Proyecto de tesis doctoral para optar al grado de Doctor en Neurociencia en la Pontificia Universidad Católica de Chile

Luis Nicolás Luarte Rodríguez - Claudio Peréz-Leighton

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. Because the environment is in constant change, food seeking strategies must be in constant adaptation (Bartumeus et al. 2016; Kölzsch et al. 2015). However, animals are never completly aware of these environmental changes, food location, quality and the probability of obtaining food are only partially known (G. H. Pyke 1984). Thus, the main challenge for the foraging animal is how to determine an optimal strategy to procure food while having only partial knowledge of food location.

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 and Kello 2021; A. Reynolds et al. 2018; Viswanathan et al. 1996; Kölzsch et al. 2015). 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 and 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 (A. M. Reynolds and Frye 2007; Nauta, Khaluf, and Simoens 2020). This suggests that food seeking behavior likely evolved to deal with food location uncertainty that could not be reduced by perceptual information.

Starting from the base that animals have only partial knowledge about food location, we asked what strategy could provide optimal results given this constraint. Lévy-walks fit within this constraint providing a way to seeking for food without relying on perceptual information. Moreover, we proposed that this strategy could have emerged to lead with environment uncertainty. In the next section, we will explore how considering uncertainty within the food-seeking strategy can be useful when perceptual abilities are limited, and how uncertainty itself can modify food-seeking

behavior (Anselme and Robinson 2019; Anselme, Otto, and Güntürkün 2017; Anselme and Güntürkün 2019).

Foraging and uncertainty

As discussed previously, animals adopted food-seeking strategies to procure food within uncertain environments (Humphries and Sims 2014; Sims et al. 2019). Here, we will show how environmental uncertainty can regulate food-seeking behavior, its easily computed, and how brain structures readily can represent it.

To show how environmental uncertainty regulates food-seeking behavior, we will expose two case, one without considering uncertainty within food-seeking behavior, and the alternative case when it is considered. For the first case, we can assume that the animal will randomly sample its alternatives and then pick the one that, at that time, delivered the greatest value. This sampling and action selection strategy is called greedy (Tokic and Palm 2011). Acting greedily implies that the animal will always choose the option that yielded the most value in an initial random sampling, neglecting 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 and Barto 2018). An improvement over acting greedily is to store in memory the mean rewards, this can provide a way to compare if the selected alternative is good or bad compared to a global reference of what is the overall quality of the environment. The classical model by Charnov (1976), proposes that animals should decide whether to stay or leave a certain location based on the mean value of the whole environment. However, such model does not provide good fit to behavioral data, showing a systematic bias to stay more time within a certain location that the model predicts (Nonacs 2001; Le Heron et al. 2020; G. H. Pyke 2010; Hayden, Pearson, and Platt 2011).

For the second case, one way of incorporating uncertainty is through considering the spread (i.e. standard deviation) over possible outcomes (Rothwell and Stock 1988). This computation of uncertainty implies an average outcome from which the spread is observed, the anterior cingulate cortex (ACC) might support this, as is able to represent the magnitude of the spread in decision making tasks (Christopoulos et al. 2009). The representation of spread within the ACC is quite precise, as is able to integrate cue information for its representations, that is, a measure of uncertainty, as spread over the average value, is considered for every cue-related option or context (van Holstein and Floresco 2020). On the other hand, uncertainty can appear when the outcomes generated by certain action changes. The ACC activity increases when action-outcome contingencies change drastically (Behrens et al. 2007). Thus, ACC activity can signal whether an outcome is within the expected values or the value obtained is totally different from expected. Being able to distinguish between this two cases is extremely important to increase success while searching for food, an animal should not leave a good spot if it ocasionally bring bad results, and should leave a good spot if it starts to have consistenly bad results. The orbitofrontal cortex (OFC) has a role in this regulation. Stolyarova and Izquierdo (2017) showed that rats could choose the option with larger mean value despite large variability. However, a lesion to the orbitofrontal cortex (OFC) impaired this ability, making rat unable to change their option when the mean value was decreased. Together, these data show that different measures of uncertainty, are tracked by brain structures, allowing to choose the best option filtering out the noise.

The notion, previously presented, that an animals should behave differently when the outcome is within expected values or has changed drastically, is formalized in the models of expected and unexpected uncertainty (Yu and Dayan 2005). The model proposes expected uncertainty as the

uncertainty regarding outcomes when contingencies (outcome given a particular action) remain stable but is subject to some noise. On the other hand, unexpected uncertainty represents a drastic change in the contingencies likely due to 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 (Yu and Dayan 2005; Soltani and Izquierdo 2019).

As mentioned previously, animals should whether changes in the outcome are expected or unexpected. If the action-outcome contingencies never present abrupt changes the environment is described as stationary, whereas if from time to time contingencies change or the mean value goes up or down, then the environment is defined as non-stationary (Raj and Kalyani 2017). Animals, tend to behave as if the environment was non-stationary (Wu, Iyer, and Wang 2018), and they couple this with an increase in exploratory behavior (Ryali, Reddy, and Yu 2016), as predicted by the expected/unexpected model (Yu and Dayan 2005; Cohen, McClure, and Yu 2007). While the functional reasons behind this behavior are not clear, it has been proposed that acting as if the environment was non-stationary simply does not prove to be problematic in stationary ones (Ryali, Reddy, and Yu 2016), and can be used as an heuristic in natural environments (Reverdy, Srivastava, and Leonard 2014). That is, assuming a non-stationary environment and consequently increasing exploratory behavior, might prove to be useful in most cases and does not prove to be detrimental when the environment is stationary.

While exploring the environment can be useful, exploring too much can be detrimental to actually obtaining food, so exploiting current knowledge about food location is also necesary. This makes animals face the exploration-exploitation dillema (Sutton and Barto 2018). Exploiting means that behavior should be consistent with previously learned reward contingencies, choosing the option with the highest expected value. On the other hand, exploration implies re-sampling the environment to improve or re-learn current contingencies. Uncertainty is a key variable to the resolution of this dillema. Unexpected uncertainty increases the exploratory behavior, boosting learning of new contingencies, whereas expected uncertainty bias towards explotation (Cohen, McClure, and Yu 2007; Aston-Jones and Cohen 2005).

Here we discussed how sensing uncertainty in the food environment regulates food-seeking behavior. Animals act as if the food environments are uncertain and non-stationary even if this is not the case, supporting 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. We argue against the idea of perceptual abilities being dominant over the simple search strategies presented (LaScala-Gruenewald et al. 2019), because sensory perception is always limited in some aspect inroducing uncertainty (Bartumeus et al. 2016).

ALl sensory organs have a receptive field, which defines a range where information can orient food-seeking behavior (Fletcher et al. 2013). If an animal must know what is beyond such range, displacement is needed. However, to inform such displacement, memory, perceptual information integration or other cognitive process should come into play (Ranc et al. 2021), because by definition such palce lies beyond the receptive field. However, what is remembered about a certain location loses validity over time. Given that the environment is in constant change, integrating information of how food availability changes over can prove to be useful for the foraging animal

(Fagan et al. 2017).

Integrating information of location and availability over time can inform food-seeking behavior. However, obtaining good quality information requires the animal to explore its environment, 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, Bartumeus, and Méndez 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 (Raposo et al. 2011; Bartumeus et al. 2016). 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.

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, Bartumeus, and Méndez 2013), even when they are not technically needed (Ryali, Reddy, and Yu 2016) shows how relevant uncertainty is food-seeking behavior. In this section we presented the relevance of uncertainty if food-seeking behavior, and how inescapable it is, because even when sensory perception can help reduce it at a local scale, in re-introduces uncertainty on further away locations. 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 a within a set of possible actions $a \in A$ for a particular state of the environment s. Any action a causes an stochastic transition from a state s to another state s'. As such, the result of an interaction between animal and environment can be described by its value q, which is a function of both action and current environment state $q(s, a)$ (Fig. 1). Such model of action, state, and value corresponds to a Markov decision process (Sutton and 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 $p(s', r|s, a)$, where r is the obtained reward (interaction outcome), and such probabilities are defined for every pair of a and s. A Markov decision process that includes the perceptual uncertainty, which we deemed inherent to food-seeking behavior, can be included by considering that states s are paired with observations o made by the animal to infer state s, because the state cannot be directly observed or there is some sensory noise. Including perceptual uncertainty makes the animal to consider environment states as the conditional probability of any particular observation given a state $p(o|s)$, generating a belief of the current state based on perceptual information (Ma and Jazayeri 2014).

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

Finally, the mean is defined as

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

and its variance by

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

With these simple statistical properties of the Beta distribution, we can represent uncertainty over the expected rewards for any given a and s. The previously presented way to model environment uncertainty and q is the general case of Thompson sampling (Thompson 1933). To select an action, a posterior is built for every action and updated according to the previously stated rules (Fig. 2). Then, for each posterior, a single reward estimate $\hat{r}$ is sampled, resulting in an estimated value for each action. The action selected greedily, so $a = argmax_{a \in A}\hat{r}(a)$ where A is the set of possible actions within an environment (Wang and 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 $((0, 1])$. The lower the value of this parameter the more weight is given to recent rewards, on the other hand, if it's closer to 1, then all the reward history is equally considered. More complex alternatives to this problem include modeling non-stationarity as Poisson arrival process that modifies the means rewards

(Ghatak 2020), bayesian approaches to modulate past observed rewards (Raj and Kalyani 2017), and explicitly modeling environment volatility in a bayesian setup (Behrens et al. 2007).

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

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

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

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 s and one action a which initiates a food-seeking bout, and only two possible outcomes: food is found (p) or not found ($q = 1 - p$), then $H = -(plog_2 p + qlog2q)$. If an animal performs multiple food-seeking bouts and none of them are successful $H = 0$, 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 $H = 1$. Neural representation of entropy has been found in the middle cingulate cortex (MCC) specifically encoding outcome entropy (Goñi et al. 2011; Gloy, Herrmann, and Fehr 2020) 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

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

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

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

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

The Rescorla-Wagner model allows deriving a prediction error based on experience, where the learning rate can be set lower to simulate unexpected uncertainty or higher to simulate expected uncertainty. However, α in such a model is a hyperparameter, thus is not derived from experience. Pearce and Hall (1980) model proposes that α can be controlled by the prediction error magnitude $|\delta$ so

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

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

A good candidate neurotransmitter for representing uncertainty is dopamine (DA) as it encodes the prediction error (Nasser et al. 2017; C. D. Fiorillo 2003; C. D. Fiorillo 2011; Lak, Stauffer, and Schultz 2014; Glimcher 2011; Khaw, Glimcher, and Louie 2017; Gershman and Uchida 2019), 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 (Steinberg et al. 2013; Chakroun et al. 2020). 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, Eshel, and Uchida 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 and Dayan 2005). Both neurotransmitters are mainly produced in the basal forebrain (Sturgill et al. 2020) and locus coeruleus (LC) (Sales et al. 2019; Aston-Jones and Cohen 2005) 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 (Payzan-LeNestour et al. 2013; Aston-Jones and Cohen 2005).

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, Dayan, and Sejnowski 1996). If a lever was associated with food delivery, but now food is being delivered randomly 50% of the time upon lever pressing, then the reward prediction would increase until the expectation is re-adjusted, and the animal should lower the its rate of lever pressing compared to when the lever always delivered food (Kacelnik and Bateson 1996). However, uncertainty in food delivery increases lever pressing and reduces the time latency for approaching the lever (Anselme, Robinson, and Berridge 2013), 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 increases in food-seeking behavior (Hess et al. 2019; King et al. 2016; Lardeux, Kim, and Nicola 2013), increases reinforcement value upon withdrawal (McGee et al. 2010) and operant behavior without withdrawal (Wojnicki, Babbs, and Corwin 2013; Wojnicki, Stine, and Corwin 2007; Wojnicki et al. 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, Avena, and Hoebel 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, Guerrero-Vargas, and Escobar 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, Chapman, and Monfort 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, Andrews, and Bateson 2017), which are mainly cheap access foods in developed countries where this effect is more pronounced (Moradi et al. 2019; Dinour, Bergen, and Yeh 2007; Nettle and 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, Borniger, and de Lecea 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, Carter, and De Lecea 2010; Kosse and Burdakov 2014), motivation for food intake (Castro and Berridge 2017), receiving internal status information of fat deposits via leptin signaling (Pandit, Beerens, and Adan 2017; Meister 2000), 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, Fadel, and Burk 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 (Stratford and Kelley 1999; Gutierrez et al. 2011). This connectivity makes posibles that DA activity response to environmental uncertainty, to increase lateral hypothalamus activity via dopaminergic inputs to the supramammillary nucleus (Plaisier, Hume, and Menzies 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 and 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 and 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 substainance food access increases motivation for palatable foods.

2. Determine whether increased unceratinty in an obesogenic environment increases motivation for palatable foods.

3. 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.

4. Determine the differential gene expression in the hypothalamic area in uncertainty compared to certain conditions of specific objective 1. We will perform real time quantitative PCR (RT qPCR) on hypothalamic brain tissue of subjects in both conditions present in specific objective 1.

5. 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 substainance food supply and in obesogenic environments, increases food-seeking behavior.

2. Prolonged exposure to food-access uncertainty environments increases prepro-orexin and orexin receptor expression in the hypothalamus.

3. 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 $\alpha = 0.05$ the sample size per group is estimated at n = 4. For differential gene expression sample size was selected according to literature reporting sample sizes of 4-6 for orexin expression (Pankevich et al. 2010; Alcaraz-Iborra et al. 2014)

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 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 and 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 and 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 (α = 0.05 y β = 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.
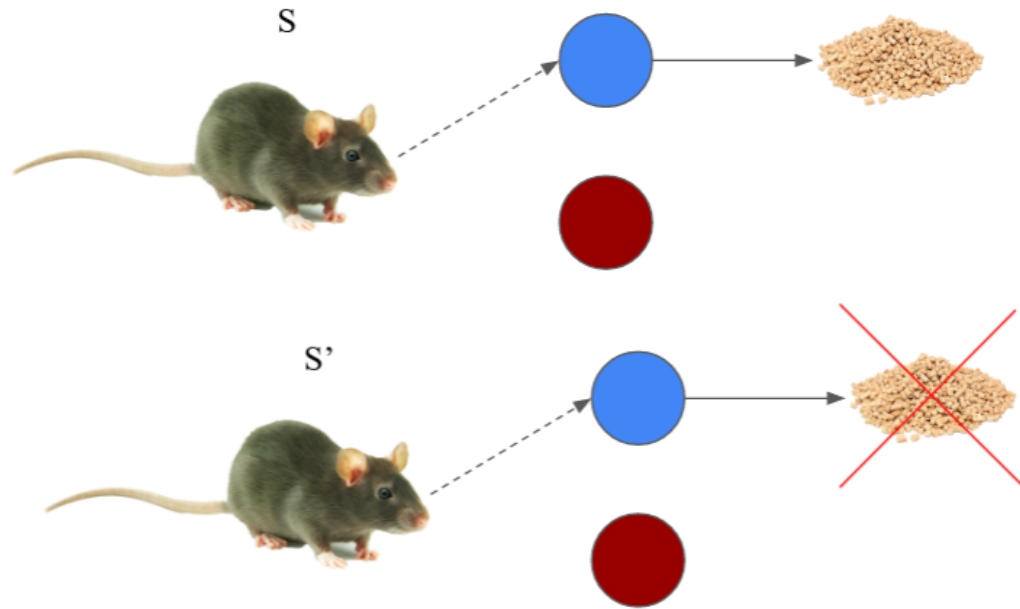
Figures

Figure 1: An animal in state s upon choosing the blue option receives a food reward. In turn, choosing the blue option changes the state to s' where choosing the same action does not deliver food. Here the value q is dependent on both action a (red or blue) and state (s or s')
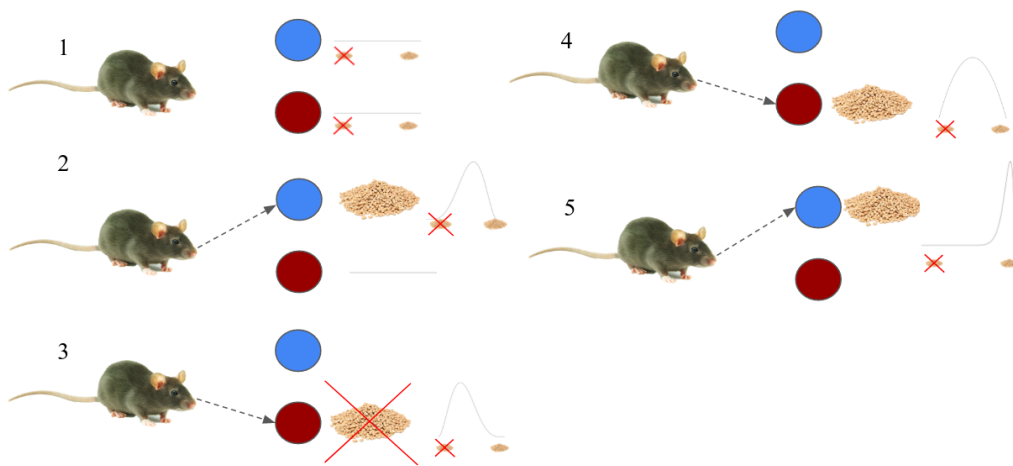
Figure 2: Creation of posterior distributions for possible actions. (1) the animal has never choosen any option so it thinks that the lever has an equal probability of given food or not. (2-3) the animal then sample both options, the blue one delivers food so the belief distribution is slighty bias toward that option delivering food, whereas the red option shows the opposite. (4) the animal re-samples the red option, but now food is delivered increasing the spread of the belief distribution, representing uncertainty on whether the red option is expected to give food or not. (5) The blue option is sampled again and food is delivered, thus increasing the bias towards the blue option delivering food, leaving less uncertainty over this option.

References

Adamantidis, Antoine, Matt C. Carter, and Luis De Lecea. 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, Manuel, Francisca Carvajal, José Manuel Lerma-Cabrera, Luis Miguel Valor, and Inmaculada Cubero. 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 (October): 93–99. https://doi.org/10.1016/j.bbr.2014.06.049.

Anselme, Patrick, and Onur Güntürkün. 2019. "How Foraging Works: Uncertainty Magnifies Food-Seeking Motivation." *Behavioral and Brain Sciences* 42: e35. https://doi.org/10.1017/S0140525X18000948.

Anselme, Patrick, Tobias Otto, and Onur Güntürkün. 2017. "How Unpredictable Access to Food Increases the Body Fat of Small Passerines: A Mechanistic Approach." *Behavioural Processes* 144 (November): 33–45. https://doi.org/10.1016/j.beproc.2017.08.013.

Anselme, Patrick, and Mike J. F. Robinson. 2019. "Evidence for motivational enhancement of sign-tracking behavior under reward uncertainty." *Journal of Experimental Psychology. Animal Learning and Cognition* 45 (3): 350–55. https://doi.org/10.1037/xan0000213.

Anselme, Patrick, Mike J.F. Robinson, and Kent C. Berridge. 2013. "Reward Uncertainty Enhances Incentive Salience Attribution as Sign-Tracking." *Behavioural Brain Research* 238 (February): 53–61. https://doi.org/10.1016/j.bbr.2012.10.006.

Aston-Jones, Gary, and Jonathan D. Cohen. 2005. "AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance." *Annual Review of Neuroscience* 28 (1): 403–50. https://doi.org/10.1146/annurev.neuro.28.061604.135709.

Barson, Jessica R. 2020. "Orexin/Hypocretin and Dysregulated Eating: Promotion of Foraging Behavior." *Brain Research* 1731 (March): 145915. https://doi.org/10.1016/j.brainres.2018.08.018.

Bartumeus, Frederic, Daniel Campos, William S. Ryu, Roger Lloret-Cabot, Vicenç Méndez, and Jordi Catalan. 2016. "Foraging Success Under Uncertainty: Search Tradeoffs and Optimal Space Use." Edited by Peter Thrall. *Ecology Letters* 19 (11): 1299–1313. https://doi.org/10.1111/ele.12660.

Beeler, Jeff A., Nathaniel D. Daw, Cristianne R. M. Frazier, and Xiaoxi Zhuang. 2010. "Tonic Dopamine Modulates Exploitation of Reward Learning." *Frontiers in Behavioral Neuroscience* 4. https://doi.org/10.3389/fnbeh.2010.00170.

Behrens, Timothy E J, Mark W Woolrich, Mark E Walton, and Matthew F S Rushworth. 2007. "Learning the Value of Information in an Uncertain World." *Nature Neuroscience* 10 (9): 1214–21. https://doi.org/10.1038/nn1954.

Burdakov, Denis. 2020. "How Orexin Signals Bias Action: Hypothalamic and Accumbal Circuits." *Brain Research* 1731 (March): 145943. https://doi.org/10.1016/j.brainres.2018.09.011.

Campos, Daniel, Frederic Bartumeus, and Vicenç Méndez. 2013. "Search Times with Arbitrary Detection Constraints." *Physical Review E* 88 (2): 022101. https://doi.org/10.1103/PhysRevE.88.022101.

Campos, Daniel, Vicenç Méndez, John Palmer, Javier Cristín, and Frederic Bartumeus. 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, Daniel C., and Kent C. Berridge. 2017. "Opioid and Orexin Hedonic Hotspots in Rat Orbitofrontal Cortex and Insula." *Proceedings of the National Academy of Sciences* 114 (43): E9125–34. https://doi.org/10.1073/pnas.1705753114.

Chakroun, Karima, David Mathar, Antonius Wiehler, Florian Ganzer, and Jan Peters. 2020. "Dopaminergic Modulation of the Exploration/Exploitation Trade-Off in Human Decision-Making." *eLife* 9 (June): e51260. https://doi.org/10.7554/eLife.51260.

Charnov, Eric L. 1976. "Optimal Foraging, the Marginal Value Theorem." *Theoretical Population Biology* 9 (2): 129–36. https://doi.org/10.1016/0040-5809(76)90040-X.

Chemelli, Richard M., Jon T. Willie, Christopher M. Sinton, Joel K. Elmquist, Thomas Scammell, Charlotte Lee, James A. Richardson, et al. 1999. "Narcolepsy in Orexin Knockout Mice: Molecular Genetics of Sleep Regulation." *Cell* 98 (4): 437–51. https://doi.org/10.1016/S009 2-8674(00)81973-X.

Chowdhury, Srikanta, Chi Jung Hung, Shuntaro Izawa, Ayumu Inutsuka, Meiko Kawamura, Takashi Kawashima, Haruhiko Bito, et al. 2019. "Dissociating Orexin-Dependent and -Independent Functions of Orexin Neurons Using Novel Orexin-Flp Knock-in Mice." *eLife* 8 (June): e44927. https://doi.org/10.7554/eLife.44927.

Christopoulos, G. I., P. N. Tobler, P. Bossaerts, R. J. Dolan, and W. Schultz. 2009. "Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making Under Risk." *Journal of Neuroscience* 29 (40): 12574–83. https://doi.org/10.1523/JNEUROSCI.2614-09.2009.

Cohen, Jonathan D, Samuel M McClure, and Angela J Yu. 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): 933–42. https://doi.org/10.1098/rstb.2007.2098.

Dang, Ruozhi, Qiuhan Chen, Jie Song, Chao He, Jun Zhang, Jianxia Xia, and Zhian Hu. 2018. "Orexin Knockout Mice Exhibit Impaired Spatial Working Memory." *Neuroscience Letters* 668 (March): 92–97. https://doi.org/10.1016/j.neulet.2018.01.013.

Davies, Kelvin J. A. 2016. "Adaptive Homeostasis." *Molecular Aspects of Medicine*, Hormetic and regulatory effects of lipid oxidation products, 49 (June): 1–7. https://doi.org/10.1016/j.mam.2016.04.007.

Dhurandhar, Emily J. 2016. "The Food-Insecurity Obesity Paradox: A Resource Scarcity Hypothesis." *Physiology & Behavior* 162 (August): 88–92. https://doi.org/10.1016/j.physbeh.2016.04.025.

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

Dinour, Lauren M., Dara Bergen, and Ming-Chin Yeh. 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): 1952–61. https://doi.org/10.1016/j.jada.2007.08.0 06.

Fagan, William F., Eliezer Gurarie, Sharon Bewick, Allison Howard, Robert Stephen Cantrell, and Chris Cosner. 2017. "Perceptual Ranges, Information Gathering, and Foraging Success in Dynamic Landscapes." *The American Naturalist* 189 (5): 474–89. https://doi.org/10.108 6/691099.

Feder, M., and N. Merhav. 1994. "Relations Between Entropy and Error Probability." *IEEE Transactions on Information Theory* 40 (1): 259–66. 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 (December): 162–71. https://doi.org/10.1016/j.neuroscience.2011.09.037.

Fletcher, Robert J., Christopher W. Maxwell, John E. Andrews, and Wendy L. Helmey-Hartman. 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/s109 80-012-9812-6.

Funamizu, Akihiro, Makoto Ito, Kenji Doya, Ryohei Kanzaki, and Hirokazu Takahashi. 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): 1180–89. https://doi.org/10.1111/j.1460-9568.2012.08025.x.

Garg, Ketika, and Christopher T Kello. 2021. "Efficient Lévy Walks in Virtual Human Foraging." *Scientific Reports* 11 (1): 5242. https://doi.org/10.1038/s41598-021-84542-w.

Gentleman, Robert C, Vincent J Carey, Douglas M Bates, Ben Bolstad, Marcel Dettling, Sandrine Dudoit, Byron Ellis, et al. 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, Samuel J., and Naoshige Uchida. 2019. "Believing in Dopamine." *Nature Reviews Neuroscience* 20 (11): 703–14. https://doi.org/10.1038/s41583-019-0220-7.

Ghatak, Gourab. 2020. "A Change-Detection Based Thompson Sampling Framework for Non-Stationary Bandits." *arXiv:2009.02791 [Cs, Eess]*, September. http://arxiv.org/abs/2009.027 91.

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): 15647–54. https://doi.org/10.1073/pnas.1014269108.

Gloy, K., M. Herrmann, and T. Fehr. 2020. "Decision Making Under Uncertainty in a Quasi Realistic Binary Decision Task – An fMRI Study." *Brain and Cognition* 140 (April): 105549. https://doi.org/10.1016/j.bandc.2020.105549.

Goñi, Joaquín, Maite Aznárez-Sanado, Gonzalo Arrondo, María Fernández-Seara, Francis R. Loayza, Franz H. Heukamp, and María A. Pastor. 2011. "The Neural Substrate and Functional Integration of Uncertainty in Decision Making: An Information Theory Approach." Edited by Matjaz Perc. *PLoS ONE* 6 (3): e17408. https://doi.org/10.1371/journal.pone.0017408.

González, J. Antonio, Lise T. Jensen, Panagiota Iordanidou, Molly Strom, Lars Fugger, and Denis Burdakov. 2016. "Inhibitory Interplay Between Orexin Neurons and Eating." *Current Biology* 26 (18): 2486–91. https://doi.org/10.1016/j.cub.2016.07.013.

Grella, Stephanie L., Jonathan M. Neil, Hilary T. Edison, Vanessa D. Strong, Irina V. Odintsova, Susan G. Walling, Gerard M. Martin, Diano F. Marrone, and Carolyn W. Harley. 2019. "Locus Coeruleus Phasic, But Not Tonic, Activation Initiates Global Remapping in a Familiar Environment." *Journal of Neuroscience* 39 (3): 445–55. https://doi.org/10.1523/JNEURO SCI.1956-18.2018.

Gutierrez, Ranier, Mary Kay Lobo, Feng Zhang, and Luis de Lecea. 2011. "Neural Integration of Reward, Arousal, and Feeding: Recruitment of VTA, Lateral Hypothalamus, and Ventral Striatal Neurons." *IUBMB Life* 63 (10): 824–30. https://doi.org/https://doi.org/10.1002/iub. 539.

Hagan, Jim J., Ron A. Leslie, Sara Patel, Martyn L. Evans, Trevor A. Wattam, Steve Holmes, Christopher D. Benham, et al. 1999. "Orexin A Activates Locus Coeruleus Cell Firing and Increases Arousal in the Rat." *Proceedings of the National Academy of Sciences* 96 (19): 10911–16. https://doi.org/10.1073/pnas.96.19.10911.

Hardaway, J. Andrew, Jennifer Jensen, Michelle Kim, Christopher M. Mazzone, Jonathan A. Sugam, Jeffrey F. Diberto, Emily G. Lowery-Gionta, et al. 2016. "Nociceptin Receptor Antagonist SB 612111 Decreases High Fat Diet Binge Eating." *Behavioural Brain Research* 307 (July): 25–34. https://doi.org/10.1016/j.bbr.2016.03.046.

Harris, Tara R., Colin A. Chapman, and Steven L. Monfort. 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, Oum Kaltoum, Matthew R. Krause, Lynda Mainville, Christopher A. Cordova, and Barbara E. Jones. 2016. "Orexin Neurons Respond Differentially to Auditory Cues Associated with Appetitive Versus Aversive Outcomes." *The Journal of Neuroscience* 36 (5): 1747–57. https://doi.org/10.1523/JNEUROSCI.3903-15.2016.

Hayden, Benjamin Y., John M. Pearson, and Michael L. Platt. 2011. "Neuronal basis of sequential foraging decisions in a patchy environment." *Nature Neuroscience* 14 (7): 933–39. https://doi.org/10.1038/nn.2856.

Hess, Andreas, Stefanie Kress, Stefan Rakete, Gerald Muench, Johannes Kornhuber, Monika Pischetsrieder, and Christian P. Müller. 2019. "Influence of the Fat/Carbohydrate Component of Snack Food on Energy Intake Pattern and Reinforcing Properties in Rodents." *Behavioural Brain Research* 364 (May): 328–33. https://doi.org/10.1016/j.bbr.2019.02.041.

Holstein, Mieke van, and Stan B. Floresco. 2020. "Dissociable Roles for the Ventral and Dorsal Medial Prefrontal Cortex in Cue-Guided Risk/Reward Decision Making." *Neuropsychopharmacology* 45 (4): 683–93. https://doi.org/10.1038/s41386-019-0557-7.

Humphries, Nicolas E., and David W. Sims. 2014. "Optimal Foraging Strategies: Lévy Walks Balance Searching and Patch Exploitation Under a Very Broad Range of Conditions." *Journal of Theoretical Biology* 358 (October): 179–93. https://doi.org/10.1016/j.jtbi.2014.05.032.

Iordanova, Mihaela D., Joanna Oi-Yue Yau, Michael A. McDannald, and Laura H. Corbit. 2021. "Neural Substrates of Appetitive and Aversive Prediction Error." *Neuroscience & Biobehavioral Reviews* 123 (April): 337–51. https://doi.org/10.1016/j.neubiorev.2020.10.029.

Johnson, A.W., A. Sherwood, D.R. Smith, M. Wosiski-Kuhn, M. Gallagher, and P.C. Holland. 2010. "An Analysis of Licking Microstructure in Three Strains of Mice." *Appetite* 54 (2): 320–30. https://doi.org/10.1016/j.appet.2009.12.007.

Kacelnik, Alex, and Melissa Bateson. 1996. "Risky Theories—The Effects of Variance on Foraging Decisions." *American Zoologist* 36 (4): 402–34. https://doi.org/10.1093/icb/36.4.402.

Kane, Gary A., Elena M. Vazey, Robert C. Wilson, Amitai Shenhav, Nathaniel D. Daw, Gary Aston-Jones, and Jonathan D. Cohen. 2017. "Increased Locus Coeruleus Tonic Activity Causes Disengagement from a Patch-Foraging Task." *Cognitive, Affective, & Behavioral Neuroscience* 17 (6): 1073–83. https://doi.org/10.3758/s13415-017-0531-y.

Kembro, Jackelyn M., Mathieu Lihoreau, Joan Garriga, Ernesto P. Raposo, and Frederic Bartumeus. 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, Mel W., Paul W. Glimcher, and Kenway Louie. 2017. "Normalized Value Coding Explains Dynamic Adaptation in the Human Valuation Process." *Proceedings of the National Academy of Sciences* 114 (48): 12696–701. https://doi.org/10.1073/pnas.1715293114.

King, S. J., T. Rodrigues, A. Watts, E. Murray, A. Wilson, and A. Abizaid. 2016. "Investigation of a role for ghrelin signaling in binge-like feeding in mice under limited access to high-fat diet." *Neuroscience* 319 (April): 233–45. https://doi.org/10.1016/j.neuroscience.2016.01.004.

Kölzsch, Andrea, Adriana Alzate, Frederic Bartumeus, Monique de Jager, Ellen J. Weerman,

Geerten M. Hengeveld, Marc Naguib, Bart A. Nolet, and Johan van de Koppel. 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, Christin, and Denis Burdakov. 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, Donald 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, Armin, William R. Stauffer, and Wolfram Schultz. 2014. "Dopamine Prediction Error Responses Integrate Subjective Value from Different Reward Dimensions." *Proceedings of the National Academy of Sciences* 111 (6): 2343–48. https://doi.org/10.1073/pnas.1321596111.

Lardeux, Sylvie, James J. Kim, and Saleem M. Nicola. 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 (April): 21–31. https://doi.org/10.1016/j.physbeh.2013.03.005.

LaScala-Gruenewald, Diana E., Rohan S. Mehta, Yu Liu, and Mark W. Denny. 2019. "Sensory Perception Plays a Larger Role in Foraging Efficiency Than Heavy-Tailed Movement Strategies." *Ecological Modelling* 404 (July): 69–82. https://doi.org/10.1016/j.ecolmodel.2019.02.015.

Lazzarino, Gisela Paola, María Florencia Acutain, Guillermina Canesini, María Florencia Andreoli, and Jorge Guillermo Ramos. 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 (December): 110542. https://doi.org/10.1016/j.mce.2019.110542.

Le Heron, Campbell, Nils Kolling, Olivia Plant, Annika Kienast, Rebecca Janska, Yuen-Siang Ang, Sean Fallon, Masud Husain, and Matthew A. J. Apps. 2020. "Dopamine Modulates Dynamic Decision-Making during Foraging." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 40 (27): 5273–82. https://doi.org/10.1523/JNEUROSCI.2586-19.2020.

Livak, Kenneth J., and Thomas D. Schmittgen. 2001. "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method." *Methods* 25 (4): 402–8. https://doi.org/10.1006/meth.2001.1262.

Luby, Matthew D., Cynthia T. Hsu, Scott A. Shuster, Christian M. Gallardo, Ralph E. Mistlberger, Oliver D. King, and Andrew D. Steele. 2012. "Food Anticipatory Activity Behavior of Mice Across a Wide Range of Circadian and Non-Circadian Intervals." Edited by Paul A. Bartell. *PLoS ONE* 7 (5): e37992. https://doi.org/10.1371/journal.pone.0037992.

Ma, Wei Ji, and Mehrdad Jazayeri. 2014. "Neural Coding of Uncertainty and Probability." *Annual Review of Neuroscience* 37 (1): 205–20. https://doi.org/10.1146/annurev-neuro-071013-014017.

Marshall, Louise, Christoph Mathys, Diane Ruge, Archy O. de Berker, Peter Dayan, Klaas E. Stephan, and Sven Bestmann. 2016. "Pharmacological Fingerprints of Contextual Uncertainty." Edited by Matthew F. S. Rushworth. *PLOS Biology* 14 (11): e1002575. https://doi.org/10.1371/journal.pbio.1002575.

McGee, Heidi M., Bekaluwa Amare, Amy L. Bennett, and Elizabeth A. Duncan-Vaidya. 2010. "Behavioral Effects of Withdrawal from Sweetened Vegetable Shortening in Rats." *Brain

Research, Neural Mechanisms of Ingestive Behaviour and Obesity, 1350 (September): 103–11. https://doi.org/10.1016/j.brainres.2010.01.033.

Meister, Björn. 2000. "Control of Food Intake via Leptin Receptors in the Hypothalamus." In *Vitamins & Hormones*, 59:265–304. Academic Press. https://www.sciencedirect.com/science/article/pii/S0083672900590104.

Montague, P. R., P. Dayan, and T. J. Sejnowski. 1996. "A Framework for Mesencephalic Dopamine Systems Based on Predictive Hebbian Learning." *Journal of Neuroscience* 16 (5): 1936–47. https://doi.org/10.1523/JNEUROSCI.16-05-01936.1996.

Moradi, Sajjad, Atieh Mirzababaei, Alireza Dadfarma, Shahabeddin Rezaei, Hamed Mohammadi, Behrooz Jannat, and Khadijeh Mirzaei. 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., R. Nogueiras, M. L. Andermann, Z. B. Andrews, S. D. Anker, J. Argente, R. L. Batterham, et al. 2015. "Ghrelin." *Molecular Metabolism* 4 (6): 437–60. https://doi.org/10.1016/j.molmet.2015.03.005.

Muñoz-Escobar, Geovanni, Natalí N. Guerrero-Vargas, and Carolina Escobar. 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, Helen M., Donna J. Calu, Geoffrey Schoenbaum, and Melissa J. Sharpe. 2017. "The Dopamine Prediction Error: Contributions to Associative Models of Reward Learning." *Frontiers in Psychology* 8 (February). https://doi.org/10.3389/fpsyg.2017.00244.

Nauta, Johannes, Yara Khaluf, and Pieter Simoens. 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, Daniel, Clare Andrews, and Melissa Bateson. 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, Daniel, and Melissa Bateson. 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 (October): 1–9. https://doi.org/10.1155/2019/7174058.

Nguyen, Katrina P., Mohamed A. Ali, Timothy J. O'Neal, Ilona Szczot, Julia A. Licholai, and Alexxai V. Kravitz. 2017. "Feeding Experimentation Device (FED): Construction and Validation of an Open-Source Device for Measuring Food Intake in Rodents." *Journal of Visualized Experiments*, no. 120 (February): 55098. https://doi.org/10.3791/55098.

Nonacs, Peter. 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., Yves Kremer, Sandrine Lefort, Masaya Harada, Vincent Pascoli, Clément Rohner, and Christian Lüscher. 2015. "Accumbal D1r Neurons Projecting to Lateral Hypothalamus Authorize Feeding." *Neuron* 88 (3): 553–64. https://doi.org/10.1016/j.neuron.2015.09.038.

Ouden, H. E. M. den, J. Daunizeau, J. Roiser, K. J. Friston, and K. E. Stephan. 2010. "Striatal Prediction Error Modulates Cortical Coupling." *Journal of Neuroscience* 30 (9): 3210–19. https://doi.org/10.1523/JNEUROSCI.4458-09.2010.

Pandit, R., S. Beerens, and R. a. H. Adan. 2017. "Role of Leptin in Energy Expenditure: The Hypothalamic Perspective." *American Journal of Physiology-Regulatory, Integrative and Com-*

parative Physiology 312 (6): R938–47. https://doi.org/10.1152/ajpregu.00045.2016.

Pankevich, D. E., S. L. Teegarden, A. D. Hedin, C. L. Jensen, and T. L. Bale. 2010. "Caloric Restriction Experience Reprograms Stress and Orexigenic Pathways and Promotes Binge Eating." *Journal of Neuroscience* 30 (48): 16399–407. https://doi.org/10.1523/JNEUROSCI.1955-10.2010.

Parkes, Shauna L., Teri M. Furlong, Alanna D. Black, and Bernard W. Balleine. 2017. "Intermittent Feeding Alters Sensitivity to Changes in Reward Value." *Appetite* 113 (June): 1–6. https://doi.org/10.1016/j.appet.2017.02.009.

Payzan-LeNestour, Elise, Simon Dunne, Peter Bossaerts, and John P. O'Doherty. 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, John M., and Geoffrey Hall. 1980. "A Model for Pavlovian Learning: Variations in the Effectiveness of Conditioned but Not of Unconditioned Stimuli." *Psychological Review* 87 (6): 532–52. https://doi.org/10.1037/0033-295X.87.6.532.

Perez-Leighton, Claudio, Morgan R. Little, Martha Grace, Charles Billington, and Catherine M. Kotz. 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): R338–46. https://doi.org/10.1152/ajpregu.00339.2016.

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

Plaisier, Fabrice, Catherine Hume, and John Menzies. 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., M. C. Lacroix, K. Badonnel, S. Gougis, C. Baly, R. Salesse, and M. Caillol. 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): 1287–98. https://doi.org/10.1016/j.neuroscience.2009.05.043.

Puigbò, J.-Y., X. D. Arsiwalla, M. A. González-Ballester, and P. F. M. J. Verschure. 2020. "Switching Operation Modes in the Neocortex via Cholinergic Neuromodulation." *Molecular Neurobiology* 57 (1): 139–49. 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): 523–75. 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*, 596–600. Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780080453378002783.

Qi, Jia, Shiliang Zhang, Hui-Ling Wang, David J. Barker, Jorge Miranda-Barrientos, and Marisela Morales. 2016. "VTA Glutamatergic Inputs to Nucleus Accumbens Drive Aversion by Acting on GABAergic Interneurons." *Nature Neuroscience* 19 (5): 725–33. https://doi.org/10.1038/nn.4281.

Rada, P., N. M. Avena, and B. G. Hoebel. 2005. "Daily Bingeing on Sugar Repeatedly Releases Dopamine in the Accumbens Shell." *Neuroscience* 134 (3): 737–44. https://doi.org/10.1016/j.neuroscience.2005.04.043.

Raj, Vishnu, and Sheetal Kalyani. 2017. "Taming Non-Stationary Bandits: A Bayesian Approach." *arXiv:1707.09727 [Cs, Stat]*, July. http://arxiv.org/abs/1707.09727.

Ranc, Nathan, Paul R. Moorcroft, Federico Ossi, and Francesca Cagnacci. 2021. "Experimental Evidence of Memory-Based Foraging Decisions in a Large Wild Mammal." *Proceedings of the National Academy of Sciences* 118 (15): e2014856118. https://doi.org/10.1073/pnas.2014856118.

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

Rescorla, Robert A, Allan R Wagner, Abraham H Black, and William Frederick Prokasy. 1972. "Classical Conditioning II: Current Research and Theory."

Reverdy, Paul B., Vaibhav Srivastava, and Naomi Ehrich Leonard. 2014. "Modeling Human Decision Making in Generalized Gaussian Multiarmed Bandits." *Proceedings of the IEEE* 102 (4): 544–71. https://doi.org/10.1109/JPROC.2014.2307024.

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

Reynolds, Andy, Eliane Ceccon, Cristina Baldauf, Tassia Karina Medeiros, and Octavio Miramontes. 2018. "Lévy Foraging Patterns of Rural Humans." Edited by Alastair Smith. *PLOS ONE* 13 (6): e0199099. https://doi.org/10.1371/journal.pone.0199099.

Rothwell, Nancy J., and Michael J. Stock. 1988. "The Cafeteria Diet as a Tool for Studies of Thermogenesis." *The Journal of Nutrition* 118 (8): 925–28. https://doi.org/10.1093/jn/118.8.925.

Ryali, Chaitanya K., Gautam Reddy, and Angela J. Yu. 2016. "Demystifying Excessively Volatile Human Learning: A Bayesian Persistent Prior and a Neural Approximation." Neuroscience. http://biorxiv.org/lookup/doi/10.1101/077719.

Sakurai, Takeshi, Ruby Nagata, Akihiro Yamanaka, Hiroko Kawamura, Natsuko Tsujino, Yo Muraki, Haruaki Kageyama, et al. 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, Anna C., Karl J. Friston, Matthew W. Jones, Anthony E. Pickering, and Rosalyn J. Moran. 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, Anne-Marike, Christiane Ahlheim, Moritz F. Wurm, and Ricarda I. Schubotz. 2012. "Surprised at All the Entropy: Hippocampal, Caudate and Midbrain Contributions to Learning from Prediction Errors." Edited by Stefan J. Kiebel. *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, Jerome M. 2004. "Hypocretin (OREXIN): Role in Normal Behavior and Neuropathology." *Annual Review of Psychology* 55 (1): 125–48. https://doi.org/10.1146/annurev.psych.55.090902.141545.

Sims, David W, Nicolas E Humphries, Nan Hu, Violeta Medan, and Jimena Berni. 2019. "Optimal Searching Behaviour Generated Intrinsically by the Central Pattern Generator for Locomotion." *eLife* 8 (November): e50316. https://doi.org/10.7554/eLife.50316.

Soltani, Alireza, and Alicia Izquierdo. 2019. "Adaptive Learning Under Expected and Unexpected Uncertainty." *Nature Reviews Neuroscience* 20 (10): 635–44. https://doi.org/10.1038/

s41583-019-0180-y.

Steinberg, Elizabeth E., Ronald Keiflin, Josiah R. Boivin, Ilana B. Witten, Karl Deisseroth, and Patricia H. Janak. 2013. "A Causal Link Between Prediction Errors, Dopamine Neurons and Learning." *Nature Neuroscience* 16 (7): 966–73. https://doi.org/10.1038/nn.3413.

Stinson, Emma J., Susanne B. Votruba, Colleen Venti, Marisol Perez, Jonathan Krakoff, and Marci E. Gluck. 2018. "Food Insecurity Is Associated with Maladaptive Eating Behaviors and Objectively Measured Overeating: Food Insecurity and Overeating." *Obesity* 26 (12): 1841–48. https://doi.org/10.1002/oby.22305.

Stolyarova, Alexandra, and Alicia Izquierdo. 2017. "Complementary Contributions of Basolateral Amygdala and Orbitofrontal Cortex to Value Learning Under Uncertainty." *eLife* 6 (July): e27483. https://doi.org/10.7554/eLife.27483.

Stratford, Thomas R., and Ann E. Kelley. 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): 11040–48. https://doi.org/10.1523/JNEUROSCI.19-24-11040.1999.

Sturgill, J. F., P. Hegedus, S. J. Li, Q. Chevy, A. Siebels, M. Jing, Y. Li, B. Hangya, and A. Kepecs. 2020. "Basal Forebrain-Derived Acetylcholine Encodes Valence-Free Reinforcement Prediction Error." *bioRxiv*, February, 2020.02.17.953141. https://doi.org/10.1101/2020.02.17.953141.

Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. Second edition. Adaptive Computation and Machine Learning Series. Cambridge, Massachusetts: The MIT Press.

Thompson, William 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, Ju, Ryan Huang, Jeremiah Y. Cohen, Fumitaka Osakada, Dmitry Kobak, Christian K. Machens, Edward M. Callaway, Naoshige Uchida, and Mitsuko Watabe-Uchida. 2016. "Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons." *Neuron* 91 (6): 1374–89. https://doi.org/10.1016/j.neuron.2016.08.018.

Tokic, Michel, and Günther Palm. 2011. "Value-Difference Based Exploration: Adaptive Control Between Epsilon-Greedy and Softmax." In *KI 2011: Advances in Artificial Intelligence*, edited by Joscha Bach and Stefan Edelkamp, 7006:335–46. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-642-24455-1_33.

Toshinai, Koji, Yukari Date, Noboru Murakami, Mitsushi Shimada, Muhtashan S. Mondal, Takuya Shimbara, Jian-Lian Guan, et al. 2003. "Ghrelin-Induced Food Intake Is Mediated via the Orexin Pathway." *Endocrinology* 144 (4): 1506–12. https://doi.org/10.1210/en.2002-220788.

Tyree, Susan M., Jeremy C. Borniger, and Luis de Lecea. 2018. "Hypocretin as a Hub for Arousal and Motivation." *Frontiers in Neurology* 9 (June): 413. https://doi.org/10.3389/fneur.2018.00413.

Viswanathan, G. M., V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley. 1996. "Lévy Flight Search Patterns of Wandering Albatrosses." *Nature* 381 (6581): 413–15. https://doi.org/10.1038/381413a0.

Wang, Zhendong, and Mingyuan Zhou. 2020. "Thompson Sampling via Local Uncertainty." *arXiv:1910.13673 [Cs, Stat]*, August. http://arxiv.org/abs/1910.13673.

Watabe-Uchida, Mitsuko, Neir Eshel, and Naoshige Uchida. 2017. "Neural Circuitry of Reward Prediction Error." *Annual Review of Neuroscience* 40 (1): 373–94. https://doi.org/10.1146/

annurev-neuro-072116-031109.

Wilson, Robert C, and Anne GE Collins. 2019. "Ten Simple Rules for the Computational Modeling of Behavioral Data." *eLife* 8 (November): e49547. https://doi.org/10.7554/eLife.49547.

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

Wojnicki, F. H. E., D. S. Johnson, G. Charny, and R. L. W. Corwin. 2015. "Development of Bingeing in Rats Altered by a Small Operant Requirement." *Physiology & Behavior* 152 (December): 112–18. https://doi.org/10.1016/j.physbeh.2015.09.009.

Wojnicki, F. H. E., J. G. Stine, and R. L. W. Corwin. 2007. "Liquid Sucrose Bingeing in Rats Depends on the Access Schedule, Concentration and Delivery System." *Physiology & Behavior* 92 (4): 566–74. https://doi.org/10.1016/j.physbeh.2007.05.002.

Wosniack, Marina E., Marcos C. Santos, Ernesto P. Raposo, Gandhi M. Viswanathan, and Marcos G. E. da Luz. 2017. "The Evolutionary Origins of Lévy Walk Foraging." Edited by Frederic Bartumeus. *PLOS Computational Biology* 13 (10): e1005774. https://doi.org/10.1371/journal.pcbi.1005774.

Wu, Qingyun, Naveen Iyer, and Hongning Wang. 2018. "Learning Contextual Bandits in a Non-Stationary Environment." In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 495–504. Ann Arbor MI USA: ACM. https://doi.org/10.1145/3209978.3210051.

Yamanaka, Akihiro, Carsten T Beuckmann, Jon T Willie, Junko Hara, Natsuko Tsujino, Michihiro Mieda, Makoto Tominaga, et al. 2003. "Hypothalamic Orexin Neurons Regulate Arousal According to Energy Balance in Mice." *Neuron* 38 (5): 701–13. https://doi.org/10.1016/S0896-6273(03)00331-3.

Yoon, Hyung-Jin, Donghwan Lee, and Naira Hovakimyan. 2018. "Hidden Markov Model Estimation-Based Q-Learning for Partially Observable Markov Decision Process." *arXiv:1809.06401 [Cs, Stat]*, September. http://arxiv.org/abs/1809.06401.

Yu, Angela J., and Peter Dayan. 2005. "Uncertainty, neuromodulation, and attention." *Neuron* 46 (4): 681–92. https://doi.org/10.1016/j.neuron.2005.04.026.

Zaghloul, Kareem A., Justin A. Blanco, Christoph T. Weidemann, Kathryn McGill, Jurg L. Jaggi, Gordon H. Baltuch, and Michael J. Kahana. 2009. "Human substantia nigra neurons encode unexpected financial rewards." *Science (New York, N.Y.)* 323 (5920): 1496–99. https://doi.org/10.1126/science.1167342.

Zajo, Kristin N., Jim R. Fadel, and Joshua A. Burk. 2016. "Orexin A-Induced Enhancement of Attentional Processing in Rats: Role of Basal Forebrain Neurons." *Psychopharmacology* 233 (4): 639–47. https://doi.org/10.1007/s00213-015-4139-z.

Zink, A. N., P. E. Bunney, A. A. Holm, C. J. Billington, and C. M. Kotz. 2018. "Neuromodulation of Orexin Neurons Reduces Diet-Induced Adiposity." *International Journal of Obesity* 42 (4): 737–45. https://doi.org/10.1038/ijo.2017.276.