57/BL6 mice will be used in this experiment. Sample size was selected by estimating the effect of uncertain access to food in operant behavior reported in (Parkes et al., 2017), with a statistical power of 80% and α=0.05. The sample size per group is estimated at n = 4. For

differential gene expression sample size was selected according to literature reporting sample sizes of 4-6 for orexin expression (Alcaraz-Iborra et al., 2014; Pankevich et al., 2010)

Experimental design

Mice (n = 16) will be housed individually with a 12/12 hr (light/dark) schedule with an automatic feeding device delivering nutritionally complete food-pellets. Animals will acclimate for a week to the feeding device. Next, mice will be split in two groups (n = 8), balanced for total food intake. The control group will maintain the same conditions as the acclimation phase, receiving pellets with a delay of 15 seconds, whereas the treatment group delay will be randomly selected from either 15, 45 seconds or 2 minutes. After 6 weeks, 4 randomly selected mice from each group will be euthanized and their brains extracted for quantitative analysis of gene expression. Remaining animals (n = 4, per group) will complete 10 sessions of a progressive ratio (PR) task for sucrose.

Random delivery of food pellets

The feeding device delivers exactly one pellet with a delay of 15 seconds once a pellet is removed from its cup. The random delay will be modeled as random samples from a truncated normal distribution with 15 second as the mean and a standard deviation of 1 minute. As such, once the food pellet is removed from the cup, in the random delay condition, the next pellet will be delivered with a mean of 15 seconds, but with a possible range from 0 to 1 minutes.

Progressive ratio task (PR)

The PR task will be conducted in a cage with two spouts, one with water and the other one with a sucrose solution, alternating positions between sessions. All sessions will last 1 h. Mice will be trained in 5 sessions to receive either water or sucrose after 5 licks (fixed-ratio 5, FR5) with a 20 second timeout after each sucrose or water delivery. In the PR task, the required number of licks to obtain sucrose will increase exponentially starting from 5 for the first sucrose reward, while water will continue in a FR5.

Quantitative analysis of gene expression

Mice will be euthanized with isoflurane. Brains will be removed, hypothalamus dissected and stored at -20C in RNAlater. RNA will be isolated by using TRIzol (Invitrogen) according to the manufacturer's instructions. Primers for prepro-orexin, orexin 1 and orexin 2 receptors will be prepared for real-time quantitative PCR according to (Lazzarino et al., 2019).

Materials and instruments

Feeding experimentation device 2 (FED 2)

The FED2 an open-source automatic feeding device, and all its parts 3D printed in polylactic acid (PLA). The device consists of two main parts: a reservoir where food pellets are stored, and a cup where food-pellets are delivered. The cup contains a photo-interrupter that detects if a pellet is within the cup, each time a pellet is removed and after a determined delay another food-pellet comes down from the reservoir. This device allows us to measure (1) the exact number of food-pellets removed (Nguyen et al. 2017), (2) the time when each of the pellets were removed, and (3) how many times the animal tried to reach for a food-pellet.

Lickometer

This device consists of a cage with two spouts through which a predetermined amount of liquid solution can be delivered. Each spout detects contact with the animal tongue, allowing us to measure the number of licks per spout, and when they were performed. Additionally the liquid solution delivery is conditioned upon a determined number of licks. This allows us to characterize the intake behavior by considering (1) the time between successive licks, (2) the distribution of the time between licks, and (3) differentiate the number of licks and the number of rewards obtained. Our pilot studies showed that between 5 to 7 sessions are sufficient for animals to stabilize their behavior and learn the operant task.

Data analysis

We want to understand if food motivation changes due to prolonged exposure to uncertainty in food-access. Food motivation will be measured as the breaking point, that is, the total number of events performed in the PR task- an event is when the animal performs the required number of licks when the spout is active. Additionally, we will analyze lick microstructure to determine for other effects related to food motivation.

Data analysis pipeline

A negative binomial generalized mixed model will be used to estimate the difference between control and treatment group in number of events. The negative binomial distribution allows us to explicitly account for individual differences in response to treatment while adjusting for variables such as weight and mean food intake. Additionally, the negative binomial distribution is the appropriate way to model count data (number of events).

Sucrose-seeking behavior will be characterized using lick microstructure analysis (A. W. Johnson et al., 2010). Lick microstructure mainly accounts for the pattern analysis of inter-lick intervals (ILI), which corresponds to the time elapsed between a determined lick and the next one. Our main patterns of interest are bursts, which correspond to rapid successive licks that are within an ILI of 500 ms between each other. Bursts are described by number and burst length, which correspond to the number of clusters under the burst definition, and how many licks compose each of those clusters, respectively. This allows us to differentiate between 'motivation' related behavior (burst number) and 'liking' response (burst size) (Alexander W. Johnson, 2018; A. W. Johnson et al., 2010). The lick microstructure features will be used as dependent variables in the negative binomial model to assess group differences in food-seeking behavior. Additionally, we will compare the main and temporal course effect of uncertainty in food-pellet intake within the home cage, effectively estimating daily caloric intake.

Using qPCR we will obtain measures of transcriptional differences between groups in prepro-orexin and orexin receptors. We aim to obtain an average relative normalized expression per group, and to statistically determine differences, we will perform a Student's t-test between groups for previously mentioned genes of interest. Statistical analysis will be performed using the bioconductor package (Gentleman et al., 2004) and the ddCt algorithm (Livak & Schmittgen, 2001).

Specific hypothesis 2, Experiment 1

Animals

8 normal C57/BL6 mice will be used in this experiment. Sample size was calculated with the same criteria of experiment 1 for differential gene expression. For behavioral effects of uncertainty, a pilot experiment determined that a sample size of 8 (4 per group) is necessary to observe a statistically significant effect.

Experimental design

Mice (n = 8) will be housed individually with a 12/12 hr (light/dark) schedule, with nutritionally complete food-pellets and water provided ad-lib. For two weeks animals will be acclimated to the lickometer device delivering a sucrose reward in both spouts. At the end of the acclimation phase animals will be splitted into two groups (n = 4) balancing on the total number of events. The control will continue with the same setup as the acclimation phase, whereas the treatment group will have one spout delivering the sucrose reward randomly 50% of the times upon 5 licks, and the other spout delivering the same sucrose reward 100% of the time for 60 minutes, alternating positions between sessions. This phase will last for 2 weeks, at the end of which animals will be euthanized with isoflurane, its brain extracted, and samples from the hypothalamus taken and prepared for RT qPCR following the protocols of the previous experiment.

Materials and instruments consist of the same lickometer described in experiment 1, variables of interest are also the same as experiment 1.

Data analysis

Our main interest is to measure how the food-seeking behavior of animals is affected by introducing uncertainty, while retaining a certain alternative. This will allow us to model an obesogenic environment where uncertainty is present. Main analyses are similar to experiment 2, but a reinforcement model is included as part of the analysis pipeline in order (1) determine the inclusion of uncertainty modifies decision processes within food-seeking behavior, and (2) compare actual food-seeking behavior to optimal models of food acquisition to determine if exploration of uncertain options is rewarding for the animal.

Data analysis pipeline

As we are expecting a general trend towards reduction of both events and licks due to sucrose devaluation from previous pilot studies, we will fit the same model as experiment 1, but including the interaction between treatment, session number and group, while adjusting for baseline number of events and licks. To model food-seeking behavior we will consider each lick as a choice for one of the two spouts.

We will use two reinforcement learning algorithms to model uncertainty effects in food-seeking behavior during the task. Epsilon-greedy choice rule will be used to model the intuition that if uncertainty does not affect reward value (i.e. the modeled estimation of the reinforcing properties of the reward), then choice should be mostly directed towards the spout with reward probability of 1, but with some stochastic deviation from this, for example, due to spout position preference. To model an effect of uncertainty in reward value we will use the Thompson sampling model, allowing us to test if uncertainty increases the reward value. Model parameters will be determined

using maximum likelihood estimation; parameter recoverability and model validation will be performed according to standard statistical techniques (Wilson et al., 2014). Model comparison will be performed using bayesian information criterion to determine the best fit to animal choice data, between epsilon-greedy and Thompson sampling models. Gene expression fold change will follow the same procedure as the previous experiment, effectively allowing us to observe if there are differences when uncertainty is introduced in an obesogenic environment.

Specific hypothesis 3, Experiment 1

Animals

18 normal C57/BL6 mice will be used in this experiment. Preliminary results from our laboratory indicate that 4 mice per group are necessary to observe Fos differences due to food intake (Coehn's $d = 0.25$, $p < 0.05$) with immunofluorescence techniques (Carolina Sandoval, unpublished data). Assuming an effect size of 80% of preliminary studies for positive control, resulting in 6 required animals (= 0.05 y = 0.8). Assuming a 80'% success rate in the bilateral tracer injection, this results in 9 required animals per experimental group, 18 in total.

Experimental design

18 mice will be injected in the VTA with a retrograde fluorescent tracer using stereotaxic surgery. Animals will be maintained with food and water ad-lib and in a 12/12 (light/dark) schedule, without any kind of intervention until they are fully recovered from surgery. Using the previously described lickometer, animals will be exposed to 30 minutes sessions with two alternating spouts delivering water and a sucrose solution upon 5 licks for treatment group ($n = 9$), and the two spouts delivering water for control group ($n = 9$). After 6 sessions, animals will be euthanized 90 minutes after the beginning of the session by isoflurane. Brain will be extracted under standard techniques for immunofluorescence to evaluate orexin neurons present in both the retrograde tracer mark and Fos neuronal activity marker.

Data analysis

This experiment will test the hypothesis that there exists the necessary functional connectivity between orexin neurons and reward processing brain structures in hedonic intake of sucrose.

Data analysis pipeline

The percent of orexin and co-released dynorphin neurons expressing Foss that are labeled with retrograde tracers will be analyzed with a two-way ANOVA, with the experimental group as a dependent variable. This will allow us to determine if this functional connectivity is present exclusively in hedonic intake compared to only water intake.

Specific hypothesis 4, Experiment 1

Animals

8 normal C57/BL6 mice will be used in this experiment. The number of animals was determined by previous pilot experiment, showing that sample size of 8 is necessary to observe the behavioral effects.

Experimental design

Mice (n = 8) will be housed individually with a 12/12 hr (light/dark) schedule, nutritionally complete food-pellets and water will be provided ad-lib. Cannulas will be implanted unilaterally aiming at the VTA, and will be left to recover without any intervention for one week. Next, for two weeks animals will be acclimated to the lickometer device, with the same setup as experiment 2 except that here, choosing any of the two spouts makes both spouts inactive until the animal makes them active by staying on top of a sensor plate located equidistant from both spouts for 1 second. After the learning phase, for 5 sessions, the spouts will be changed to one delivering a 5% sucrose solution 100% of the time, whereas the other will randomly deliver 50% of the time, alternating positions between sessions. For the next 10 sessions, spread 48 hours apart, animals will perform the same task after VTA injections of orexin-A and an orexin receptor antagonist TCS1105 (0.3 nmol / side) in random sequence for 5 sessions each. All sessions will be recorded with an infrared camera to obtain movement data.

Materials and instruments are the same as previous experiments. In addition, a sensor plate will be placed equidistant to each spout, this plate sensor allows us to enforce a trial structure to the task, where each trial begins after the animal touches the plate for 1 second and ends when a spout is chosen. Infrared video data will be processed with a custom-made image processing software, to obtain animal centroid. Allowing us to track animal movement at a resolution of ~30 frames per second. Similar to previous experiments using the lickometer, the main measures obtained are the number of events and licks per spout with a timestamp, and positional "x-y" data from video recordings.

Data analysis

Our main interest is to measure how orexin-induced activity modifies the choice ratio between uncertain and certain alternatives, in a forced-choice paradigm. Additionally, behavioral trajectories between trial start and end will inform us of locomotor activity up to decision, which can be considered as the exploratory activity.

Data analysis pipeline

We will compare the ratios of uncertain to certain options between conditions without drug, orexin antagonist and orexin agonist using a two-ways repeated measures ANOVA for choice ratio, using drug and session number as covariates. A mixed effects logistic regression will be fitted to data to obtain a more fine-grained effect of orexin in the choice between spouts. The logistic model will allow us to determine the change in odds for choosing the certain or uncertain option, while controlling for individual differences, weight, and baseline number of events. Video recordings will allow us to compute the area under the trajectory curve, which basically measures the distance between an optimal (straight) path from the sensor plate to the spout chosen and the actual path taken, so more deviation accounts for a larger area. Area under the trajectory curve can inform us about invigoration of exploitative behavior between conditions (reduction of area under the trajectory curve). Instantaneous speed and traveled distance, will allow us to infer time spent exploring and exploiting. All previous features derived from video recording will be tested using repeated measures ANOVA with Tukey post-hoc pairwise comparison, allowing us to determine differences between the drug conditions.

Work plan

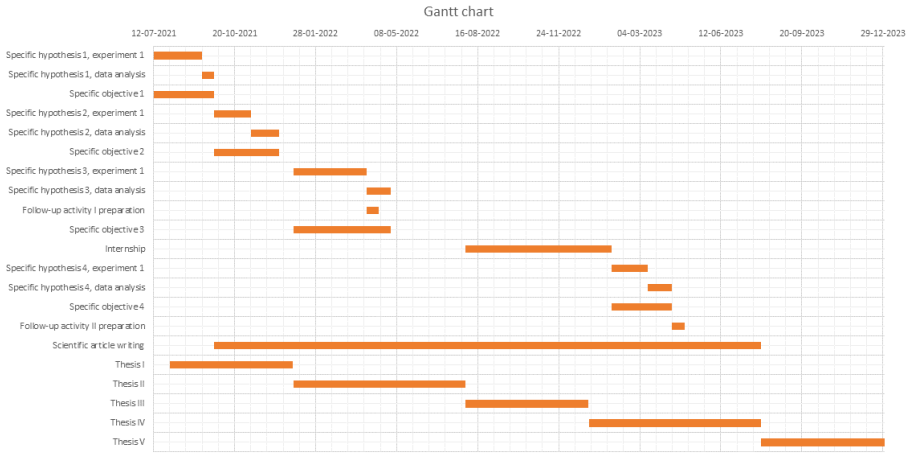

Figure 1: Work plan gantt chart, Thesis I-VI lengths are approximated, and are included to guide the time scales

Activity	Start date	Planned work days	Finish date
Specific hypothesis 1, experiment 1	7/12/2021	60	9/10/2021
Specific hypothesis 1, data analysis	9/10/2021	15	9/25/2021
Specific objective 1	7/12/2021	75	9/25/2021
Specific hypothesis 2, experiment 1	9/25/2021	45	11/9/2021
Specific hypothesis 2, data analysis	11/9/2021	35	12/14/2021
Specific objective 2	9/25/2021	80	12/14/2021
Specific hypothesis 3, experiment 1	1/1/2022	90	4/1/2022
Specific hypothesis 3, data analysis	4/1/2022	30	5/1/2022
Follow-up activity I preparation	4/1/2022	15	4/16/2022
Specific objective 3	1/1/2022	120	5/1/2022
Internship	8/1/2022	180	1/28/2023
Specific hypothesis 4, experiment 1	1/28/2023	45	3/14/2023
Specific hypothesis 4, data analysis	3/14/2023	30	4/13/2023
Specific objective 4	1/28/2023	75	4/13/2023
Follow-up activity II preparation	4/13/2023	15	4/28/2023
Scientific article writing	9/25/2021	675	8/1/2023

Figure 2: Work plan with specific dates and planned amount of work days

Figures

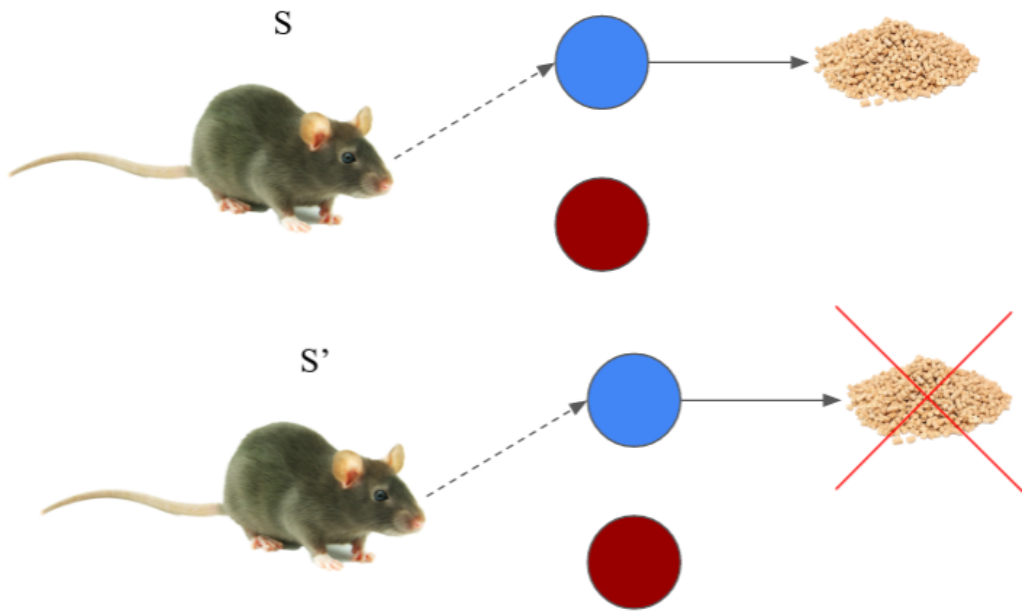

Figure 3: An animal in state s upon choosing the blue option receives a food reward. In turn, choosing the blue option changes the state to s' where choosing the same action does not deliver food. Here the value q is dependent on both action a (red or blue) and state (s or s')

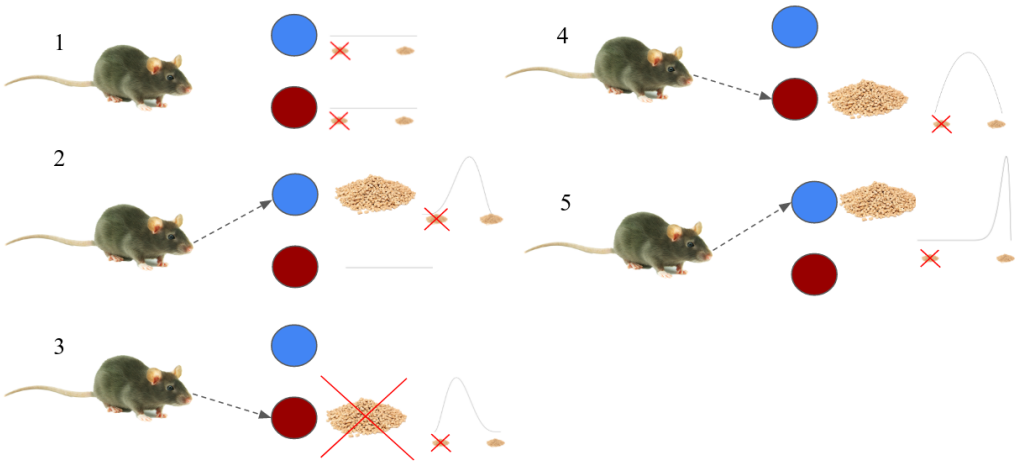

Figure 4: Creation of posterior distributions for possible actions. (1) the animal has never choosen any option so it thinks that the lever has an equal probability of given food or not. (2-3) the animal then sample both options, the blue one delivers food so the belief distribution is slighty bias toward that option delivering food, whereas the red option shows the opposite. (4) the animal re-samples the red option, but now food is delivered increasing the spread of the belief distribution, representing uncertainty on whether the red option is expected to give food or not. (5) The blue option is sampled again and food is delivered, thus increasing the bias towards the blue option delivering food, leaving less uncertainty over this option.

References

Adamantidis, A., Carter, M. C., & De Lecea, L. (2010). Optogenetic deconstruction of sleep-wake circuitry in the brain. *Frontiers in Molecular Neuroscience*, *2*. https://doi.org/10.3389/neuro.02.031.2009

Alcaraz-Iborra, M., Carvajal, F., Lerma-Cabrera, J. M., Valor, L. M., & Cubero, I. (2014). Binge-like consumption of caloric and non-caloric palatable substances in ad libitum-fed C57BL/6J mice: Pharmacological and molecular evidence of orexin involvement. *Behavioural Brain Research*, *272*, 93–99. https://doi.org/10.1016/j.bbr.2014.06.049

Almeneessier, A. S., Alzoghaibi, M., BaHammam, A. A., Ibrahim, M. G., Olaish, A. H., Nashwan, S. Z., & BaHammam, A. S. (2018). The effects of diurnal intermittent fasting on the wake-promoting neurotransmitter orexin-A. *Annals of Thoracic Medicine*, *13*(1), 48–54. https://doi.org/10.4103/atm.ATM_181_17

Anselme, P., & Güntürkün, O. (2019). How foraging works: Uncertainty magnifies food-seeking motivation. *Behavioral and Brain Sciences*, *42*, e35. https://doi.org/10.1017/S0140525X18000948

Anselme, P., Otto, T., & Güntürkün, O. (2017). How unpredictable access to food increases the body fat of small passerines: A mechanistic approach. *Behavioural Processes*, *144*, 33–45. https://doi.org/10.1016/j.beproc.2017.08.013

Anselme, P., & Robinson, M. J. F. (2019). Evidence for motivational enhancement of sign-tracking behavior under reward uncertainty. *Journal of Experimental Psychology. Animal Learning and Cognition*, *45*(3), 350–355. https://doi.org/10.1037/xan0000213

Anselme, P., Robinson, M. J. F., & Berridge, K. C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behavioural Brain Research*, *238*, 53–61. https://doi.org/10.1016/j.bbr.2012.10.006

Arrigoni, E., Mochizuki, T., & Scammell, T. E. (2010). Activation of the basal forebrain by the orexin/hypocretin neurones. *Acta Physiologica*, *198*(3), 223–235. https://doi.org/https://doi.org/10.1111/j.1748-1716.2009.02036.x

Aston-Jones, G., & Cohen, J. D. (2005). AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Smith, R. J., Sartor, G. C., Moorman, D. E., Massi, L., Tahsili-Fahadan, P., & Richardson, K. A. (2010). Lateral hypothalamic orexin/hypocretin neurons: A role in reward-seeking and addiction. *Brain Research*, *1314*, 74–90. https://doi.org/10.1016/j.brainres.2009.09.106

Baimel, C., Lau, B. K., Qiao, M., & Borgland, S. L. (2017). Projection-Target-Defined Effects of Orexin and Dynorphin on VTA Dopamine Neurons. *Cell Reports*, *18*(6), 1346–1355. https://doi.org/10.1016/j.celrep.2017.01.030

Ballinger, Elizabeth C., Ananth, M., Talmage, David A., & Role, Lorna W. (2016). Basal Forebrain Cholinergic Circuits and Signaling in Cognition and Cognitive Decline. *Neuron*, *91*(6), 1199–1218. https://doi.org/10.1016/j.neuron.2016.09.006

Bartumeus, F., Campos, D., Ryu, W. S., Lloret-Cabot, R., Méndez, V., & Catalan, J. (2016). Foraging success under uncertainty: search tradeoffs and optimal space use. *Ecology Letters*, *19*(11), 1299–1313. https://doi.org/10.1111/ele.12660

Beeler, J. A., Daw, N. D., Frazier, C. R. M., & Zhuang, X. (2010). Tonic Dopamine Modulates Exploitation of Reward Learning. *Frontiers in Behavioral Neuroscience*, *4*. https://doi.org/10.3389/fnbeh.2010.00170

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Brunstrom, J. M., & Cheon, B. K. (2018). Do humans still forage in an obesogenic environment? Mechanisms and implications for weight maintenance. *Physiology & Behavior*, *193*, 261–267. https://doi.org/10.1016/j.physbeh.2018.02.038

Burdakov, D. (2020). How orexin signals bias action: Hypothalamic and accumbal circuits. *Brain Research*, *1731*, 145943. https://doi.org/10.1016/j.brainres.2018.09.011

Campos, D., Bartumeus, F., & Méndez, V. (2013). Search times with arbitrary detection constraints. *Physical Review E*, *88*(2), 022101. https://doi.org/10.1103/PhysRevE.88.022101

Castro, D. C., & Berridge, K. C. (2017). Opioid and orexin hedonic hotspots in rat orbitofrontal cortex and insula. *Proceedings of the National Academy of Sciences*, *114*(43), E9125–E9134. https://doi.org/10.1073/pnas.1705753114

Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., & Peters, J. (2020). Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *eLife*, *9*, e51260. https://doi.org/10.7554/eLife.51260

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

Chemelli, R. M., Willie, J. T., Sinton, C. M., Elmquist, J. K., Scammell, T., Lee, C., Richardson, J. A., Williams, S. C., Xiong, Y., Kisanuki, Y., Fitch, T. E., Nakazato, M., Hammer, R. E., Saper, C. B., & Yanagisawa, M. (1999). Narcolepsy in orexin Knockout Mice: Molecular Genetics of Sleep Regulation. *Cell*, *98*(4), 437–451. https://doi.org/10.1016/S0092-8674(00)81973-X

Chou, T. C., Lee, C. E., Lu, J., Elmquist, J. K., Hara, J., Willie, J. T., Beuckmann, C. T., Chemelli, R. M., Sakurai, T., Yanagisawa, M., Saper, C. B., & Scammell, T. E. (2001). Orexin (hypocretin) neurons contain dynorphin. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(19), RC168.

Chowdhury, S., Hung, C. J., Izawa, S., Inutsuka, A., Kawamura, M., Kawashima, T., Bito, H., Imayoshi, I., Abe, M., Sakimura, K., & Yamanaka, A. (2019). Dissociating orexin-dependent and -independent functions of orexin neurons using novel Orexin-Flp knock-in mice. *eLife*, *8*, e44927. https://doi.org/10.7554/eLife.44927

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making under Risk. *Journal of Neuroscience*, *29*(40), 12574–12583. https://doi.org/10.1523/JNEUROSCI.2614-09.2009

Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A. R., & Khamassi, M. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Scientific Reports*, *9*(1), 6770. https://doi.org/10.1038/s41598-019-43245-z

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098

Dang, R., Chen, Q., Song, J., He, C., Zhang, J., Xia, J., & Hu, Z. (2018). Orexin knockout mice exhibit impaired spatial working memory. *Neuroscience Letters*, *668*, 92–97. https://doi.org/10.1016/j.neulet.2018.01.013

de Jager, M., Weissing, F. J., Herman, P. M. J., Nolet, B. A., & van de Koppel, J. (2011). Levy Walks Evolve Through Interaction Between Movement and Environmental Complexity. *Science*, *332*(6037), 1551–1553. https://doi.org/10.1126/science.1201187

Dhurandhar, E. J. (2016). The food-insecurity obesity paradox: A resource scarcity hypothesis. *Physiology & Behavior*, *162*, 88–92. https://doi.org/10.1016/j.physbeh.2016.04.025

Diano, S., Horvath, B., Urbanski, H. F., Sotonyi, P., & Horvath, T. L. (2003). Fasting activates the nonhuman primate hypocretin (orexin) system and its postsynaptic targets. *Endocrinology*, *144*(9), 3774–3778. https://doi.org/10.1210/en.2003-0274

Diederen, K. M. J., & Fletcher, P. C. (2021). Dopamine, Prediction Error and Beyond. *The Neuroscientist*, *27*(1), 30–46. https://doi.org/10.1177/1073858420907591

Dinour, L. M., Bergen, D., & Yeh, M.-C. (2007). The Food Insecurity–Obesity Paradox: A Review of the Literature and the Role Food Stamps May Play. *Journal of the American Dietetic Association*, *107*(11), 1952–1961. https://doi.org/10.1016/j.jada.2007.08.006

Drewnowski, A., & Darmon, N. (2005). The economics of obesity: dietary energy density and energy cost. *The American Journal of Clinical Nutrition*, *82*(1), 265S–273S. https://doi.org/10.1093/ajcn/82.1.265S

España, R. A., Melchior, J. R., Roberts, D. C. S., & Jones, S. R. (2011). Hypocretin 1/orexin A in the ventral tegmental area enhances dopamine responses to cocaine and promotes cocaine self-administration. *Psychopharmacology*, *214*(2), 415–426. https://doi.org/10.1007/s00213-010-2048-8

Fadel, J., & Burk, J. A. (2010). Orexin/hypocretin modulation of the basal forebrain cholinergic system: Role in attention. *Brain Research*, *1314*, 112–123. https://doi.org/10.1016/j.brainres.2009.08.046

Fagan, W. F., Gurarie, E., Bewick, S., Howard, A., Cantrell, R. S., & Cosner, C. (2017). Perceptual Ranges, Information Gathering, and Foraging Success in Dynamic Landscapes. *The American Naturalist*, *189*(5), 474–489. https://doi.org/10.1086/691099

Feder, M., & Merhav, N. (1994). Relations between entropy and error probability. *IEEE Transactions on Information Theory*, *40*(1), 259–266. https://doi.org/10.1109/18.272494

Fiorillo, C. D. (2003). Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. *Science*, *299*(5614), 1898–1902. https://doi.org/10.1126/science.1077349

Fiorillo, C. D. (2011). Transient activation of midbrain dopamine neurons by reward risk. *Neuroscience*, *197*, 162–171. https://doi.org/10.1016/j.neuroscience.2011.09.037

Fletcher, R. J., Maxwell, C. W., Andrews, J. E., & Helmey-Hartman, W. L. (2013). Signal detection theory clarifies the concept of perceptual range and its relevance to landscape connectivity. *Landscape Ecology*, *28*(1), 57–67. https://doi.org/10.1007/s10980-012-9812-6

Funamizu, A., Ito, M., Doya, K., Kanzaki, R., & Takahashi, H. (2012). Uncertainty in action-value estimation affects both action choice and learning rate of the choice behaviors of rats. *The European Journal of Neuroscience*, *35*(7), 1180–1189. https://doi.org/10.1111/j.1460-9568.2012.08025.x

Futatsuki, T., Yamashita, A., Ikbar, K. N., Yamanaka, A., Arita, K., Kakihana, Y., & Kuwaki, T. (2018). Involvement of orexin neurons in fasting- and central adenosine-induced hypothermia. *Scientific Reports*, *8*(1), 2717. https://doi.org/10.1038/s41598-018-21252-w

Garg, K., & Kello, C. T. (2021). Efficient Lévy walks in virtual human foraging. *Scientific Reports*, *11*(1), 5242. https://doi.org/10.1038/s41598-021-84542-w

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., … Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, *5*(10), R80. https://doi.org/10.1186/gb-2004-5-10-r80

Gershman, S. J. (2019). Uncertainty and exploration. *Decision*, *6*(3), 277–286. https://doi.org/10

.1037/dec0000101

Ghatak, G. (2020). A Change-Detection Based Thompson Sampling Framework for Non-Stationary Bandits. *arXiv:2009.02791 [Cs, Eess]*. http://arxiv.org/abs/2009.02791

Glennon, E., Carcea, I., Martins, A. R. O., Multani, J., Shehu, I., Svirsky, M. A., & Froemke, R. C. (2019). Locus coeruleus activation accelerates perceptual learning. *Brain Research*, *1709*, 39–49. https://doi.org/10.1016/j.brainres.2018.05.048

Gloy, K., Herrmann, M., & Fehr, T. (2020). Decision making under uncertainty in a quasi realistic binary decision task – An fMRI study. *Brain and Cognition*, *140*, 105549. https://doi.org/10.1016/j.bandc.2020.105549

Goñi, J., Aznárez-Sanado, M., Arrondo, G., Fernández-Seara, M., Loayza, F. R., Heukamp, F. H., & Pastor, M. A. (2011). The Neural Substrate and Functional Integration of Uncertainty in Decision Making: An Information Theory Approach. *PLoS ONE*, *6*(3), e17408. https://doi.org/10.1371/journal.pone.0017408

González, J. Antonio., Jensen, Lise T., Iordanidou, P., Strom, M., Fugger, L., & Burdakov, D. (2016). Inhibitory Interplay between Orexin Neurons and Eating. *Current Biology*, *26*(18), 2486–2491. https://doi.org/10.1016/j.cub.2016.07.013

Grella, S. L., Neil, J. M., Edison, H. T., Strong, V. D., Odintsova, I. V., Walling, S. G., Martin, G. M., Marrone, D. F., & Harley, C. W. (2019). Locus Coeruleus Phasic, But Not Tonic, Activation Initiates Global Remapping in a Familiar Environment. *Journal of Neuroscience*, *39*(3), 445–455. https://doi.org/10.1523/JNEUROSCI.1956-18.2018

Hagan, J. J., Leslie, R. A., Patel, S., Evans, M. L., Wattam, T. A., Holmes, S., Benham, C. D., Taylor, S. G., Routledge, C., Hemmati, P., Munton, R. P., Ashmeade, T. E., Shah, A. S., Hatcher, J. P., Hatcher, P. D., Jones, D. N. C., Smith, M. I., Piper, D. C., Hunter, A. J., … Upton, N. (1999). Orexin A activates locus coeruleus cell firing and increases arousal in the rat. *Proceedings of the National Academy of Sciences*, *96*(19), 10911–10916. https://doi.org/10.1073/pnas.96.19.10911

Hardaway, J. A., Jensen, J., Kim, M., Mazzone, C. M., Sugam, J. A., Diberto, J. F., Lowery-Gionta, E. G., Hwa, L. S., Pleil, K. E., Bulik, C. M., & Kash, T. L. (2016). Nociceptin receptor antagonist SB 612111 decreases high fat diet binge eating. *Behavioural Brain Research*, *307*, 25–34. https://doi.org/10.1016/j.bbr.2016.03.046

Harris, G. C., Wimmer, M., & Aston-Jones, G. (2005). A role for lateral hypothalamic orexin neurons in reward seeking. *Nature*, *437*(7058), 556–559. https://doi.org/10.1038/nature04071

Harris, T. R., Chapman, C. A., & Monfort, S. L. (2010). Small folivorous primate groups exhibit behavioral and physiological effects of food scarcity. *Behavioral Ecology*, *21*(1), 46–56. https://doi.org/10.1093/beheco/arp150

Hassani, O. K., Krause, M. R., Mainville, L., Cordova, C. A., & Jones, B. E. (2016). Orexin Neurons Respond Differentially to Auditory Cues Associated with Appetitive versus Aversive Outcomes. *The Journal of Neuroscience*, *36*(5), 1747–1757. https://doi.org/10.1523/JNEUROSCI.3903-15.2016

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939. https://doi.org/10.1038/nn.2856

Hess, A., Kress, S., Rakete, S., Muench, G., Kornhuber, J., Pischetsrieder, M., & Müller, C. P. (2019). Influence of the fat/carbohydrate component of snack food on energy intake pattern and reinforcing properties in rodents. *Behavioural Brain Research*, *364*, 328–333. https://doi.org/10.1016/j.bbr.2019.02.041

Humphries, N. E., Queiroz, N., Dyer, J. R. M., Pade, N. G., Musyl, M. K., Schaefer, K. M.,

Fuller, D. W., Brunnschweiler, J. M., Doyle, T. K., Houghton, J. D. R., Hays, G. C., Jones, C. S., Noble, L. R., Wearmouth, V. J., Southall, E. J., & Sims, D. W. (2010). Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature*, *465*(7301), 1066–1069. https://doi.org/10.1038/nature09116

Humphries, N. E., & Sims, D. W. (2014). Optimal foraging strategies: Lévy walks balance searching and patch exploitation under a very broad range of conditions. *Journal of Theoretical Biology*, *358*, 179–193. https://doi.org/10.1016/j.jtbi.2014.05.032

Johnson, Alexander W. (2018). Characterizing ingestive behavior through licking microstructure: Underlying neurobiology and its use in the study of obesity in animal models. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, *64*, 38–47. https://doi.org/10.1016/j.ijdevneu.2017.06.012

Johnson, A. W., Sherwood, A., Smith, D. R., Wosiski-Kuhn, M., Gallagher, M., & Holland, P. C. (2010). An analysis of licking microstructure in three strains of mice. *Appetite*, *54*(2), 320–330. https://doi.org/10.1016/j.appet.2009.12.007

Kacelnik, A., & Bateson, M. (1996). Risky Theories—The Effects of Variance on Foraging Decisions. *American Zoologist*, *36*(4), 402–434. https://doi.org/10.1093/icb/36.4.402

Kane, G. A., Vazey, E. M., Wilson, R. C., Shenhav, A., Daw, N. D., Aston-Jones, G., & Cohen, J. D. (2017). Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(6), 1073–1083. https://doi.org/10.3758/s13415-017-0531-y

Kembro, J. M., Lihoreau, M., Garriga, J., Raposo, E. P., & Bartumeus, F. (2019). Bumblebees learn foraging routes through exploitation–exploration cycles. *Journal of The Royal Society Interface*, *16*(156), 20190103. https://doi.org/10.1098/rsif.2019.0103

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696–12701. https://doi.org/10.1073/pnas.1715293114

King, S. J., Rodrigues, T., Watts, A., Murray, E., Wilson, A., & Abizaid, A. (2016). Investigation of a role for ghrelin signaling in binge-like feeding in mice under limited access to high-fat diet. *Neuroscience*, *319*, 233–245. https://doi.org/10.1016/j.neuroscience.2016.01.004

Kölzsch, A., Alzate, A., Bartumeus, F., de Jager, M., Weerman, E. J., Hengeveld, G. M., Naguib, M., Nolet, B. A., & van de Koppel, J. (2015). Experimental evidence for inherent Lévy search behaviour in foraging animals. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1807), 20150424. https://doi.org/10.1098/rspb.2015.0424

Kosse, C., & Burdakov, D. (2014). A unifying computational framework for stability and flexibility of arousal. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00192

Kramer, D. L. (2001). Foraging Behavior. In *Evolutionary Ecology*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780195131543.001.0001/isbn-9780195131543-book-part-24

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences*, *111*(6), 2343–2348. https://doi.org/10.1073/pnas.1321596111

Lardeux, S., Kim, J. J., & Nicola, S. M. (2013). Intermittent access to sweet high-fat liquid induces increased palatability and motivation to consume in a rat model of binge consumption. *Physiology & Behavior*, *114-115*, 21–31. https://doi.org/10.1016/j.physbeh.2013.03.005

LaScala-Gruenewald, D. E., Mehta, R. S., Liu, Y., & Denny, M. W. (2019). Sensory perception plays a larger role in foraging efficiency than heavy-tailed movement strategies. *Ecological*

Modelling, *404*, 69–82. https://doi.org/10.1016/j.ecolmodel.2019.02.015

Lazzarino, G. P., Acutain, M. F., Canesini, G., Andreoli, M. F., & Ramos, J. G. (2019). Cafeteria diet induces progressive changes in hypothalamic mechanisms involved in food intake control at different feeding periods in female rats. *Molecular and Cellular Endocrinology*, *498*, 110542. https://doi.org/10.1016/j.mce.2019.110542

Le Heron, C., Kolling, N., Plant, O., Kienast, A., Janska, R., Ang, Y.-S., Fallon, S., Husain, M., & Apps, M. A. J. (2020). Dopamine Modulates Dynamic Decision-Making during Foraging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *40*(27), 5273–5282. https://doi.org/10.1523/JNEUROSCI.2586-19.2020

Li, X., Marchant, N. J., & Shaham, Y. (2014). Opposing roles of cotransmission of dynorphin and hypocretin on reward and motivation. *Proceedings of the National Academy of Sciences*, *111*(16), 5765–5766. https://doi.org/10.1073/pnas.1403603111

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Luby, M. D., Hsu, C. T., Shuster, S. A., Gallardo, C. M., Mistlberger, R. E., King, O. D., & Steele, A. D. (2012). Food Anticipatory Activity Behavior of Mice across a Wide Range of Circadian and Non-Circadian Intervals. *PLoS ONE*, *7*(5), e37992. https://doi.org/10.1371/journal.pone.0037992

Ma, W. J., & Jazayeri, M. (2014). Neural Coding of Uncertainty and Probability. *Annual Review of Neuroscience*, *37*(1), 205–220. https://doi.org/10.1146/annurev-neuro-071013-014017

Mahler, Stephen V., Moorman, D. E., Smith, R. J., James, M. H., & Aston-Jones, G. (2014). Motivational activation: a unifying hypothesis of orexin/hypocretin function. *Nature Neuroscience*, *17*(10), 1298–1303. https://doi.org/10.1038/nn.3810

Mahler, Stephen V., Smith, R. J., & Aston-Jones, G. (2013). Interactions between VTA orexin and glutamate in cue-induced reinstatement of cocaine seeking in rats. *Psychopharmacology*, *226*(4), 687–698. https://doi.org/10.1007/s00213-012-2681-5

Marshall, L., Mathys, C., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological Fingerprints of Contextual Uncertainty. *PLOS Biology*, *14*(11), e1002575. https://doi.org/10.1371/journal.pbio.1002575

Maye, A., Hsieh, C., Sugihara, G., & Brembs, B. (2007). Order in Spontaneous Behavior. *PLoS ONE*, *2*(5), e443. https://doi.org/10.1371/journal.pone.0000443

McGee, H. M., Amare, B., Bennett, A. L., & Duncan-Vaidya, E. A. (2010). Behavioral effects of withdrawal from sweetened vegetable shortening in rats. *Brain Research*, *1350*, 103–111. https://doi.org/10.1016/j.brainres.2010.01.033

Meister, B. (2000). Control of food intake via leptin receptors in the hypothalamus. In *Vitamins & Hormones* (Vol. 59, pp. 265–304). Academic Press. https://www.sciencedirect.com/science/article/pii/S0083672900590104

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/JNEUROSCI.16-05-01936.1996

Moradi, S., Mirzababaei, A., Dadfarma, A., Rezaei, S., Mohammadi, H., Jannat, B., & Mirzaei, K. (2019). Food insecurity and adult weight abnormality risk: a systematic review and meta-analysis. *European Journal of Nutrition*, *58*(1), 45–61. https://doi.org/10.1007/s00394-018-1819-6

Müller, T. D., Nogueiras, R., Andermann, M. L., Andrews, Z. B., Anker, S. D., Argente, J., Batterham, R. L., Benoit, S. C., Bowers, C. Y., Broglio, F., Casanueva, F. F., D'Alessio,

D., Depoortere, I., Geliebter, A., Ghigo, E., Cole, P. A., Cowley, M., Cummings, D. E., Dagher, A., … Tschöp, M. H. (2015). Ghrelin. *Molecular Metabolism*, *4*(6), 437–460. https://doi.org/10.1016/j.molmet.2015.03.005

Muñoz-Escobar, G., Guerrero-Vargas, N. N., & Escobar, C. (2019). Random access to palatable food stimulates similar addiction-like responses as a fixed schedule, but only a fixed schedule elicits anticipatory activation. *Scientific Reports*, *9*(1), 18223. https://doi.org/10.1038/s41598-019-54540-0

Muschamp, J. W., Hollander, J. A., Thompson, J. L., Voren, G., Hassinger, L. C., Onvani, S., Kamenecka, T. M., Borgland, S. L., Kenny, P. J., & Carlezon, W. A. (2014). Hypocretin (orexin) facilitates reward by attenuating the antireward effects of its cotransmitter dynorphin in ventral tegmental area. *Proceedings of the National Academy of Sciences*, *111*(16), E1648–E1655. https://doi.org/10.1073/pnas.1315542111

Nakamura, K., & Ono, T. (1986). Lateral hypothalamus neuron involvement in integration of natural and artificial rewards and cue signals. *Journal of Neurophysiology*, *55*(1), 163–181. https://doi.org/10.1152/jn.1986.55.1.163

Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The Dopamine Prediction Error: Contributions to Associative Models of Reward Learning. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00244

Naudé, J., Didienne, S., Takillah, S., Prévost-Solié, C., Maskos, U., & Faurej, P. (2018). *Acetylcholine-dependent phasic dopamine activity signals exploratory locomotion and choices*. Neuroscience. http://biorxiv.org/lookup/doi/10.1101/242438

Naudé, Jérémie, Tolu, S., Dongelmans, M., Torquet, N., Valverde, S., Rodriguez, G., Pons, S., Maskos, U., Mourot, A., Marti, F., & Faure, P. (2016). Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking. *Nature Neuroscience*, *19*(3), 471–478. https://doi.org/10.1038/nn.4223

Nauta, J., Khaluf, Y., & Simoens, P. (2020). Hybrid foraging in patchy environments using spatial memory. *Journal of The Royal Society Interface*, *17*(166), 20200026. https://doi.org/10.1098/rsif.2020.0026

Nettle, D., & Bateson, M. (2019). Food-Insecure Women Eat a Less Diverse Diet in a More Temporally Variable Way: Evidence from the US National Health and Nutrition Examination Survey, 2013-4. *Journal of Obesity*, *2019*, 1–9. https://doi.org/10.1155/2019/7174058

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

O'Connor, Eoin C., Kremer, Y., Lefort, S., Harada, M., Pascoli, V., Rohner, C., & Lüscher, C. (2015). Accumbal D1R Neurons Projecting to Lateral Hypothalamus Authorize Feeding. *Neuron*, *88*(3), 553–564. https://doi.org/10.1016/j.neuron.2015.09.038

Pandit, R., Beerens, S., & Adan, R. a. H. (2017). Role of leptin in energy expenditure: the hypothalamic perspective. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *312*(6), R938–R947. https://doi.org/10.1152/ajpregu.00045.2016

Pankevich, D. E., Teegarden, S. L., Hedin, A. D., Jensen, C. L., & Bale, T. L. (2010). Caloric Restriction Experience Reprograms Stress and Orexigenic Pathways and Promotes Binge Eating. *Journal of Neuroscience*, *30*(48), 16399–16407. https://doi.org/10.1523/JNEUROSCI.1955-10.2010

Parkes, S. L., Furlong, T. M., Black, A. D., & Balleine, B. W. (2017). Intermittent feeding alters sensitivity to changes in reward value. *Appetite*, *113*, 1–6. https://doi.org/10.1016/j.appet.2017.02.009

Perez-Leighton, C., Little, M. R., Grace, M., Billington, C., & Kotz, C. M. (2017). Orexin signal-

ing in rostral lateral hypothalamus and nucleus accumbens shell in the control of spontaneous physical activity in high- and low-activity rats. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *312*(3), R338–R346. https://doi.org/10.1152/ajpregu.00339.2016

Peyron, C., Tighe, D. K., van den Pol, A. N., de Lecea, L., Heller, H. C., Sutcliffe, J. G., & Kilduff, T. S. (1998). Neurons Containing Hypocretin (Orexin) Project to Multiple Neuronal Systems. *The Journal of Neuroscience*, *18*(23), 9996–10015. https://doi.org/10.1523/JNEUROSCI.18-23-09996.1998

Piray, P., & Daw, N. D. (2020). A simple model for learning in volatile environments. *PLoS computational biology*, *16*(7), e1007963. https://doi.org/10.1371/journal.pcbi.1007963

Plaisier, F., Hume, C., & Menzies, J. (2020). Neural connectivity between the hypothalamic supramammillary nucleus and appetite- and motivation-related regions of the rat brain. *Journal of Neuroendocrinology*, *32*(2), e12829. https://doi.org/https://doi.org/10.1111/jne.12829

Prud'homme, M. J., Lacroix, M. C., Badonnel, K., Gougis, S., Baly, C., Salesse, R., & Caillol, M. (2009). Nutritional status modulates behavioural and olfactory bulb Fos responses to isoamyl acetate or food odour in rats: roles of orexins and leptin. *Neuroscience*, *162*(4), 1287–1298. https://doi.org/10.1016/j.neuroscience.2009.05.043

Pyke, G. H. (1984). Optimal Foraging Theory: A Critical Review. *Annual Review of Ecology and Systematics*, *15*(1), 523–575. https://doi.org/10.1146/annurev.es.15.110184.002515

Pyke, G. H. (2010). Optimal Foraging and Plant–Pollinator Co-Evolution. In *Encyclopedia of Animal Behavior* (pp. 596–600). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780080453378002783

Qi, J., Zhang, S., Wang, H.-L., Barker, D. J., Miranda-Barrientos, J., & Morales, M. (2016). VTA glutamatergic inputs to nucleus accumbens drive aversion by acting on GABAergic interneurons. *Nature Neuroscience*, *19*(5), 725–733. https://doi.org/10.1038/nn.4281

Rada, P., Avena, N. M., & Hoebel, B. G. (2005). Daily bingeing on sugar repeatedly releases dopamine in the accumbens shell. *Neuroscience*, *134*(3), 737–744. https://doi.org/10.1016/j.neuroscience.2005.04.043

Raj, V., & Kalyani, S. (2017). Taming Non-stationary Bandits: A Bayesian Approach. *arXiv:1707.09727 [Cs, Stat]*. http://arxiv.org/abs/1707.09727

Ranc, N., Moorcroft, P. R., Ossi, F., & Cagnacci, F. (2021). Experimental evidence of memory-based foraging decisions in a large wild mammal. *Proceedings of the National Academy of Sciences*, *118*(15), e2014856118. https://doi.org/10.1073/pnas.2014856118

Raposo, E. P., Bartumeus, F., da Luz, M. G. E., Ribeiro-Neto, P. J., Souza, T. A., & Viswanathan, G. M. (2011). How Landscape Heterogeneity Frames Optimal Diffusivity in Searching Processes. *PLoS Computational Biology*, *7*(11), e1002233. https://doi.org/10.1371/journal.pcbi.1002233

Rescorla, R. A., Wagner, A. R., Black, A. H., & Prokasy, W. F. (1972). *Classical conditioning II: current research and theory*.

Reverdy, P. B., Srivastava, V., & Leonard, N. E. (2014). Modeling Human Decision Making in Generalized Gaussian Multiarmed Bandits. *Proceedings of the IEEE*, *102*(4), 544–571. https://doi.org/10.1109/JPROC.2014.2307024

Reynolds, A. M., & Frye, M. A. (2007). Free-Flight Odor Tracking in Drosophila Is Consistent with an Optimal Intermittent Scale-Free Search. *PLoS ONE*, *2*(4), e354. https://doi.org/10.1371/journal.pone.0000354

Reynolds, A., Ceccon, E., Baldauf, C., Karina Medeiros, T., & Miramontes, O. (2018). Lévy foraging patterns of rural humans. *PLOS ONE*, *13*(6), e0199099. https://doi.org/10.1371/journal.pone.0199099

Rothwell, N. J., & Stock, M. J. (1988). The Cafeteria Diet as a Tool for Studies of Thermogenesis. *The Journal of Nutrition*, *118*(8), 925–928. https://doi.org/10.1093/jn/118.8.925

Ryali, C. K., Reddy, G., & Yu, A. J. (2016). *Demystifying excessively volatile human learning: A Bayesian persistent prior and a neural approximation*. Neuroscience. http://biorxiv.org/lookup/doi/10.1101/077719

Sakurai, T., Nagata, R., Yamanaka, A., Kawamura, H., Tsujino, N., Muraki, Y., Kageyama, H., Kunita, S., Takahashi, S., Goto, K., Koyama, Y., Shioda, S., & Yanagisawa, M. (2005). Input of Orexin/Hypocretin Neurons Revealed by a Genetically Encoded Tracer in Mice. *Neuron*, *46*(2), 297–308. https://doi.org/10.1016/j.neuron.2005.03.010

Sales, A. C., Friston, K. J., Jones, M. W., Pickering, A. E., & Moran, R. J. (2019). Locus Coeruleus tracking of prediction errors optimises cognitive flexibility: An Active Inference model. *PLOS Computational Biology*, *15*(1), e1006267. https://doi.org/10.1371/journal.pcbi.1006267

Schiffer, A.-M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012). Surprised at All the Entropy: Hippocampal, Caudate and Midbrain Contributions to Learning from Prediction Errors. *PLoS ONE*, *7*(5), e36445. https://doi.org/10.1371/journal.pone.0036445

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Siegel, J. M. (2004). Hypocretin (OREXIN): Role in Normal Behavior and Neuropathology. *Annual Review of Psychology*, *55*(1), 125–148. https://doi.org/10.1146/annurev.psych.55.090902.141545

Sims, D. W., Humphries, N. E., Hu, N., Medan, V., & Berni, J. (2019). Optimal searching behaviour generated intrinsically by the central pattern generator for locomotion. *eLife*, *8*, e50316. https://doi.org/10.7554/eLife.50316

Soares Naufel, M. F., Martin Truzzi, G. de, & Santos Coelho, F. M. (2020). Hypocretin (Orexin): What It Does and How It Links With Narcolepsy and Food Choices. In *Neurological Modulation of Sleep* (pp. 159–168). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128166581000168

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, *20*(10), 635–644. https://doi.org/10.1038/s41583-019-0180-y

Soya, S., Takahashi, T. M., McHugh, T. J., Maejima, T., Herlitze, S., Abe, M., Sakimura, K., & Sakurai, T. (2017). Orexin modulates behavioral fear expression through the locus coeruleus. *Nature Communications*, *8*(1), 1606. https://doi.org/10.1038/s41467-017-01782-z

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Stinson, E. J., Votruba, S. B., Venti, C., Perez, M., Krakoff, J., & Gluck, M. E. (2018). Food Insecurity is Associated with Maladaptive Eating Behaviors and Objectively Measured Overeating: Food Insecurity and Overeating. *Obesity*, *26*(12), 1841–1848. https://doi.org/10.1002/oby.22305

Stolyarova, A., & Izquierdo, A. (2017). Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *eLife*, *6*, e27483. https://doi.org/10.7554/eLife.27483

Sturgill, J. F., Hegedus, P., Li, S. J., Chevy, Q., Siebels, A., Jing, M., Li, Y., Hangya, B., & Kepecs, A. (2020). Basal forebrain-derived acetylcholine encodes valence-free reinforcement prediction error. *bioRxiv*, 2020.02.17.953141. https://doi.org/10.1101/2020.02.17.953141

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: an introduction* (Second edition).

The MIT Press.

Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., & Schoenbaum, G. (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. *Neuron*, *95*(6), 1395–1405.e3. https://doi.org/10.1016/j.neuron.2017.08.025

Thompson, W. R. (1933). On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika*, *25*(3/4), 285. https://doi.org/10.2307/2332286

Tian, J., Huang, R., Cohen, Jeremiah Y., Osakada, F., Kobak, D., Machens, Christian K., Callaway, Edward M., Uchida, N., & Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. *Neuron*, *91*(6), 1374–1389. https://doi.org/10.1016/j.neuron.2016.08.018

Tokic, M., & Palm, G. (2011). Value-Difference Based Exploration: Adaptive Control between Epsilon-Greedy and Softmax. In J. Bach & S. Edelkamp (Eds.), *KI 2011: Advances in Artificial Intelligence* (Vol. 7006, pp. 335–346). Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-642-24455-1_33

van Holstein, M., & Floresco, S. B. (2020). Dissociable roles for the ventral and dorsal medial prefrontal cortex in cue-guided risk/reward decision making. *Neuropsychopharmacology*, *45*(4), 683–693. https://doi.org/10.1038/s41386-019-0557-7

Villano, I., Messina, A., Valenzano, A., Moscatelli, F., Esposito, T., Monda, V., Esposito, M., Precenzano, F., Carotenuto, M., Viggiano, A., Chieffi, S., Cibelli, G., Monda, M., & Messina, G. (2017). Basal Forebrain Cholinergic System and Orexin Neurons: Effects on Attention. *Frontiers in Behavioral Neuroscience*, *11*. https://doi.org/10.3389/fnbeh.2017.00010

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., & Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature*, *381*(6581), 413–415. https://doi.org/10.1038/381413a0

Wang, Z., & Zhou, M. (2020). Thompson Sampling via Local Uncertainty. *arXiv:1910.13673 [Cs, Stat]*. http://arxiv.org/abs/1910.13673

Wearmouth, V. J., McHugh, M. J., Humphries, N. E., Naegelen, A., Ahmed, M. Z., Southall, E. J., Reynolds, A. M., & Sims, D. W. (2014). Scaling laws of ambush predator 'waiting' behaviour are tuned to a common ecology. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782), 20132997. https://doi.org/10.1098/rspb.2013.2997

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Wojnicki, F. H. E., Babbs, R. K., & Corwin, R. L. W. (2013). Environments predicting intermittent shortening access reduce operant performance but not home cage binge size in rats. *Physiology & Behavior*, *116-117*, 35–43. https://doi.org/10.1016/j.physbeh.2013.03.015

Wojnicki, F. H. E., Johnson, D. S., Charny, G., & Corwin, R. L. W. (2015). Development of bingeing in rats altered by a small operant requirement. *Physiology & Behavior*, *152*, 112–118. https://doi.org/10.1016/j.physbeh.2015.09.009

Wojnicki, F. H. E., Stine, J. G., & Corwin, R. L. W. (2007). Liquid sucrose bingeing in rats depends on the access schedule, concentration and delivery system. *Physiology & Behavior*, *92*(4), 566–574. https://doi.org/10.1016/j.physbeh.2007.05.002

Wu, Q., Iyer, N., & Wang, H. (2018). Learning Contextual Bandits in a Non-stationary Environment. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 495–504. https://doi.org/10.1145/3209978.3210051

Yoon, H.-J., Lee, D., & Hovakimyan, N. (2018). Hidden Markov Model Estimation-Based Q-

learning for Partially Observable Markov Decision Process. *arXiv:1809.06401 [Cs, Stat]*. http://arxiv.org/abs/1809.06401

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692. https://doi.org/10.1016/j.neuron.2005.04.026

Yu, J., Ishikawa, M., Wang, J., Schlüter, O. M., Sesack, S. R., & Dong, Y. (2019). Ventral Tegmental Area Projection Regulates Glutamatergic Transmission in Nucleus Accumbens. *Scientific Reports*, *9*(1), 18451. https://doi.org/10.1038/s41598-019-55007-y

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science (New York, N.Y.)*, *323*(5920), 1496–1499. https://doi.org/10.1126/science.1167342

Zajo, K. N., Fadel, J. R., & Burk, J. A. (2016). Orexin A-induced enhancement of attentional processing in rats: role of basal forebrain neurons. *Psychopharmacology*, *233*(4), 639–647. https://doi.org/10.1007/s00213-015-4139-z

Zink, A. N., Bunney, P. E., Holm, A. A., Billington, C. J., & Kotz, C. M. (2018). Neuromodulation of orexin neurons reduces diet-induced adiposity. *International Journal of Obesity*, *42*(4), 737–745. https://doi.org/10.1038/ijo.2017.276