# Introduction

Mammals evolved in an environment where food sources are limited and often scarce; 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 food access uncertainty predicts future food scarcity, triggering food-seeking behavior to prevent starvation. This prediction effectively allows animals to act upon proximal cues without the need to know the complete state of the environment.

Currently, it is unclear how animals use uncertainty to drive food-seeking behavior. 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. Then, exploratory behavior should increase proportionally to the prediction error.

Relating uncertainty to food scarcity is a successful adaptive strategy when food is limited, but food sources are ubiquitous in modern environments. These environments increase food-seeking behavior leading to excessive intake, raising the risk of overweight and obesity. Even though the effects of uncertainty (through intermittent diet schemes) are known to increase food intake in a binge-like fashion, the specific behavioral events and the neurobiological substrate mediating these effects remain unclear.

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 its activation increases physical activity and food intake, but this effect seems to be brain site-specific. Together, these data support that orexin might promote foraging behavior by increasing locomotor activity in response to a nutrient deficit before food intake. Orexin could provide the mechanism to drive increased food-seeking behavior when uncertainty increases.

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-drive increased food-seeking behavior.

# General research strategy

To determine how uncertainty in food-access increases food-seeking behavior we will use mainly two strategies: (1) behavioral modeling using the reinforcement learning framework to assess changes in reward processing in conditions with uncertainty compared to those without 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 through 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), t 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 are all behaviors related to obtaining food, including feeding and hoarding (Kramer, 2001). Within foraging, food-seeking behavior precedes actual intake and is the re-orientation of attention and locomotor activity to acquire food. A successful food-seeking behavior uses the least amount of resources while reducing exposure to potential predators to find food. Food-seeking is a dynamic behavior: the environment is constantly changing along with food locations, so strategies must constantly adapt to obtain food (Bartumeus et al., 2016; Kölzsch et al., 2015). Moreover, animals have incomplete knowledge about location, quality, and the probability of obtaining food (G. H. Pyke, 1984 ). Thus, a foraging animal must determine its food-seeking behavior with only partial knowledge of where food is located.

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 such resources, and what is the expected outcome of foraging here or there. These properties should not be considered in isolation. For example, not considering changes in food quality could lead to a search for food in an area with an overall low food quality, increasing the risk of starvation. This problem begs the question: what kind of strategy could accomplish this while having only partial or incomplete knowledge about the food environment?

Distribution fitting of movement data of different species (including human) has shown that food-seeking behavior tends to follow 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 describes multiple concentrated movements with sharp turning angles followed by few ballistic displacements. These patterns produce optimal searches in various environments with a patchy distribution of resources (clumped resources distant to one another) (Wosniack et al., 2017). Although how animals produce Lévy search patterns is unknown (A. Reynolds et al., 2018), this might be partially independent of sensory information (Humphries & Sims, 2014; Sims et al., 2019). This strategy probably evolved to optimize food searching with partial or complete lack of knowledge of food location (Wosniack et al., 2017).

Lévy-walks is a ubiquitous strategy that does not rely on sensory information, and might be an ancient ‘base’ strategy to deal with locating food in uncertain environments (Kölzsch et al., 2015; Sims et al., 2019). However, after a food encounter, the search strategy switches to a more focused one similar to Brownian-motion, following random trajectories within a more enclosed space, possibly reflecting an adaptation to the usual distribution of food within patches: if one encounters food there is a good chance that more food will be nearby (Nauta et al., 2020; A. M. Reynolds & Frye, 2007). This suggests that food seeking behavior likely evolved to deal with food location uncertainty that could not be reduced by perceptual information.

Here, we proposed that Levy-walks could capture the intuition that food-seeking behavior relies on strategies adapted to deal with food location uncertainty. Having a ‘base’ strategy that does not rely on perceptual information might seem to end in an inflexible strategy. However, uncertainty itself can modify food-seeking vigor (Anselme et al., 2013a), increasing the number of food-seeking bouts (Anselme et al., 2017; Anselme & Robinson, 2019). In the next section, we will explore how animals sense uncertainty and why this is adaptive when perceptual abilities are limited.

## Foraging and uncertainty

As discussed, animals likely adapted its food-seeking behavior to find food within an uncertain food environment. Here, we will show how environmental uncertainty can regulate food-seeking behavior, and that uncertainty can be processed without the need for extensive computation in certain brain structures.

An animal that does not consider the stochastic nature of its food environment in its food-seeking behavior will act greedily upon sampled values. Acting greedily implies that the animal will always choose the option that yielded the most value in an initial random sampling, which neglects that the samples come from a stochastic reward process that changes over time. This is problematic as it makes food-seeking strategies insensitive to reward variation Sutton & Barto (2018)]. To improve upon the greedy case one could propose that animals act upon the mean rewards, such as the classical model by Charnov (1976), which determines the strategy based on the current patch value against the global mean of the environment. Nevertheless, such ‘rule-based’ models do not provide good fit to behavioral data (Hayden et al., 2011; Le Heron et al., 2020; Nonacs, 2001; G. H. Pyke, 2010). ‘Rule-based’ models propose that animals guide their behavior by choosing foraging areas where the mean rate of food acquisition is equal or larger than the global average (e.g. Charnov (1976)). However, due to limited perceptual abilities and environment stochastic nature, keeping track of a global average is not possible. An alternative is that animals do not keep track of the global average, but the uncertainty of their food environment. Nevertheless, to propose that perceived uncertainty modulates food-seeking behavior requires that animals have the ability to track uncertainty-related features.

One way of incorporating uncertainty is through risk, which considers the spread (i.e. standard deviation) of possible outcomes (Rothwell & Stock, 1988). In humans, the anterior cingulate cortex (ACC) activity tracks risk in decision tasks (Christopoulos et al., 2009). Moreover, ACC is able to track risk in a context-dependent fashion: it considers cue-related information to determine the risk and expected value of a given option (van Holstein & Floresco, 2020). While risk is the component of uncertainty that measures the spread of outcome, volatility indicates how often an environment changes how an action and an outcome are related, which is also readily tracked by ACC (Behrens et al., 2007). The main goal of tracking uncertainty is to augment the chance of success while searching for food. In that regard, filtering out uncertainty regarding outcomes can prove beneficial, as the actual outcome prediction does not become affected by noise. This means that if a given area is likely to provide good quality and amount of food, we should keep searching for food in that area, even when sometimes our search results are bad. However, how many bad outcomes we should tolerate can only be determined if we know the associated outcome uncertainty. Stolyarova & Izquierdo (2017) showed that rats could optimally choose options with more value (less mean reward waiting time) despite large associated variability. Furthermore, lesions in the orbitofrontal cortex (OFC) showed an impaired ability to change preference when the mean rewards were up or downshifted; that is, value inference became noisy. Together, these data show that measures of uncertainty, such as risk and volatility can be tracked by brain structures, and this contributes to determining the expected value of an option.

Being able to track uncertainty can provide information about how to behave in any given environment. Thus, an increase in the variability of the value of a food source should trigger an increase in exploratory behavior to find new food sources and update the expected value of the current food source. The intuition that an animal optimizes its food choices according to the mean expected is captured in the models of expected and unexpected uncertainty Yu & Dayan (2005), they propose expected uncertainty as the uncertainty regarding outcomes when contingencies (outcome given a particular action) remains stable but its subject to some noise, and unexpected uncertainty representing a drastic change in the contingencies that likely reflects a structural change in the environment. More precisely, the expected/unexpected uncertainty model proposes that if the obtained rewards present some variation it should not modify learned contingencies that map actions with rewards in a given environment, so top-down control (where value representations are) should dominate over bottom-up input (sensory information). On the other hand, if obtained rewards present a large amount of variation, the balance should switch in order to prioritize bottom-up input increasing learning about new contingencies (Soltani & Izquierdo, 2019; Yu & Dayan, 2005).

An interesting observation is that even when tasks have no abrupt change in the reward contingencies, animals behave as if the task was non-stationary (Wu et al., 2018), that is, as if the contingencies do indeed presented abrupt changes, increasing exploratory behavior (Ryali et al., 2016). While the functional reasons behind this behavior is not clear, it has been proposed that they emerge because sustaining the belief that environments are non-stationary does not prove to be problematic in stationary ones (Ryali et al., 2016) or because stochasticity in decisions may provide a sufficient heuristic in many natural environments (Reverdy et al., 2014) which are effectively non-stationary. This points that animals do not only use uncertainty to properly guide its behavior, but assume that all environments are uncertain.

Under non-stationary environments, contingencies change, so animals are faced with the dilemma of either exploiting or exploring (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 tries to re-sample the environment to improve or re-learn current contingencies. The exploitation-exploration balance has been linked to the expected and unexpected uncertainty (Cohen et al., 2007), as unexpected uncertainty should increase exploratory behavior to boost learning of new contingencies, whereas if only expected uncertainty is present, behavior should exploit current knowledge. However, exploration could also be triggered by ‘boredom’ when environment properties are extensible learned (Aston-Jones & Cohen, 2005).

Here we discussed that an animal senses the uncertainty in its food environment, and this information regulates food-seeking behavior. That animals act as if the food environment is uncertain and non-stationary even if this is not the case, further supports the idea of a deeply rooted strategy to deal with uncertain environments. One of the key assumptions behind this idea is that sensing uncertainty is useful because the perceptual abilities of an animal are limited. Yet, there is the position that perceptual abilities of animals are sufficient to inform food seeking strategies and that perception of environmental uncertainty is unnecessary to adapt food-seeking behavior. Next, we will argue in favor that perceptual abilities are not sufficient to adapt food-seeking behavior and that perception of uncertainty is necessary.

Animals use perceptual information to track for food, informing where to go next. However, perceptual information is alway limited in some regard, establishing a perceptual range where food location information can be obtained with enough quality to inform food-seeking strategies (Fletcher et al., 2013). If an animal must know what is beyond such perceptual range, displacement is needed. However, to inform such displacement memory, perceptual information integration or other cognitive processes should come into play, because displacement from the perceptual range necessarily stops the animals from sensing the previous perceptual range. In many classical models, animals are assumed to somehow integrate information about resource quality and distribution into an environmental mean that is kept on memory (Charnov, 1976). However, even if this was possible, integration of environmental information is dependent on time, as the environment changes information must be constantly updated. So integrating information over time, such as how food availability changes, can prove to be helpful for the foraging animal (Fagan et al., 2017), because it includes information about the temporal dimension.

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, which at the same time can modify the quality of information obtained. The speed-perception tradeoff describes how perceptual abilities are degraded as speed is increased (Campos et al., 2013), rapid approaching speeds are required to capture prey or obtain food resources, otherwise food location can change or be consumed by another animal, rendering the information useless; however, moving fast prevents animals to make an exhaustive exploration of its environment (Fagan et al., 2017). On the other hand, the intensive-extensive tradeoff points how finding food-resources nearby impairs finding resources far-away, if the animal is performing and exhaustive search in the nearby area, information about areas far away is more difficult to obtain (Bartumeus et al., 2016 ; Raposo et al., 2011). Both tradeoffs imply that in order to obtain knowledge about food location and actually achieving success in obtaining it, 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 are required to deal with.

In order to deal with the inescapable uncertainty, animals establish two distinct behavioral modes: (a) local search and (b) relocation. Local search consists of small displacements and is predominantly informed by perceptual information, while relocation behavior implies traveling larger distances, using Lévy-like patterns (Bazazi et al., 2012). While there has been discussion of Brownian-like random movements guiding local search, this is most likely an emergent property caused by frequent food encounters (de Jager et al., 2014), so the shifts from a random relocation process to a perceptually guided local search, are in part result of an increase in food encounters. Thus, the overall food-seeking strategy derives from a combination of random movements, that allows to optimally sample the environment, and perceptually informed ones when food sources are near (Balogh et al., 2020).

Given the limitation in perceptual abilities, uncertainty seems to be inescapable. Even when an experienced animal can remember optimal foraging paths, random searches with distinct cycles of exploration/exploitation phases persist (Kembro et al., 2019). The reason why animals still explore, even when having knowledge of food location, might be that introducing stochasticity in food-seeking behavior improves success as it makes strategies more resilient to cognitive errors derived from perception (Campos et al., 2020). The persistence of strategies that permits the balance between exploitation, exploration (Campos et al., 2013), even when they are not technically needed (Ryali et al., 2016) shows how relevant uncertainty is food-seeking behavior. In this section, we presented how stochastic properties of the environment lead to behavioral adaptations to deal with the resultant uncertainty, by favoring strategies that aim to reduce uncertainty (Peters et al., 2017). In the following 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

A model that presents how animals include 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 within a set of possible actions for a particular state of the environment . Any action causes an stochastic transition from a state to another state . As such, the result of an interaction between animal and environment can be described by its value , which is a function of both action and current environment state . Such model of action, state, and value corresponds to a Markov decision process (Sutton & Barto, 2018) that 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 , where is the obtained reward (interaction outcome), and such probabilities are defined for every pair of and . A Markov decision process that includes the perceptual uncertainty, which we deemed inherent to food-seeking behavior, can be included by considering that states are paired with observations made by the animal to infer state , 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 , 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 in an uncertain environment, the value of the option becomes a distribution over possible values that is updated every time an action is performed. For the simple case where rewards are obtained or not (without any difference in reward magnitude), has a Bernoulli distribution and . Then, these probabilities can be modeled with the Beta distribution, which takes parameters and . With the Beta distributions produce a uniform distribution over 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 increases by 1, biasing the distribution towards 1. On the other hand, if no reward is obtained, the parameter increases by 1, biasing the distribution towards 0.

Finally, the mean is defined as

and its variance by

With these simple statistical properties of the Beta distribution, we can represent uncertainty over the expected rewards for any given and . If the exploration is defined by the posterior distribution, then it can be considered a Thompson sampling strategy (Thompson, 1933). To select an action, a posterior is built for every action and updated according to the previously stated rules. Then, for each posterior, a single reward estimate is sampled, resulting in an estimated value for each action. The action selected greedily, so where 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. In general terms, a solution for this is to consider the reward vector as a weighted average over past rewards, with a step-size parameter . 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 . To include this, a probability for every by state is necessary, where is the set of all particular observations . To model state beliefs, the goal is to obtain the function that maps observations to action , given an underlying model that relates states with observations. A hidden Markov model (HMM) represents this. HMM generates conditional probability distributions . Bayesian, among other methods for obtaining such model given only actions and observations has been proposed (Funamizu et al., 2012; Piray & Daw, 2020; Yoon et al., 2018), that allows us to explicitly model how an animal infers the current state given perceptual information.

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

Uncertainty arises from having more than one option, with the motivation to opt for one of those somewhat distributed. Considering that 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

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.

If we consider a simple environment with only one state and one action which initiates a food-seeking bout, and only two possible outcomes: food is found () or not found (), then . If an animal performs multiple food-seeking bouts and none of them are successful , the same is true if all are successful. However, if the probability of a successful food-seeking bout is 0.5, then entropy is maximized . Neural representation of entropy has been found in the middle cingulate cortex (MCC) specifically encoding outcome entropy (Gloy et al., 2020; Goñi et al., 2011) so this computation seems to be biologically plausible. However, entropy is not directly available as sensory input, and must be derived from actions and outcomes, which are dependent on the environment state. Previously, through Thompson sampling, we provided a way in which action outcome entropy could be encoded as variance in the posterior distribution. Nevertheless, a more direct way to compute entropy is possible through the prediction error.

The classical model of Rescorla-Wagner (Rescorla et al., 1972) modeled how animals could predict the reinforcing value of a given stimuli

where the value representation of the stimuli is obtained by considering the previously estimated value $y\_{t-1} weighted by a learning rate , and a prediction error . is the simple difference between the expected reward and the actual reward where is the obtained reward. An extension to this model has been proposed by (Sutton & Barto, 2018) where the prediction error considers an estimate of the rewards that gives more weight to current rewards, while still considering past rewards

where is a discount factor for all the history of rewards, and is an estimate for the true value of the reward. Finally, is the learning rate which effectively weights the reward prediction error so to make small updates or rather large ones 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, in such a model is a hyperparameter, thus is not derived from experience. Pearce & Hall (1980) model proposes that can be controlled by the prediction error magnitude so

Higher entropy on reward outcomes increases the probability of error (Feder & Merhav, 1994), thus increasing and, consequently, . 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).

A good candidate neurotransmitter for representing uncertainty is dopamine (DA) as it encodes the prediction error (C. D. Fiorillo, 2003; C. D. Fiorillo, 2011; Gershman & Uchida, 2019; Glimcher, 2011; Khaw et al., 2017; Lak et al., 2014; Nasser et al., 2017), more specifically, DA phasic median activity encodes reward probability as positive linear relationship for the conditioned stimulus, and as a negative linear relationship for the unconditioned stimulus (C. D. Fiorillo, 2003). Effectively encoding the prediction of the conditioned stimulus and the surprise for the unconditioned stimulus (if a reward has a low probability, obtaining it should be ‘surprising’). The change in DA sustained activation encodes rewards probability analogous to entropy, that is, displaying a peak of activity when the reward is obtained with a probability of 0.5, where maximum entropy is attained, and lower relative levels of activity at a probability of 1 and 0 (C. D. Fiorillo, 2003). Similar encoding of uncertainty has been observed in hippocampal activity (Schiffer et al., 2012), on the striatum (den Ouden et al., 2010), substantia nigra (Zaghloul et al., 2009), and VTA (Iordanova et al., 2021). Ventral tegmental area DA activity is well modeled by classical reinforcement models for cue-reward learning tasks, showing predicted changes in activity when previously learned cue-reward paired are modified (Chakroun et al., 2020; Steinberg et al., 2013). Moreover, DA activity is capable of regulating exploratory/exploitative behavior via tonic and phasic signaling, respectively (Beeler et al., 2010).

There is substantial evidence that DA neurons, specifically in the VTA, serve the functional role of computing reward prediction errors (Watabe-Uchida et al., 2017 ), by weighing inputs from multiple brain areas, most remarkably the lateral hypothalamus, dorsal and ventral striatum, ventral pallidum, and subthalamic nucleus (Tian et al., 2016). Acetylcholine (ACh) and norepinephrine (NA) activity has been proposed to signal expected and unexpected uncertainty, respectively (Yu & Dayan, 2005). Both neurotransmitters are mainly produced in the basal forebrain (Sturgill et al., 2020) and locus coeruleus (LC) (Aston-Jones & Cohen, 2005; Sales et al., 2019) and also play an important role in the reward prediction error signal. ACh antagonists increase the response sensitivity to expected uncertainty (Marshall et al., 2016), which supports the role of ACh as signaling expected uncertainty, or in other words, that changes in the outcomes are part of chance fluctuations and there are no structural changes present. On the other hand, LC tonic activity signals unexpected uncertainty promoting the learning of new contingencies by increasing exploratory behavior (Aston-Jones & Cohen, 2005; Payzan-LeNestour et al., 2013).

In this section, we presented simple models showing how uncertainty can be computed using the reward prediction error without any meaningful computational complexity. One of the main candidates for supporting the computation of the reward prediction error is DA activity within the VTA, paired with ACh and NA activity modulating the sensitivity of DA according to whether we are in presence of expected or unexpected uncertainty. In the following two sections, we will show empirical evidence on how food-access uncertainty increases food-seeking behavior, which in turn could contribute to overweight related issues 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

Food-seeking behavior adapted to environments with limited food. When food is limited animals approach cues with the highest associative strength to actual food (Montague et al., 1996). Reward prediction error then signals whether the current state is better or worse than expected, and animals should move or stay where the highest expected value is (Kacelnik & Bateson, 1996). However, uncertainty in food delivery increases lever pressing and reduces the time latency for approaching the lever (Anselme et al., 2013b), which in fact means the opposite of preferring cues with highest associative strength food. Intermittent access to high-fat diets generates binge-eating behavior corresponding to drastic increase in food-seeking behavior (Hess et al., 2019; King et al., 2016; Lardeux et al., 2013), increases reinforcement value upon withdrawal (McGee et al., 2010) and operant behavior without withdrawal (Wojnicki et al., 2013; Wojnicki et al., 2007, 2015) attenuating the effects of food devaluation, general psychomotor behavior increases (Hardaway et al., 2016), and DA-Ach release also increases, augmenting motivated behavior towards food (Rada et al., 2005). Mainly because of logistic reasons, intermittent access is provided at the same time of the day (in most cases), and that allows animals to predict the arrival of food, although with some inconsistencies within animals (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 food-seeking behavior increases in response to uncertainty in action-outcome association, and food availability.

Food-seeking behavior is increased to avoid starvation when a food shortage is predicted, and as previously noted, this derives in increased exploration, number foraging bouts, and time expended in foraging (Harris et al., 2010). Food insecurity is defined as the perception of how secure or certain food access is going to be in the future (Dhurandhar, 2016), and positively correlates with greater energetic intake (Dhurandhar, 2016) and preference for high-fat alternatives (Nettle et al., 2017), which are mainly cheap access foods in developed countries where this effect is more pronounced (Dinour et al., 2007; Moradi et al., 2019; Nettle & Bateson, 2019). The effect of food insecurity on increased intake are probably due to overeating fats and carbohydrates (cheap access and calorically dense foods) in periods of high food availability, leading to overweight (Stinson et al., 2018). 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, increasing food-seeking behavior is likely to result in excessive caloric intake.

In this section we showed how action-outcome uncertainty, intermittent food access, and food insecurity show converging evidence on how food access uncertainty invigorates food-seeking behavior. However, coupled with an easy access to highly energetic foods, an adaptive behavior for preventing starvation, turns maladaptive resulting in overweight. 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 (Tyree et al., 2018).

## Orexin as a potential mediator of uncertainty-drive foraging

Animals respond to the environment by seeking to preserve certain physiological parameters within a normal range, this homeostatic process is highly adaptive and can generate transient changes within animals physiology (Davies, 2016). 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; Toshinai et al., 2003). Together, this functionality takes the hypothalamus to a pivotal role in the homeostasis process, specifically related to food intake control.

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 (Chowdhury et al., 2019). Its functions range from regulating sleep/wakefulness states (Chemelli et al., 1999) to energetic balance (Yamanaka et al., 2003). Thus, lateral hypothalamus orexin neurons are in a well-suited spot to control foraging-related behaviors. Orexin activity promotes locomotor activity but is rapidly inhibited upon contact with food (González et al., 2016), activity increases upon sucrose predictive cues (Hassani et al., 2016), and in certain subpopulations, the increased spontaneous physical activity is directed towards food sources (Zink et al., 2018). Thus, orexin-related activity can be interpreted as food procuring signal (Zink et al., 2018). Further support for this interpretation comes from orexin increasing olfactory activity (Prud’homme et al., 2009), enhancing visual attention (Zajo et al., 2016), impairment of spatial working memory in orexin knockout mice (Dang et al., 2018), among others which have been classified as foraging-related behavior by Barson (2020).

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.

VTA has glutamatergic projections into the nucleus accumbens shell, specifically into parvalbumin GABAergic interneurons. Activation of GABAergic interneurons results in inhibition of medium spiny neurons activity into the lateral hypothalamus Qi et al. (2016)], which are known to be inhibitory (O’Connor et al., 2015’Connor\_Etal\_2015; Perez-Leighton et al., 2017). This inhibition results in a increase of lateral hypothalamus activity (Gutierrez et al., 2011; Stratford & Kelley, 1999). This connectivity makes posibles that DA activity response to environmental uncertainty, to increase lateral hypothalamus activity via dopaminergic inputs to the supramammillary nucleus (Plaisier et al., 2020). Additionally, a positive feedback loop between VTA and lateral hypothalamus can be established with the inhibition of the nucleus accumbens shell, as previously described. This could result in a net increase of food-seeking behavior via uncertainty-driven activity in the VTA.

Orexin signal depolarizes the LC (Hagan et al., 1999) promoting 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). Task disengagement and updating of action-outcome pairings point to a role of orexin in modulating uncertainty-related behavior, specifically regarding unexpected uncertainty. On the other hand, orexin is also related to ACh activity via a positive feedback loop with the basal forebrain (Sakurai et al., 2005). ACh activity somewhat mirrors DA prediction error (Sturgill et al., 2020), so DA and ACh connectivity with orexin could generate a similar uncertainty-driven activity. However, ACh activity might have an additional role by signaling a smaller variance in outcome results(Puigbò et al., 2020; Yu & Dayan, 2005) effectively sustaining an expected uncertainty signal.

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 increased uncertainty in food access increases motivation for palatable foods. We will model uncertainty by using an automated feeder to deliver food using fixed and variable schedules over 10 weeks. Next, we will use a progressive ratio test for sucrose as a measure of motivation for palatable food and whether uncertainty increases orexin expression and its ability to increase motivation for sucrose. Additionally, we will model an environment with both a certain and an uncertain alternative available at the same time over 2 weeks. For this we will use a lickometer with a spout randomly delivering a sucrose reward 50% of the time (uncertain spout) and a spout delivering it 100% of the time (certain reward). We will measure whether introducing an uncertain option increases the amount of licks and events, even when a certain option is present.
2. Determine whether orexin promotes uncertainty-driven motivation for a sucrose solution. We will model uncertainty by presenting two alternatives one with uncertainty and the other with certainty about reward delivery. However, here choosing one option implies not choosing the other. Uncertainty-driven motivation will be considered as a change in the selection ratio between certain and uncertain alternatives in saline and orexin injected mice. We will also test the ratio change while blocking orexin function.
3. Determine the differential gene expression in the hypothalamic area in uncertainty compared to certain conditions of specific objective 1. We will perform mRna sequencing on hypothalamic brain tissue of subjects in both conditions present in specific objective 1.
4. Determine if orexin/dynorphin neurons projecting to the paraventricular nucleus and VTA are active during sucrose intake. We will inject a retrograde tracer into PVN and VTA and identify fos expressing projections after an operant task of sucrose acquisition.

# 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 normal and 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. Orexin, during hedonic intake, presents functional connectivity to the VTA.
4. Orexin agonists increase the choice of uncertain options when presented with a certain and uncertain alternative, while orexin antagonists reduce the choice of uncertain options, relative to control.

# Methods

## Specific hypothesis 1, Experiment 1

### Animals

16 normal C57/BL6 mice. 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 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, within each cage an automatic feeding device will be placed, delivering nutritionally complete food-pellets. Animals will be in this condition for a week to acclimate to the feeding device. After the acclimation period, mice will be split in two groups (n = 8), balancing for food-pellet intake. The control group will maintain the same conditions as the acclimation phase, whereas the treatment group will receive food pellets with a random delay centered on the mean delay of the control group (15 seconds). 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 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

The progressive ratio task consists of a cage with two spouts, one with water and the other one with a sucrose solution, alternating positions between sessions. Mice will be trained in 5 sessions (60 min each) to receive either water or sucrose after 5 licks, throughout the session. In the progressive ratio task, the required amount of licks will increase logarithmically starting from 5 within the 1 hour duration of the task, after every reward acquisition the spout will be inactive for 20 seconds.

### Quantitative analysis of gene expression

Mice will be euthanized with isoflurane. Brain will be removed and stored at -80 C until hypothalamus dissection. 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 and mRNA quantification according to (Lazzarino et al., 2019).

### Materials and instruments

### Feeding experimentation device 2 (FED 2)

Is an automatic feeding device with an open-source design, 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 that can deliver a predetermined amount of liquid solution. Each spout detects contact with the animal tongue, allowing us to measure the number of licks per spout, and when they were performed. 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 laboratory previous pilot studies have shown that between 5 to 7 sessions are sufficient for animals to stabilize their behavior and learn the operant task. For acclimation purposes the lickometer is equipped with two spouts that deliver water and a sucrose solution upon 5 licks.

### Data analysis

Our main interest is to measure food motivation given prolonged exposure to uncertainty in food-access. Food motivation will be measured as the number of events for sucrose in the progressive ratio task- an event is when the animal performs the required number of licks when the spout is active.

### Data analysis pipeline

Estimate the difference between control and treatment group regarding the number of events. To do this we will fit a negative binomial generalized mixed model, which allows us to appropriately model count data (number of events) under the negative binomial distribution, and explicitly account for individual response to treatment while adjusting for meaningful variables such as weight and mean food intake. Characterize the food-seeking behavior in the treatment and control group using a lick microstructure analysis (Nguyen et al., 2017). 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 further divided into burst 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) (Johnson et al., 2010). Previously described features will be used as dependent variables in the mixed effects models to assess group differences. Compare food-pellet intake between groups. Using the previously described model we can compare the main and temporal course effect of uncertainty in food-pellet count within the home cage, allowing us to compare an estimation of daily caloric intake. Obtain a measure of transcriptional differences between groups in prepro-orexin and orexin receptors. Using qPCR we aim to obtain an average relative normalized expression per group, 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 1, Experiment 2

### Animals

8 normal C57/BL6 mice. 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 4 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, nutritionally complete food-pellets and water will be provided ad-lib. For two weeks animals will be acclimated to the lickometer device. 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

Compare the number of events and licks, relative to baseline, of both control and treatment groups. As we are expecting a general trend towards reduction of both events and licks due to sucrose devaluation, 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. Because the expected (average) value of a choice corresponds to the probability of reward delivery of a given spout and animals will be well trained in the task, we will assume that animals know the expected value of each spout. We will compare an epsilon-greedy choice rule to model the intuition that if uncertainty does not affect reward value, 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. On the other hand, to model an effect of uncertainty in reward value we will use the Thompson sampling model, which establishes a belief distribution of the value of each spout, and then samples that distribution so beliefs centered around larger values would be sampled often, however uncertain options will also be sampled as the belief of the value can have a large variance between a small and a very large value. Model parameters will be determined using maximum likelihood estimation among standar statistical techniques (Wilson & Collins, 2019), and model comparison 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 or within the home cage food.

## Specific hypothesis 3, Experiment 1

### Animals

18 normal C57/BL6 mice. Preliminary results from our laboratory indicated that 4 mice per group are necessary to observe Fos differences in food intake (Coehn’s d = 0.25, p < 0.05) with immunofluorescence techniques (Carolina Sandoval, datos no publicados). 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

18 normal C57/BL6 mice

### 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. After recovery and 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, each camera frame will have a timestamp to synchronize the video with events during the task offline.

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

Compare the ratios of uncertain to certain options between conditions without drug, orexin antagonist and orexin agonist. This will be performed 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 in order 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. Additionally, adding session time courses as a covariate, will allow us to observe differences in choice behavior within each session. Video recording 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, and coupled with instantaneous speed and traveled distance, the time spent in exploratory behavior versus exploitative behavior can be inferred. Repeated measures ANOVA with Tukey post-hoc pairwise comparison, will allow us to determine the difference between these measures between the drug conditions.

# 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>

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. (2013b). 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>

Anselme, P., Robinson, M. J. F., & Berridge, K. C. (2013a). 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>

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>

Balogh, S. L., Björklund, N., Huber, D. P. W., & Lindgren, B. S. (2020). Random and Directed Movement by Warren Root Collar Weevils (Coleoptera: Curculionidae), Relative to Size and Distance of Host Lodgepole Pine Trees. *Journal of Insect Science*, *20*(4), 9. <https://doi.org/10.1093/jisesa/ieaa063>

Barson, J. R. (2020). Orexin/hypocretin and dysregulated eating: Promotion of foraging behavior. *Brain Research*, *1731*, 145915. <https://doi.org/10.1016/j.brainres.2018.08.018>

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>

Bazazi, S., Bartumeus, F., Hale, J. J., & Couzin, I. D. (2012). Intermittent Motion in Desert Locusts: Behavioural Complexity in Simple Environments. *PLoS Computational Biology*, *8*(5), e1002498. <https://doi.org/10.1371/journal.pcbi.1002498>

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>

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>

Campos, D., Méndez, V., Palmer, J., Cristín, J., & Bartumeus, F. (2020). *Random searchers cope with cognitive errors and uncertainty better than path planners*. Ecology. <http://biorxiv.org/lookup/doi/10.1101/2020.08.06.239269>

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>

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>

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>

Davies, K. J. A. (2016). Adaptive homeostasis. *Molecular Aspects of Medicine*, *49*, 1–7. <https://doi.org/10.1016/j.mam.2016.04.007>

de Jager, M., Bartumeus, F., Kölzsch, A., Weissing, F. J., Hengeveld, G. M., Nolet, B. A., Herman, P. M. J., & van de Koppel, J. (2014). How superdiffusion gets arrested: ecological encounters explain shift from Lévy to Brownian movement. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1774), 20132605. <https://doi.org/10.1098/rspb.2013.2605>

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal Prediction Error Modulates Cortical Coupling. *Journal of Neuroscience*, *30*(9), 3210–3219. <https://doi.org/10.1523/JNEUROSCI.4458-09.2010>

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>

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>

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>

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., & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, *20*(11), 703–714. <https://doi.org/10.1038/s41583-019-0220-7>

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>

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement\_3), 15647–15654. <https://doi.org/10.1073/pnas.1014269108>

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>

Gutierrez, R., Lobo, M. K., Zhang, F., & Lecea, L. de. (2011). Neural integration of reward, arousal, and feeding: Recruitment of VTA, lateral hypothalamus, and ventral striatal neurons. *IUBMB Life*, *63*(10), 824–830. https://doi.org/<https://doi.org/10.1002/iub.539>

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, 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., & 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>

Iordanova, M. D., Yau, J. O.-Y., McDannald, M. A., & Corbit, L. H. (2021). Neural substrates of appetitive and aversive prediction error. *Neuroscience & Biobehavioral Reviews*, *123*, 337–351. <https://doi.org/10.1016/j.neubiorev.2020.10.029>

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>

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>

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>

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>

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>

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>

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., Andrews, C., & Bateson, M. (2017). Food insecurity as a driver of obesity in humans: The insurance hypothesis. *Behavioral and Brain Sciences*, *40*, e105. <https://doi.org/10.1017/S0140525X16000947>

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>

Nguyen, K. P., Ali, M. A., O’Neal, T. J., Szczot, I., Licholai, J. A., & Kravitz, A. V. (2017). Feeding Experimentation Device (FED): Construction and Validation of an Open-source Device for Measuring Food Intake in Rodents. *Journal of Visualized Experiments*, *120*, 55098. <https://doi.org/10.3791/55098>

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>

Payzan-LeNestour, E., Dunne, S., Bossaerts, P., & O’Doherty, John P. (2013). The Neural Representation of Unexpected Uncertainty during Value-Based Decision Making. *Neuron*, *79*(1), 191–201. <https://doi.org/10.1016/j.neuron.2013.04.037>

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. <https://doi.org/10.1037/0033-295X.87.6.532>

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

Peters, A., McEwen, B. S., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, *156*, 164–188. <https://doi.org/10.1016/j.pneurobio.2017.05.004>

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>

Puigbò, J.-Y., Arsiwalla, X. D., González-Ballester, M. A., & Verschure, P. F. M. J. (2020). Switching Operation Modes in the Neocortex via Cholinergic Neuromodulation. *Molecular Neurobiology*, *57*(1), 139–149. <https://doi.org/10.1007/s12035-019-01764-w>

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>

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>

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>

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>

Stratford, T. R., & Kelley, A. E. (1999). Evidence of a Functional Relationship between the Nucleus Accumbens Shell and Lateral Hypothalamus Subserving the Control of Feeding Behavior. *Journal of Neuroscience*, *19*(24), 11040–11048. <https://doi.org/10.1523/JNEUROSCI.19-24-11040.1999>

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.

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>

Toshinai, K., Date, Y., Murakami, N., Shimada, M., Mondal, M. S., Shimbara, T., Guan, J.-L., Wang, Q.-P., Funahashi, H., Sakurai, T., Shioda, S., Matsukura, S., Kangawa, K., & Nakazato, M. (2003). Ghrelin-Induced Food Intake Is Mediated via the Orexin Pathway. *Endocrinology*, *144*(4), 1506–1512. <https://doi.org/10.1210/en.2002-220788>

Tyree, S. M., Borniger, J. C., & de Lecea, L. (2018). Hypocretin as a Hub for Arousal and Motivation. *Frontiers in Neurology*, *9*, 413. <https://doi.org/10.3389/fneur.2018.00413>

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>

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>

Watabe-Uchida, M., Eshel, N., & Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience*, *40*(1), 373–394. <https://doi.org/10.1146/annurev-neuro-072116-031109>

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, *8*, e49547. <https://doi.org/10.7554/eLife.49547>

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>

Wosniack, M. E., Santos, M. C., Raposo, E. P., Viswanathan, G. M., & da Luz, M. G. E. (2017). The evolutionary origins of Lévy walk foraging. *PLOS Computational Biology*, *13*(10), e1005774. <https://doi.org/10.1371/journal.pcbi.1005774>

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>

Yamanaka, A., Beuckmann, C. T., Willie, J. T., Hara, J., Tsujino, N., Mieda, M., Tominaga, M., Yagami, K., Sugiyama, F., Goto, K., Yanagisawa, M., & Sakurai, T. (2003). Hypothalamic Orexin Neurons Regulate Arousal According to Energy Balance in Mice. *Neuron*, *38*(5), 701–713. <https://doi.org/10.1016/S0896-6273(03)00331-3>

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>

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>