A computational account of needing and wanting

Bosulu, J[1,2]., Pezzulo, G.[3], Hétu, S.[1,2]

[1]Faculty of Arts and Sciences, Université de Montréal
[2]Centre interdisciplinaire de recherche sur le cerveau et l'apprentissage (CIRCA)
[3]National Research Council of Italy, Institute of Cognitive Sciences and Technologies (ISTC-CNR)

ABSTRACT

Our actions and choices sometimes seem driven by our needs, sometimes by our wants and sometimes by both. Why sometimes there is association and sometimes dissociation between needing and wanting remains largely unknown. Needing is related to deprivation of something biologically significant, and wanting is linked to reward prediction and dopamine and usually has more power on behavioral activation than need states alone. This paper aims to clarify their relations using active inference. In this approach needing is related to a deviation from preferred states that living things tend to occupy in order to reduce their surprise, while wanting, i.e. reward prediction, is related to precision over policy leading to rewards. Through a series of simulations, we demonstrate the interplay between needing and wanting systems. Specifically, our simulations show that when need states increase, the tendency to occupy preferred states is enhanced independently of wanting (or reward prediction), showing a dissociation between needing and wanting. Furthermore, the simulations show that when need states increase, the value of cues that signal reward achievement and the precision of the policies that lead to preferred states increase, suggesting that need can amplify the value of a reward and its wanting. Taken together, our model and simulations help clarifying the directional and underlying influence of need states separately from reward prediction, i.e. wanting, and at the same time show how this same underlying influence of need amplifies wanting, i.e. increases the precision of reward cues that lead to the preferred state.

FULL TEXT

1. Introduction

There has been a debate around the question of if people, for instance consumers, are driven by their needs or their wants (Campbell, 1998). “Needing” is related to a state of deprivation of something important for life or survival (Bouton, 2016), and increases arousal through interoceptive salience (Craig, 2003); and not responding to a need might lead to some adverse

consequences such as physiological or psychological suffering which go beyond mere frustration (MacGregor 1960; Baumeister & Leary, 1995). “Wanting” is more related to goal achievement and reward prediction and is more closely related to dopaminergic activity and motivation (Berridge, 2004). Both have some influence, albeit differently, on the motivational value of stimuli. Importantly, motivation can have a directional effect and/or activational effect. The directional effect is linked to choice (preference or action selection) and directs towards or away from stimuli, while the activational effect is related to action and its initiation, maintenance, and vigor or effort (Salamone et al., 2018). Notably, wanting is able influence both activational and directional values of stimuli even in absence of need states: nonhuman animals tend to respond and “want” food even when satiated (Bouton, 2016), and for humans, cues of M&M or pictures of cigarettes (for smokers) have been shown to lead to more consumption even after having been consumed to satiety (Hogarth & Chase, 2011; Watson et al., 2014). On the other hand, needing seems to control directional motivation, but seems to rely on wanting for the activational part ((Dickinson & Balleine 1994; Balleine, 1992; Berridge, 2004; Berridge, 2018; Wassum et al., 2011; Salamone et al., 2018).

At the brain level, needing and wanting systems could map to partially different neural substrates. A recent fMRI meta-analysis compared consistent brain activations during needing (perception of needed stimuli, i.e. food when hungry) and during wanting (perception of a reward cue that leads to reward seeking). It showed that needing seems more related to brain regions implicated in directional motivation/value; whereas wanting seems to be more related to brain regions implicated in both directional and activational motivation/value, and to mesolimbic dopaminergic areas (Bosulu et al., 2022). Furthermore, these results suggest that needing is related to interoceptive predictions and prediction errors as well as cues related to interoceptive states, such as food cues, possibly computed within the mid-posterior insula (Bosulu, et al., 2022; Livneh et al., 2020), while wanting is instead related to exteroceptive (e.g., cue- or reward-related) predictions and prediction errors, with the latter possibly computed within the ventral tegmental area (VTA) (Schultz et al., 1997; Bosulu, et al., 2022). Hence, needing seems to be more related to the internal environment, whereas wanting seems to be more related to the external environment, although this is not necessarily a strict separation as – for example – wanting could also take into account the internal environment (Berridge, 2004).

However, needing and wanting systems are not segregated but might influence each other reciprocally. Biological needs, such as those related to states of deprivation of a biologically significant stimuli or events (Bouton, 2016), seem to influence the rewarding aspect of biologically relevant stimuli, such as wanting, pleasure and choice. For instance, food is often more wanted, liked and chosen when hungry. Thus, although it has been shown that motivation to pursue rewards depends more on expectations (e.g., reward prediction), than on needs states (Bindra, 1974; Berridge, 2004), need states tend to amplify reward predicting cues that are

relevant to such need states (Toates, 1994; Berridge, 2004). Need states can also enhance liking of, and pleasure related to, relevant stimuli that satisfy one's needs (Cabanac, 2017; Berridge & Kringelbach, 2015; Becker et al., 2019). Moreover, a need state has the capacity to give and to control the preference/choice or value of a novel food or drink or of a particular choice or action, in relation to their consequence on the organism (Dickinson & Balleine 1994; Balleine, 1992), and such need related action happens via a system that could be both dependent (Berridge, 2004; Berridge, 2018) and independent of dopamine (Wassum et al., 2011 Salamone et al., 2018). Furthermore, in absence of a reward guiding cue, a need state can (directly) activate relevant actions (Passingham & Wise, 2012) or explorative behavior (Panksepp, 2004). Overall, tThis indicates that biological need states are able to influence certain tendencies towards relevant needed stimuli, hence affecting wanting and goal directed choices (and liking).

To sum up, our previous discussion illustrates that needing and wanting systems could link to different aspects of motivation (i.e., directional and activational) and to partially different neural substrates and states (i.e., internal and external) – but at the same time, they might interact in various ways. Yet, we still lack a comprehensive computational framework that accounts for the findings reported above and that specifically explains (1) how could internal needs exert a directional influence on goal-directed behavior and choice, even in absence of wanting and (2) how could a state of needing amplify wanting (and liking).

The main goal of this paper is to provide a computationally-guided perspective on needing and wanting, which helps make sense of the fragmented literature on these topics. In the following sections, we will firstly address the two above questions conceptually, using formal methods from active inference, information theory and reinforcement learning (Parr et al., 2022; Sutton and Barto, 2018). Then, we present two simulations that address the functioning of needing and wanting systems – and their interactions – more formally.

2. Needing and wanting systems and their interactions: a conceptual perspective

2.1 The directional influence of needs on goal-directed behavior and choice

The goal of organisms is to regulate internal states and keep them within certain very limited boundaries (Barrett, 2017; Sterling & Laughlin, 2015; Friston, 2006). For instance, the average normal body temperature for humans is generally between 36.1°C (97°F) to 37.2°C (99°F), which is a very small range compared to the range of possible temperatures in the universe, from the absolute zero to trillions of degrees. The same is true for levels of glucose or the balance between water and salt in the body. The main idea is that the number of "states" that make life possible are really small compared to the very large number of other combinations that wouldn't sustain life. So, to allow a living organism to remain within its normal physiological boundaries, natural evolution might have set them as so-called (empirical) priors, which might be conceived as (possibly genetically encoded) innate preferred states that the organism always strives to achieve. These preferred states (corresponding to physiological bound) have a greater probability

to be reached from the point of view of the organism, i.e. are less surprising (Friston, 2010). Here, the surprise associated with a state, denoted h(y), is the inverse of being probable and simply means less probable. Anecdotally, for a fish, being out of water would count as a surprising state. Any self-organizing system must minimize such surprise in order to resist a natural tendency to disorder (Friston, et al., 2006; Friston, 2010) and in the case of our fish, death.

Formally, the notion of surprise is closely related to the notion of entropy. Entropy, denoted as H, is the long-term average of the surprise and (here) it expresses the uncertainty related to which state must be occupied. If an organism is endowed with a prior about the states to occupy (i.e., with one or more preferred states), achieving these high probability states reduces surprise and its long-term average: entropy (Parr et al., 2022; Friston, 2010). By the same token, higher probability translates into more preference: living creatures naturally prefer the most probable states that sustain life. Thus, distancing from those states leads (or equivalently, being in a surprising state) entails an (informational) cost that living beings strive to minimize.

Importantly, the notion of being in a surprising state (or in other words, being far from preferred states) links well to the concept of "needing" discussed in the Introduction. In the same way being in a surprising state entails an (informational) cost, a state of need entails a (biological) cost if a person does not respond to the need (see MacGregor 1960; Baumeister & Leary, 1995). When a living organism moves away from its preferred state, it is in a state of "need" - which amounts to having a tendency to occupy preferred states (again). The state of need can be represented as:

$h_n(y) = -\ lnP(y|C)$ Eq. 1;

where h_n represents the "need-related" surprise of a sensation or state *y*, which is equal to the negative log probability of being in (or observing) a state *y* given the distribution of prior preferences, denoted as *C*. Note that for simplicity, in this article we will collapse the notions of "state" and of "observation that can be obtained in the state", which are typically distinct in active inference (and more broadly, in Partially Observable Markov Decision Processes); see the simulations below.

The perception of a need state translates into a "goal" of reducing surprise by reaching the preferred states, e.g., states that represent adaptive physiological conditions (Friston, 2010). Such tendency could activate an action or a policy (i.e., an action pattern or sequence of actions) that compel creatures to seek out the (valuable) preferred states. Note that the actions or policies that resolve a state of need could in some cases correspond to (fixed) regulatory actions, such as autonomic reflexes, as opposed to action courses determined by the circumstances of the external environment (Sajid et al., 2021). The states that the creature occupies when pursuing a policy that resolves a need can become valued per se (Friston & Ao, 2012). In other words, when the

creature pursues a course of actions towards the preferred state, all the intermediate states (here intended in a broad sense that encompasses situations, actions, stimuli, etc.) can become valued and needed. For instance, when moving from a state of hunger to a state of satiety, some intermediary states could be the gustatory stimulus associated to having food and the act of eating; and such states would become valued, because they are in the path towards the preferred (satiety) state (Pezzulo et al. 2015). Through this mechanism, a creature would experience the need for food or the need to eat – and then start to prefer the valued states, stimuli or actions. In other words, the directional effect of need states on motivation could come from the tendency to occupy preferred states.

In turn, as noted above, pursuing preferred states reduces "need related entropy" and the surprise associated with occupying non-preferred states. In this sense, the tendency to occupy preferred states confers to need states the possibility to influence – and give value to – stimuli or actions that are either costly states $S(h_n)$ that lead to surprise, or in the path towards the preferred state $\pi(p)$. In other words, in an environment where there are increasingly costly/surprising states, any state (stimulus or action) that is in the path to the preferred state will become valued (needed) because it reduces entropy.

2.2 How needing amplifies wanting

The effect of needing on wanting (and on other phenomena such as pleasure and liking) could be conceptualized by appealing to the formal notion of precision in active inference. Mathematically, precision is a term used to express the inverse of the variance of a distribution which in our context can be seen (loosely speaking) as the inverse of entropy (Friston, 2010; Holmes 2022) – in the sense that the higher the entropy, the lower the precision. In predictive coding and active inference, precision acts as a multiplicative weight on prediction errors: prediction errors that are considered more precise have a greater impact on neural computations (Parr et al., 2022).

In active inference, there are different precisions associated with different forms of changes in prediction, such as interoceptive, reward or policy predictions (see Parr et al., 2022). Of particular relevance here is the precisions of policies, which indexes the confidence that we have that by following a particular policy we will reach a preferred (goal or reward) state. Crucially, while we pursue a policy, every cue or information that confirms (reduces uncertainty about the fact) that a policy will achieve a preferred state enhances policy precision. This is the case for example if we are following a route that we hope will lead to a given city (or to a restaurant that we hope is open) and we encounter a traffic sign that indicates that the direction to the city is correct (or a sign that the restaurant is open).

At the neurophysiological level, policy precision, or the confidence that a policy will lead to reward, is typically associated with the dopaminergic system in active inference (FitzGerald et

al., 2015; Parr et al. 2022, Holmes, 2022). Therefore, reward cues that enhance our policy precision and our confidence that the policy will lead to reward would trigger dopamine bursts, which will attribute incentive salience to such cues (Berridge, 2007). This is in line with the idea that dopamine is linked with incentive salience and wanting; but also with reward prediction and behavioral activation as they typically co-occur (Hamid et al., 2016). Rather, precisions regarding hedonic contact with the reward (to ask questions such as: is it good?) or the state of satiety (to ask questions such as: am I well?) might be mediated by the opioid system (Berridge & Kringelbach, 2015) and the serotonin system (Parr et al., 2022; Liu et al., 2020; Luo et al., 2016), respectively.

Interestingly, these systems are interdependent. When one is in a surprising (need) state, the presence of a cue (e.g., a traffic or restaurant sign) might reduce our uncertainty about goal/reward achievement by improving policy precision via dopamine system activity (wanting); the presence of, or contact with, the reward itself might reduce entropy by enhancing precision through the opioid system (pleasure/liking); while being directly in a preferred state or towards the preferred state could be related to precision via serotonin (well-being). All of these contexts serve as information that reduce entropy by signaling the availability of a path to preferred states ($\pi(p)$), or equivalently away from surprising states (h_n), given some prior preference.

The policies discussed so far depend on exteroceptive cues from the environment. As discussed above, these cues (e.g., cues that signal a reward) can become imbued with incentive salience and "wanting" (Berridge, 2007) in virtue of the fact that they enhance policy precision. The increased policy precision that comes from the presence of rewards (or of cues that signal incoming rewards) means more certainty that the state to which the policy leads will be rewarding, and it is this certainty that amplifies wanting. This mechanism could function relatively independent from a state of need. For instance a relatively satiated animal that smells some food, will be motivated to follow the smell because doing this increases the probability of finding food. However, its wanting (and the associated dopamine firing) could be amplified if the animal is in a need state. If the animal that smells food is hungry, it will be motivated to follow the smell, not just because there is a high probability to secure food, but also because that food will be very rewarding. That is, the animal will be more certain (i.e., have greater precision) that it should pursue a policy that leads to the predicted reward. It's in that sense that needing amplifies wanting. It is also possible to speculate that the hungry animal will likely have more pleasure while eating food than the satiated one – because it could assign more certainty (precision) to how good food is when it is hungry.

To summarize, we propose that the wanting mechanism is intrinsically related to the fact that cues afford some resolution of uncertainty (e.g., about what policy to pursue) and linked to the dopaminergic system. Indeed, “wanting” depends on external stimuli that act as pavlovian cues that predict rewards (Berridge, 2018), and the attribution of value to these cues depends on mesolimbic dopamine reactivity which can be enhanced by physiological states (needs,

emotions, drugs, etc.) (Berridge, 1996; Berridge, 2004). Furthermore, we propose that the influence of needing on wanting can be conceptualized as a need-induced enhancement of the precision of policies and thus of the saliency of the rewards, stimuli or actions that lead to the preferred state.

2.3 Summary

Our discussion so far has highlighted two important points about needing and wanting systems, and their interactions. First, need states exert a directional influence on choices separately from wanting (and reward prediction). In the perspective offered here, need states gain their underlying (motivational and saliency) effect from the tendency to occupy preferred states. When a living system moves away from such preferred states towards "costly" states, a state of need is generated. Being in costly or need states automatically increases the probability of activating policies that lead to preferred states. Hence, the mere state of deprivation of need has an underlying directional effect and influences the tendency to reach rewarding or goal states. This tendency exists irrespective of reward prediction, as conceptualized for example in model-free reinforcement learning systems, which can be related to wanting (Zhang et al., 2009; Berridge, 2004).

Second, by the same token, the underlying influence of need amplifies wanting, by increasing the value of reward cues that signal the possibility to reach goal states (hence lowering uncertainty about goal achievement) and the precision of goal-achieving policies. The simultaneous circumstance of being in a need (hence surprising) state and encountering a cue that signals that a goal state is achievable (hence that surprise and uncertainty can be reduced) is the ideal condition to confer goal-achieving policy with a very high precision. By indicating that there is a path to a goal or reward state, the cue renders the organism more confident about what to do - and this is amplified when the gain in reward (or the amount of surprise minimized) is greater, such as when one is in a need state. This is why need states amplify "wanting" by enhancing the value of cues and the precision of policies that lead to preferred goal states.

In the next sections, we move from the conceptual treatment offered in this section to a formal implementation of the proposed model of needing and wanting. Subsequently we illustrate the functioning of the model in two simulations, which illustrate how being in need/costly states influence the tendency to reach rewarding/preferred state (Simulation 1), and how the simultaneous presence of a state of need and the presence of a path to the preferred (reward or goal) state implies low entropy and high precision over which state to occupy (Simulation 2).

3. Methods and Results

3.1 Simulation environment

Our simulations regard a very simple situation, mimicking the idea of agents that have to remain within a limited physiological (e.g., food or temperature) bound. For this, we designed a grid-world environment in which only one state is rewarding/preferred, whereas the number of surprising/costly states is much greater (See Figure 1). We can draw a parallel between the grid-world and human physiological states, such as hunger or temperature: the preferred state corresponds to the optimal interval of sugar level in the bloodstream, or the temperature range (between 36.1°C (97°F) and 37.2°C (99°F)). In the grid-world, each interval of sugar level, or of temperature range, is represented as a specific state in which the agent can be (each state corresponds to a box in Figure 1).

Specifically, we used a 3x3 grid-world containing nine states, eight "costly states" (states 0, 1, 3, 5, 6, 7, 8) and a reward/preferred state (state 2), whose value is initially unknown to the agent. There are two terminal states (not shown) reachable from states 2 and 5, i.e., states that, once occupied, cannot be left by the agent. State 2 represents a reward/preferred state that gives a reward of 1 and state 5 leads to death (which is as costly as the other states, but from which the agent can never come back).

state 0	***state 1***	***state 2*** ***<u>reward</u>***
state 3	***state 4***	***state 5*** ***<u>death</u>***
state 6	***state 7***	***state 8***

Figure 1. Grid world environment used in our simulations. Each box represents one state in which the agent can be. These include eight costly states (states 0, 1, 3, 4, 5, 6, 7, 8*) and a reward state (state 2). The value of these states is initially unknown.*

In our simulations below, the only thing that we will vary is the amount of cost associated with the eight "costly states". We will do so by assigning to these eight states the same negative

reward (-1, -2, or -5) representing a significant departure from biological bounds. Note that the agents that dwell in the simulated environments do not know these values and will have to compute the *expected* biological costs, or values by themselves (see below for a description of the agents that we will use in our simulations).

3.2 Simulation agents

Our simulations will consider two agents: an active inference agent that embodies our hypotheses about needing and wanting systems and that implements tendencies or prior preferences over policies; and a reinforcement learning agent that computes reward prediction in the form of action values (Parr et al., 2022; Sutton and Barto, 2018). Note that at each time step, the two agents receive an observation about their current state and then they can self-localize in the grid map; and they can move one step vertically or horizontally, but not diagonally.

Agent 1: active inference agent

Agent 1 is a simplified version of active inference, in which the perceptual part is kept as simple as possible, by assuming that all the states of the grid world are observable. Technically, this means that we are dealing with a Markov Decision Process (MDP) and not a Partially Observable Markov Decision Process (POMDP) as more commonly done in active inference (see Friston et al., 2009; Friston et al., 2017). This simplifying assumption is motivated by the fact that our focus in this work is on action selection and not perceptual discrimination. Furthermore, keeping the perceptual part simple facilitates the comparison with the reinforcement learning agent (see Agent 2 below) that uses an MDP, too.

The costs and rewards assigned to states translate directly into a prior preference for these states (denoted below as C), with rewarding states being a-priori more probable than other states. Since the agent expects to occupy (or to move towards) these a-priori probable states, the prior over states also translates into priors over actions or action sequences (policies) that achieve such states. In this simplified setting, action (and policy) selection simply corresponds to inferring a distribution of states that it prefers to occupy and policies to reach (sequences of) these states. In other words, the active inference agent tend to select policies that lead it to achieve goal states - which in Bayesian terms corresponds to maximizing model evidence.

More formally, under the simplifying assumptions discussed above, the active inference agent strives to maximize a measure of (log) evidence, defined as:

$ln\,[P(y)]$ Eq. 2

where ln denotes a natural logarithm and $(P(y)$ the probability of observing a state (remind that in our setting, hidden states and observations are the same; hence $y = s$.) Equivalently, active inference agents strive to minimize surprise, defined as the negative of the evidence:

$-ln\,[P(y)]$ Eq. 3

Importantly, as remarked above, for active inference agents what is surprising or not surprising depends on prior preferences (e.g., a fish out of water is in a "surprising" state). In this perspective, being in a surprising state (i.e., far from prior preferences) is what defines a state of "need". To account for needs, we condition the surprise to priors C (prior preferences, i.e. biological costs or rewards) and we get a need-related surprise:

$-ln\,[P(y|C)]$ Eq. 4

This equation represents the surprise or negative "value" of an observation/state, given the organism's prior preferences. In turn, the prior preferences play a role in prioritizing policies, which correspond to courses of actions that try to reduce current and future (expected) surprise. When one accounts for policies (π), the expected surprise given the prior preferences is:

$-\,E_{Q(y|\pi)} ln\,[P(y|C)]$ Eq. 5

The $E_{Q(y|\pi)}$ part means that the probability of states/outcomes is averaged across all policies (π).

In active inference, the quantity shown in Eq. 5 $E_{Q(y|\pi)} ln\,[P(y|C)]$ (without the minus sign) is typically called a "pragmatic value" and in this setting, it corresponds to the expected free energy G(π) (an upper bound on expected surprise):

$G(\pi) = \;-\,E_{Q(y|\pi)} ln\,[P(y|C)]$ Eq. 6

For completeness, it is important to consider that the quantity shown in Eq. 6 (without the minus sign –the "pragmatic value"–is only one of the two terms of the expected free energy G(π) of active inference ; however, in our setting, the second term ("epistemic value") is zero, and hence here we simply ignore it.

The expected free energy G(π) is particularly important since it is used for policy selection. Specifically, active inference agents are equipped with a prior over policies, denoted as $P(\pi)$. The greater the expected free energy that policies are expected to minimize in the future, the greater their prior, i.e.,

$P(\pi) = \sigma(-G)$ Eq. 7

where σ represents the softmax function, bounded between 0 and 1, and enforces normalization (i.e. ensures that the probability over policies sums to one).

This $P(\pi)$ is what will be used as prior over policies by the agent 1 in order to select policies (which, in our simple scenario, reduce simply to actions) and ultimately to minimize surprise.

Agent 2: model-free reinforcement learning

The second agent makes decisions based on prediction of rewards assessed by state-action values, i.e. each decision will depend on the value of actions given the current states (see Sutton & Barto, 2018). Here the policies depend on the action values, denoted $Q^{\pi}(s,a)$, and given by:

$$Q^{\pi}(s,a) = E_{\pi}\{R_t | s_t = s,\ a_t = a\} = E_{\pi}\{\sum_{t=0} y^i r_{t+i} | s_t = s,\ a_t = a\} \quad \text{Eq. 8}$$

The equation shows the value or "quality" (Q) of the action (a) in state (s) under a policy (π). The function denoted $Q^{\pi}(s,a)$ expresses the expected (E) return (R), which is the (expected) sum of rewards, starting from state (s) and taking the action (a), and thereafter following policy (π). Here the state *s* for agent 2 is equivalent to the state/observation *y* of agent 1.

The agent's decision after learning is based on the optimal policy π_*, i.e. the one that maximizes the expected return, and therefore the optimal $Q^{\pi}(s,a)$, noted $Q^{*}(s,a)$ is equal to:

$max_a Q(s,a)$ Eq. 9,

where max_a is related to the action that maximizes $Q(s,a)$.

In sum, the two agents differ in their policy selection mechanism. Agent 1 (active inference) uses a prior over policies $P(\pi)$ illustrated in Eq 7, whereas agent 2 (reinforcement learning) uses action values $Q^{*}(s,a)$ illustrated in Eq 9 which, here, amounts to reward prediction. The goal of our first simulation, illustrated below, is to assess the effects of increasing need states on the action selection mechanisms of the two agents.

3.3 Simulation 1: Directional aspect of needing separately from reward prediction

In this simulation, we consider the effects of increasing the costs of the eight “costly states” of Fig. 1 from -1, to -2 and -5 on action selection of active inference (left panels of Fig. 2) and reinforcement learning agents (right panels of Fig. 2), respectively. For this, we calculate for each level of need/cost (i.e. -1, -2, -5) the prior probabilities of policies $P(\pi)$ that reach each of the states of the grid-world (Agent 1) and the state-action values that emerge from reinforcement learning (Agent 2). To calculate the probabilities of policies of Agent 1, we use active inference, as explained above. Rather, to calculate the action values of Agent 2, we use a standard Q learning approach and let Agent 2 learn by trial and error, by navigating in the grid map (for 1000 trials).

The results illustrated in Fig. 2 show that increasing the costs of the eight “costly states” significantly increases the probability assigned to policies that reach the rewarding state in the active inference agent (Agent 1). This is evident when considering that the probability increases from (about) 0.5, 0.8 and 1 in the three left rows. However, increasing the costs of the eight “costly states” does not affect reward prediction in the reinforcement learning agent (Agent 2). This is evident when considering that the state-action (Q) values assigned by the reinforcement learning agent to the rewarding state is always 1 in the three right rows (this is because the true reward provided by the state is 1).

These results help illustrate the idea that costly or need states might exert directional effects and impact on the probability (or tendency) to reach preferred states, irrespective of reward prediction. In other words, expected biological costs do not control or constrain values directly, but directly control probabilities, i.e. policies or tendencies, to be in (or go to) the preferred states. This directional effect of needs is well captured by active inference agents, which increases the probability to reach rewarding states when “in need”. Conversely, a reinforcement learning approach to estimate the reward guaranteed by the preferred state correctly infers the reward itself but it is not sensitive to the magnitude of the "need" of the organism.

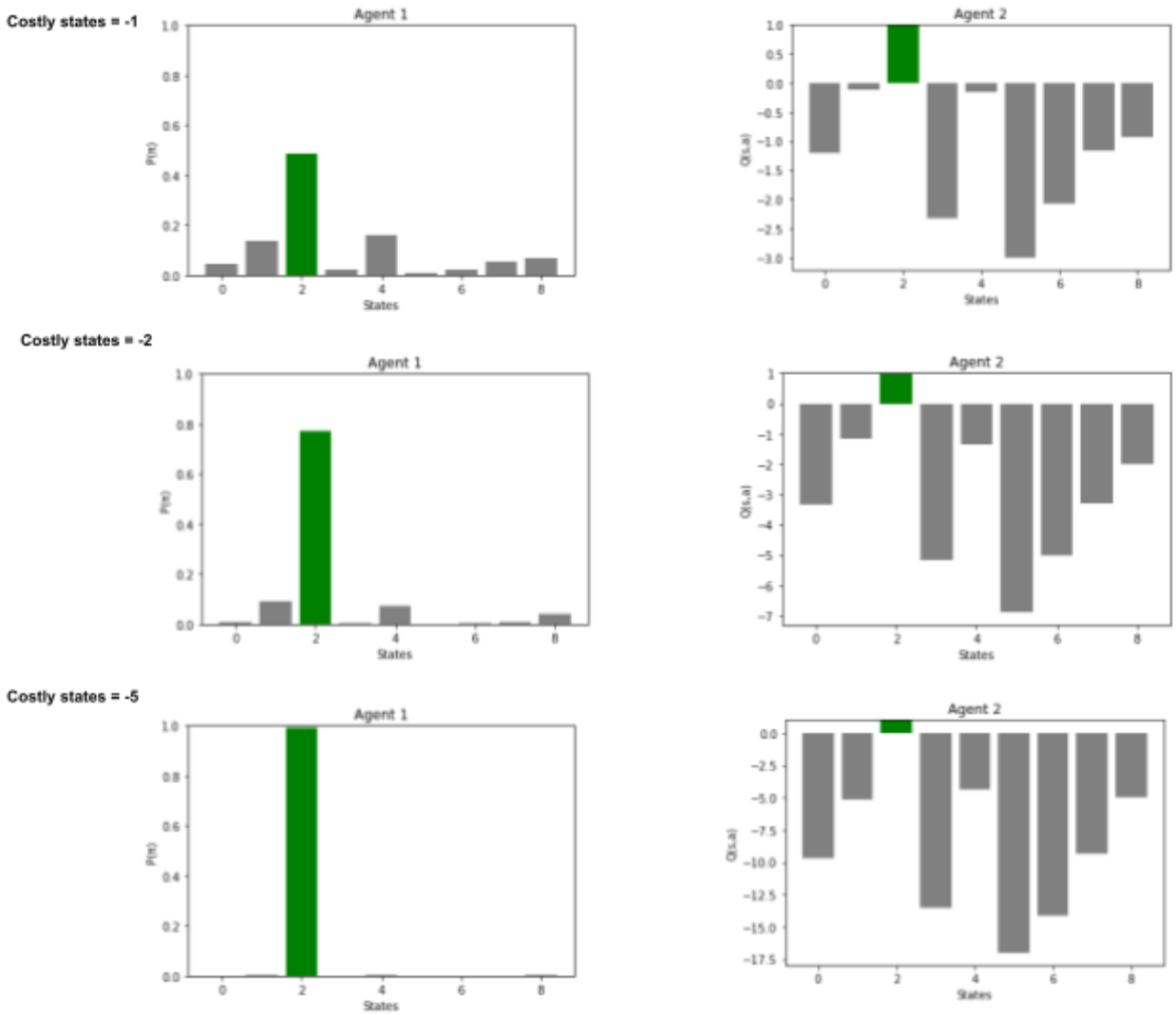

Figure 2. Effects of biological needs on policy selection under active inference (Agent 1, left panels) and reinforcement learning (Agent 2, right panels). Note that in the bottom panel of the RL agent, the value of state 2 is 1 as for its other two panels. The left and right panels show the results for active inference (Agent 1) and reinforcement learning (Agent 2), respectively. For active inference agents, the y axis plots priors over policies $P(\pi)$ *to reach each of the states of the grid-world, whereas for reinforcement learning agents, the y axis plots state-action (Q) values. The three rows show the effects of setting the costly states (states 0 to 8 except state 2, see Fig. 1) to -1, -2 and -5, respectively. The results show that increasing biological needs (across the three rows) increases the probability that Agent 1 selects policies to reach the preferred state 2, but does not increase per se the state-action value assigned by Agent 2 to state 2. This is consistent with the idea that need (directionally) influence tendencies (i.e. probabilities) more than reward prediction. See the main text for explanation.*

3.4 Simulation 2: How needing amplifies wanting

Simulation 1 showed that action selection in the active inference agent (Agent 1) is sensitive to need states. In Simulation 2, we ask if being in a state of greater need amplifies the wanting of the active inference agent, when a reward is present (note that here we do not consider Agent 2, as the results of Simulation 1 show that it was not sensitive to the magnitude of the "need").

For this, we consider an active inference agent dwelling in an environment in which the costs associated with costly states vary from -1, -2 and -5, which correspond to the fact that the agent faces milder or more severe conditions of "need". We consider (1) the entropy over the states that it plans to occupy in the future by following its inferred policy and (2) the inverse of the above entropy, i.e., the precision, which is a measure of certainty about which states to occupy in the future. We compare two conditions; namely, when there is a reward (i.e. the reward state 2 is baited with a reward of 1) and when there is no reward (i.e., the reward state 2 has the same cost as all the other costly states).

We calculate the “need related entropy” (or simply entropy) as follows:

$H_n(Y = S(h_n))$ Eq. 10,

when there is no reward, and

$H_{n,p}(Y = S(h_n), \pi(p))$ Eq. 11,

when the reward is present and the agent has a potential path towards the preferred state. Here, H denotes the entropy and it can be calculated on two sets of states. When the reward is available, the entropy is over the states occupied by the agent while following a policy $\pi(p)$ that leads to the preferred rewarding state (p). Alternatively, when there is no reward, the entropy is over the states $S(h_n)$, or the states S that lead to surprise h_n given the prior preferences (i.e., the need related surprise of Eq. 1).

$Y = S(h_n)$ means Y is $S(h_n)$, and $Y = S(h_n), \pi(p)$ simply means Y can be $S(h_n)$ or $\pi(p)$

The $\pi(p)$ and $S(h_n)$ represent states in different subsets of prior preferences, with the $\pi(p)$ representing states that are on the path to the preferred state. These can be viewed as rewarding (or cues) states or events that lead (transition) to the preferred state if one follows a policy leading to the preferred state. The $S(h_n)$ represent the states that lead to surprise. This formulation highlights that the (need-related) entropy of an agent that faces costly/surprising

states $S(h_n)$ is reduced when there is a path towards the preferred state $\pi(p)$. Thus, the inequality below holds:

$H_n \geq H_{n,p}$ Eq. 12

We calculate the precision as the inverse of the entropy:

$P_n = H_n^{-1}$ Eq. 13,

when there is no reward, and

$P_{n,p} = H_{n,p}^{-1}$ Eq. 14,

when there is a reward and hence a path to the preferred state / rewarded state. Given the inequality in Eq. 12, when states become more costly, the precision increases, providing that there is a path towards the preferred / rewarded state, which implies that:

$P_n \leq P_{n,p}$ Eq. 15

Given that we are discussing the motivational, i.e. active part, here entropy means (average) uncertainty over which state to occupy rather than uncertainty over what state is. Similarly, precision means certainty over what state to occupy. The principle is the same whether applied to what states to occupy or what policy to follow. The idea is to make it general so it can apply to incentive salience (wanting) or to hedonic sensation (liking), and also to simpler organisms that might not have a sophisticated brain.

The results of the simulations of entropy (Eq. 10 and 11) and precision (Eq. 13 and 14) can be appreciated graphically in Fig 3. These results shown indicate that compared to the case with no reward, the condition where a reward is present implies a significant decrease of the entropy over which states the active inference plans to occupy in the path to the reward (Fig. 3, left) and a significant increase of its precision, which is a measure of certainty about which states to occupy in the future (Fig. 3, right). This is because the availability of a reward makes the agent more confident about the states to occupy and the policy to select, whereas in the absence of a reward, all states are equally costly and the agent has no strong preference about which states to occupy (i.e., high entropy and low precision). The presence of a reward (which in this simulation is known by the agent) is a cue that makes it possible to pursue a preferred course of action, reducing the entropy about the states to occupy and increasing the certainty (precision) about the states to visit in the path towards the preferred state (and the precision of the relevant policy, not shown).

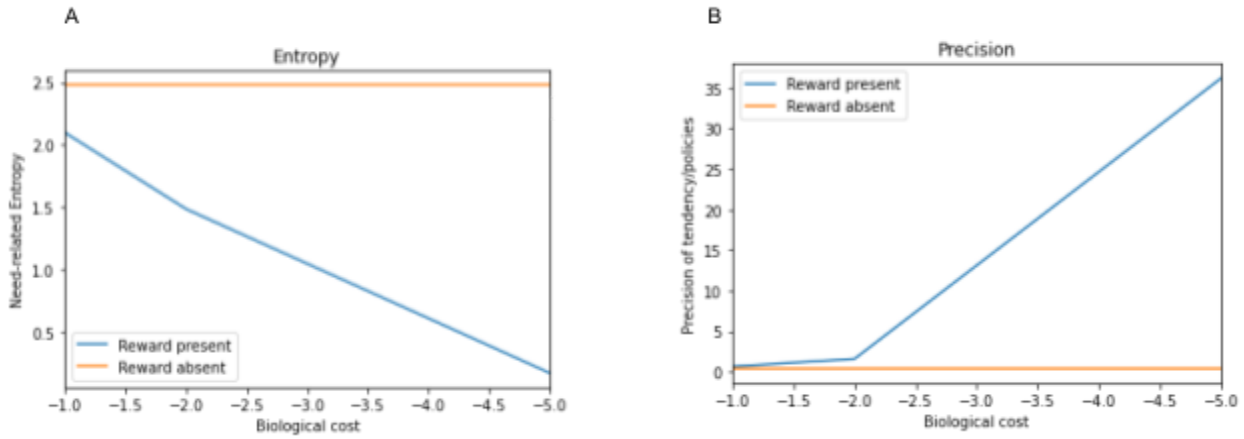

Figure 3. The impact of different need states on the entropy (left plot) and on its inverse, the precision (right plot), over which state to occupy for an active inference agent, in conditions in which a reward is available (blue lines) or no reward is available (orange lines).

Furthermore, and interestingly, being in a more costly state amplifies the above effects: with greater biological needs and a reward available, entropy reduces and confidence increases. Mathematically, the greater the biological cost of costly (or surprising) states, the lower the probability assigned to them, the higher the probability assigned to the preferred state - and ultimately, the lower the entropy over which state to occupy (as all the probability mass will be on the preferred state). In other words, if the biological cost associated with costly states increases, the agent becomes more confident about where to go and tends to occupy the preferred state with higher probability. In this sense, it is the conjoint presence of costly (surprising, need) states and of preferred (rewarding) states that maximally reduces the agent's entropy over which state to occupy (because the probability of reaching the preferred state increases) and increases its confidence about the path to pursue. To the extent that we associate the above precision (and confidence) with "wanting", these results show that a state of need amplifies the wanting, when there is a reward (or a cue) available.

4. Discussion

"Needing" and "wanting" exert significant influence on our decisions and actions. The former is related to biological costs and the deprivation of something biologically significant, while the latter is more related to reward prediction and dopamine and can usually exert a stronger influence on behavior. However, the respective roles of "needing" and "wanting" systems and the ways they interact are not completely understood. Here, we aimed to provide a

computationally-guided analysis of the mechanisms of "needing" and "wanting" and their interactions, from the perspective of active inference theory.

We firstly defined a need state as a "surprising" state in the sense assigned to the term "surprise" by theories like predictive coding and active inference, in which living creatures strive to remain within tight physiological boundaries (i.e. preferred states) and are surprised outside them - like a fish out of water. This perspective suggests that being in a costly or need state may exert a directional effect on action selection and motivation, because creatures would have an automatic tendency to select policies that avoid surprises and lead to preferred states. Importantly, this automatic tendency would be present without any reward or reward cue, which is in keeping with evidence of driving influences of "needing" without "wanting".

Furthermore, we associated "wanting" to the precision of policies that achieve preferred (goal or reward) states, consistent with previous work that linked policy precision to incentive salience and dopaminergic activity (FitzGerald et al 2015, Friston et al., 2017). In this perspective, cues that signal that there is a path to secure a reward are particularly salient and related to the "wanting". Needing and wanting systems are however related, as a state of greater need can amplify wanting: the higher the initial state of need, the greater the wanting of cues related to reward and of reward itself (and possibly also the greater the pleasure of reward consumption).

4.1 Simulation 1. The need system and its directional effect on behavior and motivation

To illustrate these arguments, we performed two simulations in a grid-world with most states associated with costs (to mimic a state of need of the organism) and one state associated with a reward. Simulation 1 illustrates the possible functioning of "needing", by focusing on the action selection mechanisms of two agents: a simplified active inference agent considers the probability of policies to go to the preferred state and a reinforcement learning agent that considers action values and reward predictions (learned using standard Q learning). This simulation shows that when the active inference agent is in more severe states of needs (i.e., the non-rewarded states of the grid world are associated with a greater cost), it assigns a greater probability to the policies that lead to the rewarding state. Hence, in active inference, a state of need can have a directional influence on behavior, leading the agent towards the preferred states and away from costly or surprising states. Rather, the reinforcement learning agent correctly estimates the reward provided by the rewarding state, but this estimate was not sensitive to the agent's need state. This latter result helps illustrate the idea that the "need" system cannot be reduced to reward prediction per se.

This simulation therefore illustrates nicely the directional and underlying effect of need states: need controls directional motivation, because of the tendency of living beings to move towards preferred states; and such tendency activates policies that lead to the preferred states. In doing so, states within the trajectory of those policies become preferred (and valued). This tendency is

mediated by homeostasis or its more general form allostasis (Sterling, 2004; Barrett, 2017; Holmes, 2022; Demekas et al., 2020), which help animals remain within viable physiological boundaries (Sterling, 2004; Holmes, 2022). From the active inference perspective, a living organism continuously strives to reach or remain in its preferred states (which could be sometimes evolutionarily defined, though homeostatic or allostatic regulation, at the somatic, autonomic and neuroendocrine levels (Parr et al., 2022; Swanson, 2000)). These preferred states act as drives or goals that, through homeostasis and allostasis, direct action (Barrett, 2017); hence the directional effect of need states. Such directional influence, dependent on tendency to occupy preferred states, is also responsible for the underlying effect through which need states amplify wanting, pleasure, interoceptive prediction, choice, etc.; by enhancing precision of their related stimuli (or actions) that are in the path towards the preferred state, in an environment of costly/surprising states. This leads to Simulation 2 below.

4.2 Simulation 2: The effects of needing on wanting

Simulation 2 illustrates how the simultaneous presence of state of need and of a reward (and a path to it) can amplify "wanting" in the active inference agent. The simulation shows that the presence (versus the absence) of (a path to) a reward decreases the entropy of the states that the active inference agent plans to occupy and increases the associated precision, or the confidence about occupying these states (and about the policies, not shown in our results). Note that while in this Simulation 2 we focused on a fully observable environment, previous (companion) simulations performed in partially observable environments showed coherent results (Friston et al., 2017). These previous simulations showed that the observation of a contextual cue - namely, of a cue that reveals reward availability and location to the agent - increases the precision of policies and that these precision dynamics can be related to dopaminergic activity. In our Simulation 2, the mechanism is similar (despite the full observability) because the agent is aware that a reward is present and hence there is a path towards the preferred state that realizes the agent's prior preferences. In other words, in both cases, something that signals a viable path to the reward increases the confidence of the agent in its course of actions.

Importantly, our Simulation 2 also shows that the decrease in entropy over which state to occupy, and the increase of associated precision, are magnified when the active inference agent is in a more severe state of need (i.e., when the costs of the non-rewarded states of the grid world are increased) and there is a path to the preferred state. In other words, the more costly (surprising) these states are, the more the agent is certain that it needs to go to the preferred state. This illustrates how need states amplify the wanting (and perhaps also the liking) of stimuli: by reducing entropy and making the agent more confident about what course of action to select. Need states have cascading effects also on the stimuli and actions in the path towards goal or reward states. When in a severe need state, relevant stimuli, reward cues and actions have a greater role in reducing entropy and increasing the confidence in the selected course of actions

(Parr et al. 2022, Holmes, 2022). These relevant stimuli, reward cues and actions are therefore assigned a greater value and a greater “need-generated” salience, which neurophysiologically could correspond to increased dopaminergic activity, too.

4.3 Relations between our proposal and previous studies

Our proposal is coherent with previous reinforcement learning models of need’s influence on reward value or on incentive salience. A representative example is the "homeostatic reinforcement learning" model of Keramati and Gutkin (2014), according to which any behavioral policy, π, that maximizes the sum of discounted rewards also minimizes the sum of discounted deviations from a setpoint. In other words, the maximization of a needed reward also minimizes the deviation caused by such need (Keramati and Gutkin, 2014). This idea is related to our approach, in that the need state is related to a deviation from a setpoint; however, in keeping with active inference, the setpoint is modeled as a prior and each deviation from it implies a "surprise", as explained above.

Another related approach is the computational model of wanting and incentive salience by Zhang and colleagues (2009). This model introduces a physiological variable, k (representing dopaminergic state), which changes the value of reward, r, in turn resulting in an amplification of the value of states, V(s); specifically, the amplified value is denoted as $\overline{V}(s_t) = \overline{r}(r_t, k) + \gamma V(s_{t+1})$, with γ representing a discount factor, t denotes the time step and $V(s_{t+1})$ denotes the value of the next state. Hence, the value of a reward cue depends on the combination of the reward value r and the dopaminergic states k. If r can be a proxy of reward that assigns state values, and *k,* is influenced by the need states (e.g., hunger or thirst, etc) then $\overline{r}$, which is the function of the combination of r and k, reduces entropy. The state value $\overline{V}(s_t)$ can be enhanced by both the physiological need state k (surprising state) and the value of the reward r, that signals the path to the preferred state (*p*). The amplification of reward is a function of the reward itself and of k, that is: $\overline{r}(r_t, k)$. Here, the reward is (on) the path to the preferred state. Hence, replacing or interpreting the path to the preferred (rewarding) state as $\pi(p)$ and the physiological state k as the need state that lead to surprise $S(h_n)$, we obtain that the function of the enhanced reward $\overline{r}$, which is encoded as confidence by dopamine, is related to precision over policies leading to reward. This is the inverse of need related entropy, i.e. inverse of the function

$$H_{n,p}(Y = S(h_n),\ \pi(p)\ ),$$

or in other words the function

$$\bar{r}(r_t, k) = H_{n,p}^{-1}\ (Y = S(h_n),\ \pi(p)\ ) = H_{n,p}^{-1} = P_{n,p}$$

This implies that there is a direct correspondence between the model discussed here and Zhang's (2009) model. However, the treatment offered here allows a more general treatment as it expresses needing, wanting and their interactions in terms of information-theoretic quantities, namely, information and uncertainty.

More broadly, a number of studies show that needing and wanting can occur both together and separately (see Bosulu et al., 2022; Berridge, 2004). As our Simulation 1 illustrates, needing (in the sense of a tendency to occupy the preferred states) can occur without wanting (which is more related to reward achievement and prediction) and can influence a preference for courses of action that lead to preferred state, even when such influence is not directly linked to reward prediction (as the results of the reinforcement learning agent show). When needing and wanting for a stimulus (or action) happen together, need states tend to amplify wanting (Berridge, 2004 ; Toates, 1994) by increasing precision, or saliency, of wanted cues - as illustrated in Simulation 2.

Furthermore, although here we are mainly concerned with the precision of future states occupied by the agent, as discussed there are different precisions associated with different forms of changes in prediction (see Parr et al., 2022). So in principle, a state of need could increase different forms of precision, or salience, of stimuli or cues that are on the path to a reward and are typically associated with wanting and "incentive salience" (see Berridge 2004). Furthermore, a state of need could increase the (precision-mediated) liking and pleasure associated with stimuli that are available. For instance, for need states to amplify pleasure, the agent must come into contact with the reward (e.g. have food in the mouth when hungry).

Importantly, a precondition for the amplifying effect of need states on wanting is the presence of a cue that predicts that the course of actions will lead to a reward. Bosulu and colleagues (2022) conducted a fMRI meta-analysis on perception of needed stimuli in absence of wanting (i.e. significant stimuli were shown but the task did not explicitly state that they would be available after) and found that need states did not seem to sufficiently and consistently activate the dopaminergic system, which is related to wanting. This might be due to the fact that the stimuli associated with needing that were used in the studies included in the meta-analysis did not act as strong cues that the same stimuli would be obtained in the future. It is in this sense that needing can happen independently of wanting. However, the same study showed consistent activity within the mid-posterior insula during the perception of needed stimuli (Bosulu et al., 2022), which might be interpreted as an indication that needing can amplify the precision relative to interoceptive predictions, even in absence of a "true" reward prediction. The same study discusses how wanting, which is more related to reward prediction and can generally happen without needing (Bindra, 1974; Berridge, 2004), can be amplified by internal states other than the need states, such as emotions, stress, drugs, etc. (Berridge, 1996; Berridge, 2004). This likely

means that those non-need related internal states can amplify wanting-related precision. Furthermore, need states are perhaps not the only ones to generate entropy or influence the reactivity of the dopaminergic system. The activation of the dopaminergic system in relation to a range of phenomena, such as stress, emotions, etc. can amplify wanting even in absence of need states (Berridge, 1996; Berridge, 2004). The scope of all these interactions between needing, wanting, bodily and emotional processing and dopaminergic activity remains to be fully mapped in future studies.

Our treatment suggests that both states of need and reward cues (related to wanting) influence behavior, but differently. When the agent is in a state of need, there is an automatic tendency to resolve it and then the organism could tend to follow preferred policies or state transitions (which can be at least in part considered to be genetically encoded and related to homeostatic and allostatic mechanisms) irrespective of cues that signal reward availability. Rather, wanting could be more related to reward prediction and its associated policies could be triggered by the availability of reward cues. The above arguments lead to the suggestion that the needing systems could be more related to internal (prior) policies that are more automatic and wanting could be more related to learned and external policies, i.e. changes in the environment related to prediction of reward which can be attained by the behavior, likely related to the somatic motor system (Swanson, 2000). This might be one of the reasons why wanting has more control over behavioral activation (Bosulu et al., 2022; Salamone et al., 2018; Berridge, 2004), but at the same time need can modulate wanting (Berridge, 1996 ; 2004) just as it can modulate liking (Berridge, 2007), preference (Balleine, 1992), etc. - in a way that here we characterize formally as a need-related precision.

Yet, as our Simulation 2 showed, needing and wanting systems can act synergistically. For example, need states increased precision when there is a reward leading to the preferred state. In that sense, if the reward determined by the need states, i.e. by the prior (internal) policy, is the same as the reward predicted by a signal of reward which specifies the external policy, then the precision will be enhanced and needing and wanting will be synchronized.

In sum, this study aimed at providing a conceptual model that defines needing and wanting systems and their interactions; and to demonstrate some of the peculiarities of these systems with the aid of simulations based on the framework of active inference. However, the relations between the relatively abstract notions introduced in this article, such as need-related entropy, and their biological substrates, remain to be fully clarified and tested empirically. A more systematic mapping between the information-theoretical notions used here and neurobiological evidence is an open objective for future research.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGMENTS

The research was supported in part by NSERC Discovery Grant #RGPIN-2018-05698 and UdeM institutional funds.

AUTHOR CONTRIBUTION

Juvenal Bosulu: Designed the study, performed the experiment, data analysis, interpretation, and wrote the manuscript. **Giovanni Pezzulo**: revised the manuscript and provided critical feedback. **Sébastien Hétu**: revised the manuscript and provided critical feedback. All authors contributed to and approved the final manuscript version.

DATA AVAILABILITY STATEMENT

All Data is available upon request.

REFERENCES

Balleine, B. (1992). Instrumental performance following a shift in primary motivation depends on incentive learning. *Journal of Experimental Psychology: Animal Behavior Processes, 18(3),* 236.
https://doi.org/10.1037/0097-7403.18.3.236

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in neurosciences*, *29*(5), 272-279.
https://doi.org/10.1016/j.tins.2006.03.002

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.
https://doi.org/10.1093/scan/nsw154

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature reviews neuroscience, 16(7),* 419-429.
https://doi.org/10.1038/nrn3950

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Becker, S., Bräscher, A. K., Bannister, S., Bensafi, M., Calma-Birling, D., Chan, R. C., ... & Wang, Y. (2019). The role of hedonics in the Human Affectome. *Neuroscience & Biobehavioral Reviews*, *102*, 221-241.
https://doi.org/10.1016/j.neubiorev.2019.05.003

Berridge, K. C. (1996). Food reward: brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews, 20(1),* 1-25.
https://doi.org/10.1016/0149-7634(95)00033-B

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & behavior, 81(2),* 179-209.
https://doi.org/10.1016/j.physbeh.2004.02.004

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391-431.
https://doi.org/10.1007/s00213-006-0578-x

Berridge, K. C. (2018). Evolving concepts of emotion and motivation. Frontiers in Psychology, 9, 1647.
https://doi.org/10.3389/fpsyg.2018.01647

Berridge, K. C., & Aldridge, J. W. (2009). Decision utility, incentive salience, and cue-triggered "wanting". *Oxford series in social cognition and social neuroscience, 2009,* 509.

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron*, *86*(3), 646-664. https://doi.org/10.1016/j.neuron.2015.02.018

Bindra, D. (1974). A motivational view of learning, performance, and behavior modification. *Psychological review, 81(3),* 199.
https://doi.org/10.1037/h0036330

Bosulu, J., Allaire, M. A., Tremblay-Grénier, L., Luo, Y., Eickhoff, S., & Hétu, S. (2022). 'Wanting'versus' Needing'related value: an fMRI meta-analysis. *Brain and Behavior*. https://doi.org/10.1002/brb3.2713

Bouton, M. E. (2016). *Learning and behavior: A contemporary synthesis* (2nd ed.). Sinauer Associates.

Cabanac, M. (2017). Pleasure and joy, and their role in human life. In *Creating the productive workplace* (pp. 73-82). Routledge.

Campbell, C. (1998). Consumption and the Rhetorics of Need and Want. *Journal of design history, 11(3),* 235-246. https://doi.org/10.1093/jdh/11.3.235

Chen, J., Papies, E. K., & Barsalou, L. W. (2016). A core eating network and its modulations underlie diverse eating phenomena. *Brain and cognition, 110*, 20-42. https://doi.org/10.1016/j.bandc.2016.04.004

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current opinion in neurobiology, 13(4),* 500-505. https://doi.org/10.1016/S0959-4388(03)00090-4

Demekas, D., Parr, T., & Friston, K. J. (2020). An investigation of the free energy principle for emotion recognition. *Frontiers in Computational Neuroscience*, *14*, 30. https://doi.org/10.3389/fncom.2020.00030

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior, 22(1),* 1-18. https://doi.org/10.3758/BF03199951

Doya, K. (2000). Metalearning, neuromodulation, and emotion. In *Proc. Conf. Affect. Minds* (Vol. 46, p. 47).

FitzGerald, T. H., Dolan, R. J., & Friston, K. (2015). Dopamine, reward learning, and active inference. *Frontiers in computational neuroscience*, 136. https://doi.org/10.3389/fncom.2015.00136

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of physiology-Paris*, *100*(1-3), 70-87. https://doi.org/10.1016/j.jphysparis.2006.10.001

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, *11*(2), 127-138.
https://doi.org/10.1038/nrn2787

Friston, K., & Ao, P. (2012). Free energy, value, and attractors. *Computational and mathematical methods in medicine*, *2012*.
https://doi.org/10.1155/2012/937860

Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference?. *PloS one*, *4*(7), e6421.
https://doi.org/10.1371/journal.pone.0006421

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. *Neural computation*, *29*(1), 1-49.

https://doi.org/10.1162/NECO_a_00912

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4-26.
https://doi.org/10.1038/npp.2009.129

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., ... & Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nature neuroscience, 19(1),* 117-126.
https://doi.org/10.1038/nn.4173

Hogarth, L., & Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behavior Processes, 37(3),* 261.
https://doi.org/10.1037/a0022913

Holmes, J. (2022). Friston's free energy principle: new life for psychoanalysis?. *BJPsych Bulletin*, *46*(3), 164-168.
https://doi:10.1192/bjb.2021.6

Keramati, M., & Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife, 3.*
https://doi:10.7554/eLife.04811

Li, J., & Daw, N. D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. *Journal of Neuroscience*, *31*(14), 5504-5511.
https://doi.org/10.1523/JNEUROSCI.6316-10.2011

Liu, Z., Lin, R., & Luo, M. (2020). Reward contributions to serotonergic functions. *Annu. Rev. Neurosci*, *43*(1), 141-162.
https://doi.org/10.1146/annurev-neuro-093019-112252

Livneh, Y., Sugden, A. U., Madara, J. C., Essner, R. A., Flores, V. I., Sugden, L. A., ... & Andermann, M. L. (2020). Estimation of current and future physiological states in insular cortex. *Neuron, 105(6),* 1094-1111.
https://doi.org/10.1016/j.neuron.2019.12.027

Lex, A., & Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learning & Memory*, *15*(7), 483-491.
https://doi.org/10.1101/lm.978708

Luo, M., Li, Y., & Zhong, W. (2016). Do dorsal raphe 5-HT neurons encode "beneficialness"?. *Neurobiology of learning and memory*, *135*, 40-49.
https://doi.org/10.1016/j.nlm.2016.08.008

MacGregor, D. (1960). *The human side of enterprise* (Vol. 21, No. 166.1960). McGraw‐Hill: New York.

Mohebi, A., Pettibone, J. R., Hamid, A. A., Wong, J. M. T., Vinson, L. T., Patriarchi, T., ... & Berke, J. D. (2019). Dissociable dopamine dynamics for learning and motivation. *Nature*, *570*(7759), 65-70.
https://doi.org/10.1038/s41586-019-1235-y

Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press.

Parr, T., & Friston, K. J. (2018). The anatomy of inference: generative models and brain structure. *Frontiers in computational neuroscience*, 90.
https://doi.org/10.3389/fncom.2018.00090

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press.

Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight*. OUP Oxford.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in neurobiology, 134*, 17-35.
https://doi.org/10.1016/j.pneurobio.2015.09.001

Rice, M. E., Patel, J. C., & Cragg, S. J. (2011). Dopamine release in the basal ganglia. *Neuroscience*, *198*, 112-137.
https://doi.org/10.1016/j.neuroscience.2011.08.066

Salamone, J. D., Correa, M., Yang, J. H., Rotolo, R., & Presby, R. (2018). Dopamine, effort-based choice, and behavioral economics: basic and translational research. *Frontiers in behavioral neuroscience, 12,* 52.
https://doi.org/10.3389/fnbeh.2018.00052

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, *80*(1), 1-27.
https://doi.org/10.1152/jn.1998.80.1.1

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275(5306),* 1593-1599.
https://doi.org/10.1126/science.275.5306.1593

Schwartenbeck, P., FitzGerald, T. H., Mathys, C., Dolan, R., & Friston, K. (2015). The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cerebral cortex*, *25*(10), 3434-3445.
https://doi.org/10.1093/cercor/bhu159

Sterling, P. (2004). Principles of allostasis: Optimal design, predictive regulation, pathophysiology, and rational therapeutics. *Allostasis, homeostasis, and the costs of physiological adaptation*, *17*, 17-64.

Sterling, P., & Laughlin, S. (2015). *Principles of neural design*. MIT press.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Swanson, L. W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain research*, *886*(1-2), 113-164.

Toates, F. (1994). Comparing motivational systems—An incentive motivation perspective. In C. R. Legg & D. A. Booth (Eds.), *Appetite: Neural and behavioural bases* (pp. 305–327). Oxford University Press.
https://doi.org/10.1093/acprof:oso/9780198547877.003.0013

Wassum, K. M., Ostlund, S. B., Balleine, B. W., & Maidment, N. T. (2011). Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learning & memory*, *18*(7), 475-483.
https://doi.org/10.1101/lm.2229311

Watson, P., Wiers, R. W., Hommel, B., & De Wit, S. (2014). Working for food you don't desire. Cues interfere with goal-directed food-seeking. *Appetite, 79*, 139-148.
https://doi.org/10.1016/j.appet.2014.04.005

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., & Aldridge, J. W. (2009). A neural computational model of incentive salience. *PLoS computational biology, 5(7*), e1000437.
https://doi.org/10.1371/journal.pcbi.1000437