# Theoretical and empirical framework

## The natural setting for food-seeking behavior

Foraging comprises the complete set of activities and behaviors related to

obtaining food in a wild environment. Food-seeking behavior is a particular

element of such a set, including all re-orientation and locomotion activity

related to the obtention of food. Thus, food-seeking behavior considers up to

food encounter, whereas foraging behavior is still present in future utilization

of acquired energy, including feeding or possibly hoarding

[@V3Z2UVAU#Kramer\_2001].

The food-seeking phase of foraging must ensure an optimal way to acquire food

using the least amount of resources and reduce exposure to potential predators.

If the food-resources location were static, food-seeking behavior only necessary

input would be an initial sampling of the environment and then matching

landscape cues. This is not the case. Animals do not necessarily follow

landscape cues [@9XCDNBAM#Bartumeus\_Etal\_2016], or even develop search

strategies based on them [@BWKDXXFW#Kölzsch\_Etal\_2015]. Moreover, animals are

subject to incomplete knowledge about resource location, quality, and

probability of obtention [@ZD73QGIR#Pyke\_1984 ]. Thus a foraging animal must

determine its food-seeking behavior considering an inherently stochastic

environment with only partial knowledge.

In a stochastic environment, to establish optimal food-seeking strategies,

animals should consider the overall statistical properties of the environment.

Otherwise, local environment volatility could lead to the misguided preference

for lower mean quality food resources with high variability, leading to

starvation in the long run. Empirical evidence has shown that multiple animal

species, including humans, perform searching in a Lévy-walk fashion

[@7YQKP7Z2#Garg\_Kello\_2021; @I2BS842S#Reynolds\_Etal\_2018;

@TPRPLPEC#Viswanathan\_Etal\_1996; @BWKDXXFW#Kölzsch\_Etal\_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; such patterns produce optimal searches in various

environments with dispersed resources in a patchy-fashion

[@97UESCC6#Wosniack\_Etal\_2017]. Although its generative mechanism is not precise

[@I2BS842S#Reynolds\_Etal\_2018] there is evidence that this mechanism is

partially independent of sensory information [@M5RXPXSZ#Humphries\_Sims\_2014;

@5WUMQR2H#Sims\_Etal\_2019], probably selected through evolution as it optimizes

food searching with partial or complete lack of knowledge

[@97UESCC6#Wosniack\_Etal\_2017].

Given that these food-seeking strategies are present without sensory information

and are ubiquitous in animals, food-seeking behavior probably evolved to deal

with partial knowledge in an uncertain environment. While Lévy-walks provide a

'basal' strategy when there is partial or no knowledge, upon food encounter, the

search strategy switches to a more focused one similar to Brownian-motion

[@5KMWW8NS#Reynolds\_Frye\_2007; @F9HICU4A#Nauta\_Khaluf\_Simoens\_2020].

Furthermore, computational modeling points out how this switch between informed

(Brownian-like) and random search might depend on food encounter uncertainty

[@3YWCKUUK#Anselme\_Otto\_Güntürkün\_2017]. Together this data suggests that

animal's food-seeking behavior evolved to deal with uncertain environments and

partial knowledge. Moreover, environment uncertainty itself modulates the

baseline strategy, optimally searching for food even when knowledge is not

complete.

## Foraging and uncertainty

Specific food-seeking strategies, such as the Lévy-Walks, emerge from the

animal's irreducible uncertainty when foraging. Uncertainty then informs

food-seeking behavior. However, how animals capture o process this uncertainty

while searching for food has not yet been explained; such is the aim of this

section.

If an animal performs any food-seeking strategy in a specific area of the

environment, there is a probability of success, depending on resource density

and the specific strategy. With a fixed trategy, resource density determines the

probability. Moreover, as the resource density changes over time, so does the

probability of successfully obtaining food. Moreover, in conjunction with

limited perceptual abilities [@9XCDNBAM#Bartumeus\_Etal\_2016 ], the environment

appears as being of stochastic nature [@DIMJNJV2#Caraco\_Etal\_1990].

An animal that does not consider environment stochastic nature in its

food-seeking behavior will act greedily upon sampled values. Acting greedily

implies that the animal will always choose the option that yielded the most

values in an initial random sampling. Thus, acting greedily is analogous to

neglecting that the sample comes from a stochastic distribution of rewards. This

is problematic as it makes food-seeking strategy insensitive to reward variation

[@2BEHEM7X#Sutton\_Barto\_2018]. One could relax this assumption and propose that

animals act upon the mean rewards, such as the classical model by

@ESYGCSLH#Charnov\_1976, which determines the strategy based on the current patch

value against the global mean of the environment. Nevertheless, such rule-based

models do not provide good fit to behavioral data [@GC6MVWQU#Nonacs\_2001;

@9XVRLKC3#LeHeron\_Etal\_2020; @RR87DVIX#Pyke\_2010;

@H9JQCFZA#Hayden\_Pearson\_Platt\_2011].

The competing models to rule-based ones are those which effectively consider

uncertainty into its formulation. However, for this to make biological sense,

animals should be able to track uncertainty measures precisely. Risk considers

the spread of possible outcomes, or similarly, the standard deviation of the

expected outcome [@VACKG3ZK#Rothwell\_Stock\_1988]. In humans, the anterior

cingulate cortex (ACC) tracks risk [@GLI8DY99#Christopoulos\_Etal\_2009].

Moreover, ACC tracks risk in a context-dependent fashion; that is, it considers

cue-related information to determine the risk and expected value of a given

option [@5ANLDC83#VanHolstein\_Floresco\_2020]. While risk is the component of

uncertainty that measures the spread of outcome, volatility signifies how often

an environment changes its contingencies, for example, in the action and outcome

probability pairing. In learning tasks, volatility increases ACC activity

[@BHR2NAEI#Behrens\_Etal\_2007]. The main goal of tracking uncertainty is to

augment the chance of success while searching for food. In that regard,

filtering out uncertainty regarding outcomes can prove beneficial as the actual

outcome prediction does not become affected by noise.

@X6XAHBZA#Stolyarova\_Izquierdo\_2017 showed that rats could optimally choose

options with more significant value (mean reward waiting time) despite large

associated variability. Furthermore, lesions in the orbitofrontal cortex (OFC)

showed an impaired ability to change preference when the mean rewards were up or

downshifted; that is, value inference became noisy.

All previously mentioned aspects of uncertainty can be categorized within the

notion of expected and unexpected uncertainty. @4BJ2B6KB#Yu\_Dayan\_2005 proposed

expected uncertainty as the uncertainty regarding outcomes when contingencies

(outcome given a particular action) remain stable but are subject to some noise.

In contrast, unexpected uncertainty represents a drastic change in the

contingencies that likely reflects a structural change in the environment. For

example, @X6XAHBZA#Stolyarova\_Izquierdo\_2017 experiment measured rat behavioral

modulation with expected uncertainty, and @BHR2NAEI#Behrens\_Etal\_2007 showed ACC

activity increases in the case of unexpected uncertainty. Considering both kinds

of uncertainties enables the animal to balance top-down and bottom-up

information if the obtained rewards present some variation is should not modify

learned contingencies that maps actions with rewards in a given environment, so

top-down control should dominate bottom-up input. On the other hand, if obtained

rewards present a large amount of variation, the balance should switch in order

to prioritize bottom-up input to increase learning about new contingencies

[@4BJ2B6KB#Yu\_Dayan\_2005; @P2FYNJKR#Soltani\_Izquierdo\_2019].

@4BJ2B6KB#Yu\_Dayan\_2005 dual notion of uncertainty can be considered in terms of

stationary and non-stationary environments. A non-stationary environment is one

where the outcome variance remains unchanged, whereas non-stationary presents a

variable variance [@UJSWSGH3#Wu\_Iyer\_Wang\_2018]. Thus, non-stationary

environments should be the only cause of unexpected uncertainty, and stationary

environments should only have expected uncertainty. However, while being subject

only to expected uncertainty, humans typically behave as the environment were

non-stationary, thus producing unexpected uncertainties

[@77AMCAE4#Ryali\_Reddy\_Yu\_2016]. While the functional reasons behind these

behaviors are not clear, it has been proposed that this emerge because

sustaining the belief that environments are non-stationary does not prove to be

problematic in stationary ones [@77AMCAE4#Ryali\_Reddy\_Yu\_2016] or because

stochasticity in decisions may provide a sufficient heuristic in many natural

environments [@ERS4UNTK#Reverdy\_Srivastava\_Leonard\_2014].

Under non-stationary environments, contingencies change, so animals are faced

with the dilemma of either exploiting or exploring

[@2BEHEM7X#Sutton\_Barto\_2018]. Exploiting means that behavior should be

consistent with previously learned reward contingencies. On the other hand,

exploring tries to re-sample the environment to improve or re-learn current

contingencies. The exploitation-exploration balance has been linked to the

expected and unexpected uncertainty [@T3QJH2AJ#Cohen\_Mcclure\_Yu\_2007], as

unexpected uncertainty should increase exploratory behavior to boost learning of

new contingencies, whereas if only expected uncertainty is present, behavior

should exploit current knowledge. However, exploration could be triggered by

'boredom' when environment properties are extensible learned

[@LEWESIS6#AstonJones\_Cohen\_2005].

Here we briefly exposed the different environmental properties of food-seeking

behavior to generate good strategies and maximize gain. We proposed that natural

environments are inherently stochastic, and animals adapted to sense various

aspects of uncertainty regarding their actions within the environment. However,

such pivotal importance of uncertainty in food-seeking behavior could be

contested when considering perceptual abilities as the primary means of

informing foraging.

Perceptual abilities are always limited in some regard, and such limitation can

be regarded as the perceptual range of a given animal

[@C377F7EP#Fletcher\_Etal\_2013]. If an animal must know what is beyond such

perceptual range, displacement is needed. However, to inform such displacement

memory, perceptual information integration or other strategies should come into

play. In many classical models, animals can somehow integrate information about

resource quality and distribution into an environment mean

[@ESYGCSLH#Charnov\_1976]. However, perceptual omniscience is not the case, and

integration of environmental information is dependent on times, so including

non-local information (outside perceptual range) such as resource gradients

prove helpful for the foraging animal [@LDCMV4VS#Fagan\_Etal\_2017].

Even when local and non-local information can improve foraging success, there is

still an issue on how perceptual information behaves regarding movement and

spatial distance. The speed-perception tradeoff describes how perceptual

abilities are degraded as speed is increased

[@KU6TMHRT#Campos\_Bartumeus\_Méndez\_2013], and as rapid approaching speeds are

required to capture prey or obtain food resources while relying on local

perceptual information [@LDCMV4VS#Fagan\_Etal\_2017] this tradeoff present an

issue to food-seeking behavior. Moreover, the intensive-extensive tradeoff

points how finding food-resource nearby impairs finding resources far-away

[@VYE5NZFU#Raposo\_Etal\_2011; @9XCDNBAM#Bartumeus\_Etal\_2016 ]. Both tradeoffs

imply that in order to properly obtain knowledge about the environment and

actually achieve success in obtaining food leads to perceptual uncertainty and

requires a fine balance between exploration and exploitation.

In order to deal with such tradeoffs, animals establish two distinct behavioral

modes: (a) local search and (b) relocation. Local search is predominantly

informed by perceptual information, while the relocation behavior show signs of

a stochastic process with Lévy-like distributions [@W7DBK8JN#Bazazi\_Etal\_2012].

While there has been discussion of Brownian-like random movements guiding local

search, this is most likely an emergent property caused by frequent food

encounters [@HJIF9SNA#DeJager\_Etal\_2014], so this shifts from a random

relocation process to a perceptually guided local search, are in part result of

an increased frequency of food encounters. Thus, the overall food-seeking

strategy derives from a combination of random and perceptually informed

movements [@KCAKLD2E#Balogh\_Etal\_2020].

Given the limitation in perceptual abilities, uncertainty seems to be

inescapable. Even when an experienced animal can integrate optimal foraging

paths in non-stationary environments, random searches with distinct cycles of

exploration/exploitation phases persist [@GNGBMVLA#Kembro\_Etal\_2019].

Furthermore, introducing stochasticity in food-seeking behavior improves success

as it makes strategies more resilient to cognitive errors derived from

perception [@YTHQBTQH#Campos\_Etal\_2020]. The persistence of strategies that

permits the balance between exploitation, exploration, and relocation

[@KU6TMHRT#Campos\_Bartumeus\_Méndez\_2013], even when they are not technically

needed [@77AMCAE4#Ryali\_Reddy\_Yu\_2016] shows how relevant is uncertainty in the

development of food-seeking behavior. In this section, we presented how

stochastic properties of the environment lead to behavioral adaptations to deal

with the resultant uncertainty by favoring strategies that favor the reduction

of uncertainty [@QAN6C9AM#Peters\_Mcewen\_Friston\_2017]. In the next section, we

will examine models that consider the case of foraging in uncertain

environments, which informs about the underlying processes in food-seeking

behavior.

## Foraging models and underlying processes in food-seekings behavior

One of the way classical models such as [@ESYGCSLH#Charnov\_1976] dealt with

modeling foraging in uncertain environments was with the assumption of perfect

knowledge. Animals should stay seeking food within a patch for as long the

capture rate is above the capture rate of the environment

[@ESYGCSLH#Charnov\_1976], which implicitly assumes that somehow the animal is

able to compute such capture rate. While such assumptions may sound

unrealistic, there is some support for this as an experienced forager may learn

and integrate information about the environment to closely approximate the

perfect knowledge [@9X7Z6PMX#Marshall\_Etal\_2016].

On the other hand, and in consequence of the previously exposed relation

between foraging and uncertainty, a model presented here should account for

such relation. First, the rules determining the results of the interaction

between animal and environment are assumed to be unknown or only partially

known due to the stochastic nature of the environment. Then, 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 $q(s, a)$. Such model of

action, state, and value corresponds to a Markov decision process

[@2BEHEM7X#Sutton\_Barto\_2018]. 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 $s$. We could consider a Markov decision process to include the

perceptual noise which we deemed inherent to food-seeking behavior by

considering that states $s$ is paired with an observation $o$ made by the

animal to infer state $s$, because the state cannot be directly observed or there

is some sensory noise. As such, animals consider environment states as

the conditional probability of any particular observation given a state $p(o | s)$,

giving a belief of the current state based on perceptual information

[@W73P5HZ9#Ma\_Jazayeri\_2014].

To model how an animal represents the value of a given option $q(s, a)$ in a

non-stationary environment, this value is a distribution over possible values

that is updated every time an action $a$ is executed. For the simple case were

rewards are obtained or not $q(s, a)$ has a Bernoulli distribution $p(X =

reward) = a$ and $(p(X = no reward) = 1 - a)$. Then, these probabilities can be

modeled with the Beta distribution, which takes parameters $\alpha$ and $\beta$.

With $\alpha = 1, \beta = 1$ the Beta distributions produces a uniform

distribution over $[0, 1]$ successfully representing the uninformed prior

probability for the rewards. To generate the posterior probability every time

the reward process results in a reward, the parameter $\alpha$ increases by 1.

On the other hand, if no reward is obtained, the parameter $\beta$ increases by 1.

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 $s$. If

the exploration is defined by the posterior, then it can be considered a

Thompson sampling strategy [@ZZ9I6KCZ#Thompson\_1933]. To select an action $a$ a

posterior is built for every action and updated according to the previously

stated rules, then for each posterior, a reward estimate $\hat{r}$ is sampled

greedily so the action selected is $a = argmax\_{a \in A} \hat{r}(a)$ where $A$

is the set of possible actions within an environment [@WFYYPZ3N#Wang\_Zhou\_2020].

This processes must be performed for every state, limiting tractability by the

the number of states. In general terms, a solution fo 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 more weight is given to recent

rewards, on the other hand, if its closer to 1, then all the reward history is

equally considered. More complex consideration of this problem include modeling

non-stationarity as Poisson arrival process that modifies the means rewards

[@2S4JPDRG#Ghatak\_2020], bayesian approaches to modulate past observed rewards

[@VPX6THEN#Raj\_Kalyani\_2017], and explicitly modeling environment volatility in

a bayasian setup [@BHR2NAEI#Behrens\_Etal\_2007].

While this general model can work in non-stationary environments, it doesn't

consider explicitly the belief of the current state based on the perceptual

information received $p(o | s)$. For this addition, 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

finally 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 $p(o | s)$ and bayesian,

among other methods for obtaining such model given only actions and observation

has been proposed [@B9YLCVG6#Funamizu\_Etal\_2012;

@XBHNGIHT#Yoon\_Lee\_Hovakimyan\_2018; @TBQJ5HNA#Piray\_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 or noise. Thompson sampling was

considered as the base for this due to its simplicity and elegance in modeling

exploration/exploitation by computing uncertainty. The goal of these

consideration was not to establish or to specify a complete model but to

provide a framework relating uncertainty with the exploration/exploitation

dilemma and perceptual limitations shown theoretically and empirically in the

previous section.

## Computing uncertainty

Uncertainty arises from having more than one option, and that the motivation to

opt for one of those options is somewhat distributed, and there is no one option

that is always preferred. Considering that the probability of choosing any given

option has a uniform distribution, then uncertainty increases proportionally

with the number of options. Shannon entropy [@LUFF6VTC#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 $s$, one action $a$

initiate a food-seeking bout, and only two possible outcomes food is found

($p$) or not found ($q = 1 - p$), then $H = -(p log\_{2} p + q log{2} q)$. If an

animal performs multiple food-seeking bouts and non 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)

for the particular implementation of encoding outcome entropy

[@55XGJCH7#Goñi\_Etal\_2011; @DRJBKPPI#Gloy\_Herrmann\_Fehr\_2020] so this

computation seems to be biologically plausible. However, entropy is not

available as sensory input must derive from actions and outcomes, which are

dependent on the environment state. Previously, through Thompson sampling, we

provided a way in that 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 [@LFYTIBBR#Rescorla\_Etal\_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}, but weighted

by a learning rate $\alpha$ and a prediction error $\delta$. $delta$ 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 has

proposed by [@2BEHEM7X#Sutton\_Barto\_2018] where the prediction error consider an

estimate of the rewards that give more weight to current rewards while still

considering past rewards $$ \delta\_{n} = R{t} + \gamma \hat{V\_{n+1}} -

\hat{V\_{n}} $$ $\gamma$ is a discount factor $0 <= \gamma < 1$ for all the

history of rewards, and $\hat{V}$ is a proxy for the true value of the reward.

Finally, $\alpha : [0, 1]$ is the learning rate which effectively weights the

reward prediction error $\delta$ so to make small updates $\alpha \approx 0$ or

rather large ones $\alpha \approx 1$.

The simple model presented allows deriving a prediction error based on

experience and the learning rate can be set lower to simulate unexpected

uncertainty or higher to simulate expected uncertainty. However, $\alpha$ in

such model is a hyperparameter, thus not derived from experience.

@NZFTTQJZ#Pearce\_Hall\_1980 model proposes that $\alpha$ 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 minimal

error probability [@7JBVDKC8#Feder\_Merhav\_1994], thus increasing $|\delta|$, and

consequently $\alpha$. The behavioral result is that the animal should increase

learning for options with the uncertain outcome by directing its attention

[@BVC98GTV#Diederen\_Fletcher\_2021].

Dopamine (DA) encodes prediction error [@8C97FJFI#Nasser\_Etal\_2017;

@AR2TQB84#Fiorillo\_2003; @5BR3FL7N#Fiorillo\_2011;

@IV9MZSXR#Lak\_Stauffer\_Schultz\_2014; @HLCSQTJB#Glimcher\_2011;

@4YU7F96V#Khaw\_Glimcher\_Louie\_2017; @34HNDDL4#Gershman\_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 [@AR2TQB84#Fiorillo\_2003].

Effectively encoding the prediction of the unconditioned stimulus, and the

surprise for the unconditioned stimulus, if a reward has a low probability

obtaining it should be 'surprising'. Moreover, the change is sustained

activation encodes rewards probability analogous to entropy, that is, displaying

a peak of activity at probability 0.5 [@AR2TQB84#Fiorillo\_2003 Moreover, the

change in sustained activation encodes rewards probability analogous to entropy,

that is, displaying a peak of activity at a probability of 0.5

[@AR2TQB84#Fiorillo\_2003]. Hippocampal activity has been shown to reflect

Shannon entropy and adaptation predicted by prediction error minimization

[@7PXFUSRW#Schiffer\_Etal\_2012], similar activity is also present on the striatum

[@6FBLSRK9#DenOuden\_Etal\_2010], substantia nigra [@6IGYU34R#Zaghloul\_Etal\_2009],

and ventral tegmental area [@XHH632AS#Iordanova\_Etal\_2021]. Moreover, DA

activity fits the classical reinforcement models as ventral tegmental DA support

cue-reward learning, the modifications of previous cue-reward associations

[@LSWYXCBD#Steinberg\_Etal\_2013; @8FIHXB6G#Chakroun\_Etal\_2020], and capable of

dealing with exploration/exploitation via tonic and phasic signaling,

respectively [@TN6NBEH8#Beeler\_Etal\_2010].

There is substantial evidence that DA neurons, specifically in the VTA, serve

the functional role of computing reward prediction errors

[@NLDHLRVN#WatabeUchida\_Eshel\_Uchida\_2017 ], by weighing inputs from multiple

brain areas, most remarkably the lateral hypothalamus, dorsal and ventral

the striatum, ventral pallidum, and subthalamic nucleus [@SFZIJKFP#Tian\_Etal\_2016].

However, acetylcholine (ACh) and norepinephrine (NA) associated with expected

and unexpected uncertainty, respectively [@4BJ2B6KB#Yu\_Dayan\_2005], which are

mainly produced in the basal forebrain [@D8EGNYCV#Sturgill\_Etal\_2020] and locus

coeruleus (LC) [@TA5KB3TF#Sales\_Etal\_2019; @LEWESIS6#AstonJones\_Cohen\_2005]. ACh

antagonism has been shown to increase the response sensitivity to expected

uncertainty within a task [@9X7Z6PMX#Marshall\_Etal\_2016], providing evidence

that ACh represents expected uncertainty. On the other hand, NA for new

contingencies. On the other hand, LC tonic activity represents unexpected

uncertainty [@9Z525EYW#PayzanLenestour\_Etal\_2013;

@LEWESIS6#AstonJones\_Cohen\_2005].

In this section, we presented simple computational models that consider

uncertainty by using the information provided by the reward prediction error, which

is an extremely simple computation that is likely to be implemented by DA

activity in the VTA, with additional modulation by ACh and NA possibly

controlling the sensitivity of DA to expected and unexpected uncertainty. In

In the following two sections, we will show empirical evidence on how food-access

uncertainty increases food-seeking behavior, and propose orexin as a potential

mediator of uncertainty-driven foraging.

## An adaptive strategy in modern times

Natural environments are limited in food resources, and food-seeking behavior

results from an adaptation to such environments. If the characteristic property

is scarcity, then animals should approach cues with the highest associative

strength to actual food resources, this approach

[@B8QZIGEN#Montague\_Dayan\_Sejnowski\_1996] suggests that prediction error signal

if the current state is better or worse than expected, so animals should prefer

options with the highest expected value [@8Q6QHXGB#Kacelnik\_Bateson\_1996].

However, uncertainty in food delivery increases lever pressing and reduces the time

latency to approach lever [@S8CHV5KG#Anselme\_Robinson\_Berridge\_2013]. Creating

intermittent access to high-fat diets generates binge-eating behavior

[@QN55S25U#Hess\_Etal\_2019; @CTAFJAZ3#King\_Etal\_2016;

@MBZJPLN3#Lardeux\_Kim\_Nicola\_2013], increases reinforcement value upon withdrawal

[@BDVQLHQZ#Mcgee\_Etal\_2010] and operant behavior without withdrawal

[@8KGR8TT6#Wojnicki\_Babbs\_Corwin\_2013; @DQHTA2QH#Wojnicki\_Stine\_Corwin\_2007;

@WPLVSADN#Wojnicki\_Etal\_2015], psychomotor behavior

[@GIDJCJMP#Hardaway\_Etal\_2016], DA and Ach release

[@23K7IPBA#Rada\_Avena\_Hoebel\_2005]. Mainly because of logistic reasons

intermittent access is provided at the same time of days (in most cases), and

that allows animals to accurately predict the arrival of food, although with some

inconsistencies within animals [@KUXWM8W5#Luby\_Etal\_2012]. However, the

behavioral effects, except anticipatory behavior, are common if intermittent

access is completely random or given at certain times of the day

[@VY8ZEZB6#MuñozEscobar\_GuerreroVargas\_Escobar\_2019].

Intermittent feeding schedules, and in general uncertainty of food-access

disrupts eating patterns. Food-seeking behavior is increased to avoid starvation

when a food shortage is predicted, and as previously noted, this derives in

increasing exploration, amount of foraging bouts, and time expended in foraging

[@G83L8BXA#Harris\_Chapman\_Monfort\_2010]. This could be translated to the concept

of food insecurity, which defines the perception of how secure food access is

[@WZS8BTVL#Dhurandhar\_2016], which is positively correlated with positive energy

balance [@WZS8BTVL#Dhurandhar\_2016], increasing preference for high-fat

alternative [@EER2TNCJ#Nettle\_Andrews\_Bateson\_2017], which corresponds to the

food of cheap access in developed countries where this effect is more pronounced

[@4LJKTR3N#Moradi\_Etal\_2019] in females [@2N6SBNMK#Dinour\_Bergen\_Yeh\_2007;

@VDHUFYYV#Nettle\_Bateson\_2019]. The strategy responsible for overweight in

food-insecure individuals are to overeat fats and carbohydrates in periods of

high food availability [@2ZRH7IWR#Stinson\_Etal\_2018].

In modern urban environments, high-fat food is of easy access, coupled with a

food-seeking behavior which seeks to maximize energetic gain, when food shortage

is predicted due to food-access uncertainty, can create overweight in the

population because the mechanism is adapted to low resource environments,

however, in developed countries, caloric density is extremely high, so increasing

food-seeking behavior results in excessive caloric intake. 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

[@52DV95LU#Burdakov\_2020], and motivated behavior

[@2X7SNKS3#Tyree\_Borniger\_DeLecea\_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 [@7BNMJXRJ#Davies\_2016]. The hypothalamus

comes as a relevant structure in the homeostatic process by being capable of

controlling arousal levels [@3ZQDGXXZ#Adamantidis\_Carter\_DeLecea\_2010;

@HQFDMJJ8#Kosse\_Burdakov\_2014], motivation for food intake

[@6FUPLHIJ#Castro\_Berridge\_2017], receiving internal status information of fat

deposits via leptin signaling [@RWABER6F#Pandit\_Beerens\_Adan\_2017;

@DHNLNHHI#Meister\_2000], and gastrointestinal status via ghrelin

[@NRWCYJCX#Müller\_Etal\_2015; @W3IUQNUY#Toshinai\_Etal\_2003]. Together, this

functionality takes the hypothalamus to a pivotal role in the homeostasis

process, specifically relating to controlling food intake.

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 [@EFWJQ65B#Chowdhury\_Etal\_2019]. Its

functions range from regulating sleep/wakefulness states

[@AR3R5QLP#Chemelli\_Etal\_1999] and energetic balance

[@KPIHYUYF#Yamanaka\_Etal\_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

[@7GXEINLY#González\_Etal\_2016], activity increases upon sucrose predictive cues

[@PIJIVQER#Hassani\_Etal\_2016], and in certain subpopulation, the increased

spontaneous physical activity is directed towards food sources

[@Y4UZKQ5V#Zink\_Etal\_2018]. Thus, orexin-related activity can be interpreted as

food procuring [@Y4UZKQ5V#Zink\_Etal\_2018], further support for this

interpretation come from orexin increasing olfactory activity

[@X9STRHXX#PrudHomme\_Etal\_2009], enhancing visual attention

[@7ZXMSXYC#Zajo\_Fadel\_Burk\_2016], impairment of spatial working memory in orexin

knockout mice [@Z2G6BP2E#Dang\_Etal\_2018], among others that have been classified

as foraging-related behavior by @LBGV5NJ5#Barson\_2020.

Orexin role in food-seeking behavior is not sufficient to suggest its role in

uncertainty-driven food-seeking behavior, however its connectivity to VTA, LC

and basal forebrain [@PQWIJHAP#Siegel\_2004], which corresponds to inputs into DA,

Ach and NE activity, respectively, might hint at orexin role in promoting

food-seeking behavior for prediction-error derived uncertainty, expected and

unexpected uncertainty. VTA projects glutamatergic inputs into the nucleus

accumbens shell, specifically into parvalbumin GABAergic interneurons,

activation of such neurons results in inhibited medium spiny neurons activity

into the lateral hypothalamus [@IUF535G9#Qi\_Etal\_2016], these connections are

known to be inhibitory [@TIF7X9CT#O’Connor\_Etal\_2015;

@M2RERH96#PerezLeighton\_Etal\_2017], and input inhibition increases lateral

hypothalamus activity [@VD2436JM#Stratford\_Kelley\_1999;

@ZCDTXFJI#Gutierrez\_Etal\_2011]. Furthermore, as we mentioned in previous

sections DA activity in VTA is related to environmental uncertainty, which could

increase lateral hypothalamus activity through dopaminergic inputs to the

supramammilary nucleus [@4SMCFQII#Plaisier\_Hume\_Menzies\_2020], and indirect

action by orexin derived increased firing in VTA that further inhibit nucleus

accumbens shell. This could result in a net increase in food-seeking behavior

via uncertainty augmented activity in the VTA.

Orexin signal depolarizes LC increasing its firing [@A5ZYWRZW#Hagan\_Etal\_1999],

which can promote task disengagement [@KA4Y29AG#Kane\_Etal\_2017], alter network

representation of tasks [@GFA24ABK#Grella\_Etal\_2019], and update world models

[@TA5KB3TF#Sales\_Etal\_2019], likely via promotion of exploratory behavior

related to tonic firing [@LEWESIS6#AstonJones\_Cohen\_2005]. These data support

the integration of orexin with unexpected uncertainty and exploration. Orexin

also, synapse on cholinergic neurons in the basal forebrain, promoting

acetylcholine release, which also creates excitatory influence on orexin

neurons, generating a positive feedback circuit [@A93676U3#Sakurai\_Etal\_2005].

Cholinergic basal forebrain activity encodes valence-free prediction error

[@D8EGNYCV#Sturgill\_Etal\_2020], so DA and Ach could generate similar

uncertainty-drive excitation of orexin neurons. However, Ach activity might be

related to reducing the effects of prediction error in learning by signaling a

smaller variance [@S3YZPJJI#Puigbò\_Etal\_2020; @4BJ2B6KB#Yu\_Dayan\_2005] similar

to expected uncertainty signaling.

In this section, we provided a plausible circuit where orexin activity acts as a

hub relating, prediction error, unexpected uncertainty, and expected uncertainty.

This puts orexin as a candidate neuropeptide for dealing with uncertainty-driven

food-seeking behavior as it can both integrate environmental and internal status

information, and promote locomotor activity to procure food. We derived this

function taking theoretical and empirical finding 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.