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 what we "need", sometimes by what we "want", and sometimes by both. 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. Why sometimes there is an association and sometimes a dissociation between needing and wanting remains largely unknown. This paper aims to clarify the relations between needing and wanting, from the perspective of a normative framework to understand brain and cognition; namely, 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 is related to the precision over policies that lead to preferred (e.g., rewarding) states. Also, we hypothesized that needing exerts a pervasive effect that is embodied on the organism. 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), therefore 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 needs can amplify the value of a reward and its wanting. Taken together, our model and simulations help clarify the directional and underlying influence of need states separately from wanting and reward prediction; and how the same underlying influence of need amplifies wanting, i.e. increases the precision of available reward cues that lead to the preferred state.

FULL TEXT

1. Introduction

There is an extensive debate around the question of what motivates people (e.g., consumers): what they "need" or what they "want" (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). 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”, instead, 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 to influence both activational and directional values of stimuli, even in absence of need states. For example, nonhuman animals tend to respond and “want” food even when satiated (Bouton, 2016), and for humans, cues of M&Ms 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 could rely on wanting for the activational part (Dickinson & Balleine 1994; Balleine, 1992; Berridge, 2004; Berridge, 2018; Wassum et al., 2011; Salamone et al., 2018) for behaviors related to the somatic motor system (Swanson, 2000). (Indeed need states could have activational impact on behavior related to autonomic and neuroendocrine levels (Swanson, 2000)).

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, such as the posterior insula; 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 to cues related to interoceptive states, such as food cues, possibly computed within the mid-posterior insula (Bosulu, et al., 2022; Livneh et al., 2020). Rather, wanting is 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. However, this is not necessarily a strict separation; for example, wanting could also take into account the internal environment (Berridge, 2004).

Interestingly, 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 one is hungry. Thus, although the 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). 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, even in the absence of a reward guiding cue, a need state can directly activate relevant actions (Passingham & Wise, 2012) or explorative behavior (Panksepp, 2004). This 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 could be related 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 the toe following points: (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 in an effort to help make sense of the fragmented literature on these topics. We thus present the bird's-eye view and the piecing together of the puzzle pieces about how needing and wanting can be both dissociated and associated. We also propose the idea of pervasiveness and of needs states and how its "embodiement", i.e. incorporaion, into prior preferences distinguishes needing from wanting and allows need states to have their directional effect, even in absence of reward predicting cues. In the following sections, we will firstly address the two above points/questions conceptually, using formal methods from active inference, information theory and reinforcement learning (Parr et al., 2022; Sutton and Barto, 2018) as well as the idea of pervasiveness and its embodiment. Then, we present two simulations that address the functioning of needing and wanting – and their interactions – more formally.

2. Needing system and wanting subsystem and their interactions: a conceptual perspective

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

A key goal of organisms is to regulate internal states and keep them within certain 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 is 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 (possibly genetically encoded) innate preferred states that the organism always strives to achieve. From the formal perspective of active inference, these preferred states (corresponding for example to physiological bounds) correspond to so-called (empirical) *priors* and have a greater probability to be reached from the point of view of the organism - or in other words, they are less surprising (Friston, 2010). Here, the surprise associated with a state, denoted h(y), is the "negative" of being probable and simply means less probable. Anecdotally, for a fish, being out of water would count as a surprising state. A key claim of active inference is that 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 priors about the (preferred) states to occupy, these states have a high prior probability to be occupied from the perspective of the organisms and achieving them reduces the organism's surprise and its long-term average: entropy (Parr et al., 2022; Friston, 2010). Hence, there is a negative correlation between the availability of the relevant reward (or a path to the preferred state) and the need induced entropy, because of the prior tendency to visit (more often) those states which results in lower entropy. In other words, a greater prior probability of states translates into a greater 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 (again) the preferred states. Formally, a 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.

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 a tendency could activate an action or a policy (i.e., a sequence of actions) that compels 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). With time, 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, through a pavlovian mechanism (Berridge, 2018). For instance, when moving from a state of hunger to a state of satiety, some intermediary states, such as the gustatory stimulus associated to having food and the act of eating, would become valued, because they are in the path towards the preferred (satiety) state (Pezzulo et al. 2015). In sum, a creature that is far from preferred states would experience a need (e.g., for food) – and then start to prefer the valued states (here, intended in a broad sense that encompasses stimuli, actions, etc.) that secure food or are part of the experience of having food. The directional effect of need states on motivation could come from the tendency to occupy these 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 needed (and hence valued) because it reduces need-related 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).

There are different precisions associated with different types of predictions and prediction errors within different subsystems, such as those related to interoceptive streams, rewards, policies, etc. (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 subsystem 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 subsystem (Berridge & Kringelbach, 2015) and the serotonin subsystem (Parr et al., 2022; Liu et al., 2020; Luo et al., 2016). (These are "natural" increases in precision due to prior preferences).

This discussion helps appreciate the deep interdependence between needing (which acts on the system as a whole) and wanting as well as other subsystems such as the hedonic and interoceptive omes. 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 the dopamine subsystem (wanting). The presence of, or contact with, the reward itself might reduce entropy by enhancing precision through the opioid subsystem (pleasure/liking). Finally, being in a preferred state or moving towards the preferred state determines an increase of precision - which is due to the fulfillment of prior preferences - via the serotonin subsystem (well-being).

In all these cases, rewards and cues are useful information sources: they reduce entropy by signaling the availability of a path to preferred states ($\pi(p)$), or equivalently a path away from surprising states (h_n), given some prior preference. Indeed, from the point of view of the organism in a need state, both encountering a relevant reward and a cue that leads to that reward would reduce uncertainty about how to move to the preferred state and alleviate the need. This is

despite making contact with a reward and encountering a cue that predicts reward might be treated in different subsystems: the liking and the wanting subsystems, respectively.

The policies discussed so far depend on exteroceptive cues from the environment that signal goals or rewards and that can therefore become imbued with incentive salience and wanting (Berridge, 2007), because they enhance policy precision. The increased policy precision that comes from the presence of rewards (or of cues that signal incoming rewards) implies 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, at the neurophysiological level, 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 it predicts that the 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.

2.3 Summary

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 are linked to the dopaminergic subsystem. 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 needing and wanting differ in many ways. First, wanting is a subsystem like others (e.g., liking), while needing has a more *pervasive* and systemic effect. Second, wanting is defined by two components: the dopaminergic reactivity and the pavlovian reward predicting cues (that are in turn related to learned reward values, see our simulation of agent 2 below) (Berridge, 2004). Pavlovian reward predicting cues are necessary for wanting (but not for needing) to occur and are integral to the definition of wanting in terms of incentive salience (Berridge, 2004). Yet the reward prediction itself (albeit necessary) is not sufficient for wanting, unless it is modulated by dopaminergic states/ reactivity to reward cues (hence increasing precision over policies). Note that the dopaminergic states that engage wanting need not to be need states; indeed, other brain states, such as those linked to emotions or drugs can facilitate mesolimbic dopamine

transmission - and in all these cases, reward cues can trigger wanting even in the absence of need states. Finally, 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.

Our discussion so far has highlighted two important points about the needing effect on the system and the wanting subsystem, 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 surprising (or 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 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 a 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 an implementation of the proposed model of needing and wanting. Subsequently we illustrate the functioning of the model in two simulations, which exemplify how being in need states influences the tendency to reach preferred states independently of reward prediction (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 3x3 grid-world with 9 states, in which only one state (state 2) is rewarding/preferred, one is a need state (state 8), i.e. it is (biologically) costly/surprising; and the other seven states are "neutral"; see Figure 1 and 2. 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.

In our simulations below, we will assume that the agents are in a need state (e.g., hunger) that can be low, mild or severe. Accordingly, a negative reward (of -1, -2, or -5) is assigned to the "costly" state (state 8). Note that the agents that dwell in the simulated environment sense this biological cost and will have to compute the *expected* biological costs, or values, by themselves, in ways that different from agent to agent (see below).

Finally, another feature of the grid-world is that there are two "terminal" states (not shown) reachable from states 2 and 5, i.e., states that, once occupied, cannot be left by the agents. State 2 represents a reward/preferred state that gives a reward of 1 and resolves the need state. Rather, state 5 leads to death. It has the same negative reward as the other "costly" states, but the agents cannot leave it. Of note, the gridworld is used here for visual presentation purposes and the proposed framework can indeed be generalized beyond a gridworld.

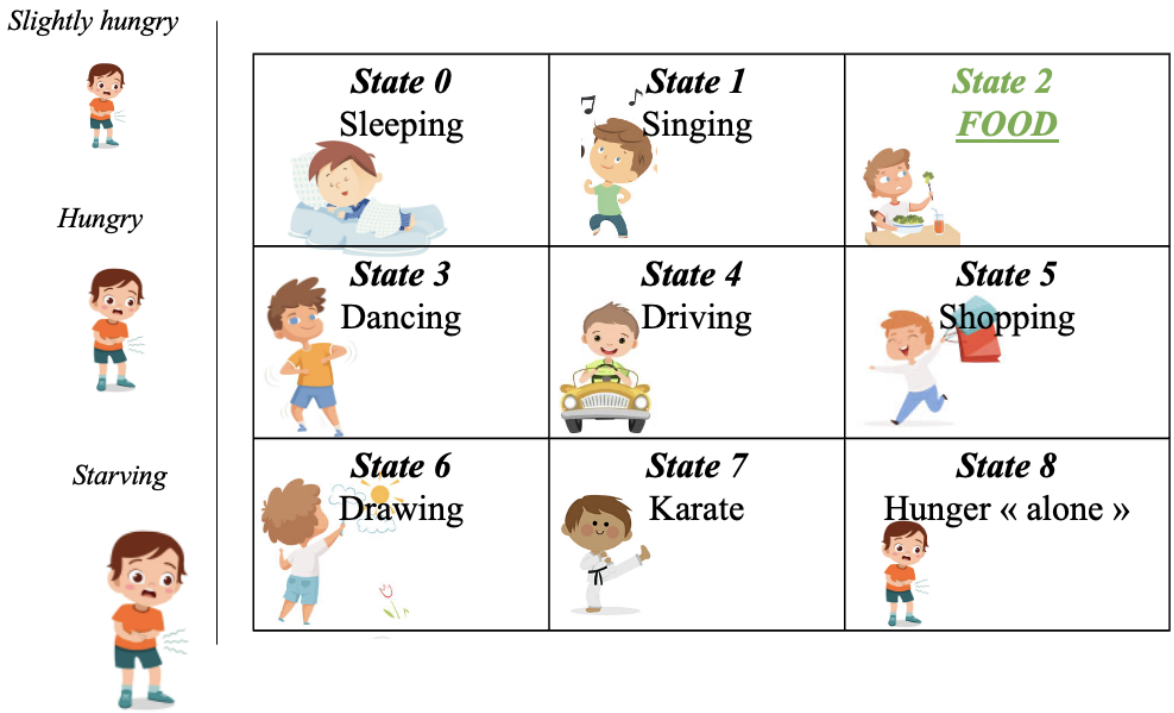

Figure 1. Grid world environment used in our simulations. Each box represents one state in which the agent can be. These include seven neutral states (states 0, 1, 3, 4, 5, 6, 7*), a reward state (state 2), and a costly state (state 8). The value of these states is initially unknown. The boxes represent different states and the images on the left represent some increase in hunger. Please note that we put hunger "alone" in order to distinguish between the presence and absence of the pervasive effect. In this case, the pervasive effect of hunger would attain all the other states, except the state 2: food.*

3.2 The pervasiveness of a need state

One might ask why we assign (the same) negative reward to all the "costly" states of the grid-world. One central point of our argument is that a need state exert a pervasive effect, i.e. it tends to affect every aspect of cognitive processing of the organism. Pervasiveness contributes to the distinction between needing and wanting. Not getting what we want may get us frustrated, but is not pervasive, whereas not getting something we need affects us beyond mere frustration, by impacting many aspects of our lives (see MacGregor 1960; Baumeister & Leary, 1995). In other words, need states do not disappear by simply changing states, contrary to aversive states (which are specific to a state). Practically speaking, if one is in a need state (e.g., a state of "hunger"), this need state propagates a negative valence to all the other states (e.g., the states of "sleep", "run", "play", etc.). The only exception to this is the state "having food in the stomach".

Formally, we represent the pervasiveness of the need state, or the fact that the need state propagates a negative valence to any other state i of the organism, in the following way:

$p(y_i, y_n)\ [Need]$ Eq. 2

where y_i is a any state, y_n is the need state and $p(y_i, y_n)$ is the joint occurrence (conjoint probability) of the ith state and the need state. So, when the organism is in a need state, all states eventually become paired with negative valence. The only exception is the rewarding state that alleviates that specific need, because the probability that it jointly occurs with the need state is zero. This is why we assign (the same) negative reward to all states, except the rewarding state 2.

We mean by "pervasiveness", the fact that a need state would potentially increase "as time goes" and impact "other states", i.e. other dimensions of life, unless one transitions to the (only) state that satisfies that need. *So, anything that is pervasive over time and over other dimensions of the individual's life, and whose negative impact is surprising with regard to prior preferences, could be defined as a need state.* (This is not strictly limited to physiological states, but includes social and other emotional aspects of life.) In this sense, depending on circumstances, some emotions and withdrawal effects of drugs could be classified as need states. For instance feeling lonely would likely have such need state (and pervasive effect) that would amplify the wanting to see loved ones. The same could be true when one is experiencing withdrawal effects, in which case the state "with drug in the body" has become a prior preference. In both cases, the pervasive emotion and the dependence on drugs usually generate need states. However, this is not the case for all "everyday" emotions (e.g. temporary joy or anger) and consumption (e.g. after having a drink) (see MacGregor 1960; Baumeister & Leary, 1995), which do not generate need states but which can still influence mesolimbic dopamine reactivity and facilitate wanting.

In order to model the pervasive effect of the need state, we assume a joint probability of 1 between the (increasing) need state and other states, except the preferred/satiety state which had a joint probability of 0 with the need state. (Please note that the behavior of the agent may have been different if the joint probability with the other states was significantly less than one.). Accordingly, to illustrate the pervasive effect, the negative reward (of -1, -2, or -5) of state 8 is propagated to the other states depending on their joint probability. For simplicity, we define a joint probability of 0 between state 8 and the rewarding state, and a joint probability of 1 between state 8 and the other states. So, Agent 1 ended up with 8 "costly" states, i.e. all states except the reward state, reflecting the fact that in all these states Agent 1 feels a pervasive effect of their need state.

Illustrating the pervasiveness effect

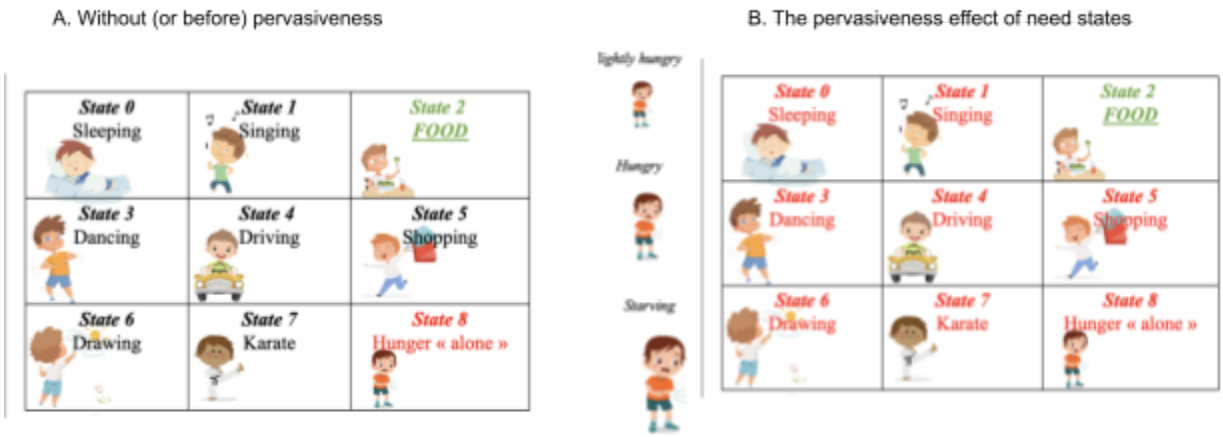

Figure 4. The impact of pervasiveness of a need state on other states. The expansion of the "red" illustrates the pervasiveness effect. In A there is no pervasiveness and hunger is more like an aversive state (state 8). In B pervasiveness occurs and hunger impacts all other states, except the (preferred) state that alleviates it. This increases preference for stimuli/events or actions that lead to the preferred state and by doing so, it increases precision (neuronal gain) or weight assigned to them within the subsystems that process them.

3.3 Simulation agents

Our simulations will consider two agents. Agent 1 incorporates the main ideas discussed in this article about needing and wanting systems. It is a simplified active inference model that bases action selection on prior preferences about states and policies (Parr et al., 2022).. Agent 2 is used as a comparison, to show that reward prediction alone is not sufficient to explain the interplay of needing and wanting. It is a simplified reinforcement learning system that based action selection on action values computed by reward prediction (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 state space, here the grid map; and they can move one step vertically or horizontally, but not diagonally.

Agent 1, representing the system : influenced by prior preferences and pervasiveness of need states

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.

In Agent 1, the need state has a pervasive effect on all the states of the grid-world (except the reward state) as described above, in . Following active inference, this pervasive effect is reflected in the prior preferences for the states to occupy (denoted as C, to follow the usual notation of active inference in discrete time), which is greater for the rewarding state than for any other state.

Using our example of hunger above, and assuming that animals have prior preferences over levels of satiety/need (e.g. a preferred level of sugar in the body), the agent's prior preference over satiety is equal to p(C_p) = 1.
Thus, the preference for each state i given the prior preference is:

$$p(y_i|C_p) = p(y_i, C_p) \quad \text{Eq. 3}$$

The prior preferences followed a normal distribution centered on the value of the preferred state with a degree of sensitivity regarding deviations from that centered value. The center, i.e. prior preference, is the mean and the sensitivity is the variance, and their values were 1 and 0.5, respectively. (Note that the agent may have behaved differently if the variance/sensitivity was different). This followed the logic of the "homeostatic reinforcement learning" model of Keramati and Gutkin (2014) which is based on minimizing the sum of discounted deviations from a setpoint.

Given the pervasiveness of the need state, almost all of the y_i acquire the negative valence and their conditional probability based on prior preference (of satiety) decreases, because their joint probability, i.e. co-occurrence, with satiety decreases. Hence, only the state "food in stomach" has high probability as it co-occurs with the prior preference.

Crucially, as it will become clear in our first simulation, this probability increases as need increases, because the sum of probability of $p(y_i|C_p)$ has to sum to 1. The y_i whose joint probability with C_p decrease are the costly or surprising states, and are denoted $S(h_n)$. Whereas the ones whose joint probability with C_p increase are the ones on the path to the preferred state, and are denoted $\pi(p)$.

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 tends to select policies that lead it to achieve goal states - which in Bayesian terms corresponds to maximizing model evidence (Parr et al 2022).

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. 4

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 $y = s$.) These states embed valence related to some need. Equivalently, the active inference agent strives to minimize surprise, defined as the negative of the evidence:

$-ln\,[P(y)]$ Eq. 5

Importantly, as remarked above, what is surprising or not surprising for an active inference agent depends on its 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" for the agent. 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. 6

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. 7

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. 7 $E_{Q(y|\pi)} ln\,[P(y|C)]$ (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. 8

For completeness, it is important to consider that the quantity shown in Eq. 8 (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 there is no ambiguity and the second term (“epistemic value”) is zero.

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.,

$$P(\pi) = \sigma(-G) \qquad \text{Eq. 9}$$

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, not to action sequences are more commonly considered in active inference) and ultimately to minimize surprise. We assume that in agent 1, the effect of needing does not need to be learned anew, but it is “actively inferred” - reflecting the fact that need states (hunger, thirst) and needed stimuli (food, water) are predefined by natural selection. Indeed, even very simple creatures that have limited (or no) learning capacities are able to navigate their environment adaptively to fulfill their current needs, e.g., by following food gradients (Sterling & Laughlin, 2015).

Agent 2, representing the wanting subsystem : influenced by reward prediction without pervasiveness of need states

Agent 2 had the need state (state 8) but did not have the pervasive effect of need states as Agent 1. To implement an agent that is guided by reward prediction, we use the reinforcement learning framework. The goal of Agent 2 is to maximize a reward function, through reward prediction. Thus, the agent makes decisions based on prediction of rewards assessed by state-action values, i.e. each decision to pursue a course of actions will depend on the value of actions given the current states (see Sutton & Barto, 2018). Here the policies depending on the action values are denoted $Q^{\pi}(s, a)$, and given by:

$$Q^{\pi}(s,a) = E_{\pi}\{R_t | s_t = s,\ a_t = a\} = E_{\pi}\{\sum_{i=0} \beta^i r_{t+i} | s_t = s,\ a_t = a\} \qquad \text{Eq. 10}$$

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) discounted (β^{i}) sum of rewards (r_{t+i}), 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.

Learning of action is updated by the Temporal Difference between the previous action value and the current one as in the equation below:

$$new\ Q(s,a)_{t} = Q(s,a) + \alpha[r + \beta * max\ Q(s,a) - Q(s,a)] \qquad \text{Eq. 11}$$

where α represents a learning rate.

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:

$$max_{a} Q(s,a) \qquad \text{Eq. 12,}$$

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

For a better comparison with the agent 1 (that embodied pervasiveness), the actions of the agent 2 were also transformed into a softmax function:

$$\sigma_Q(s,a) = \sigma(max_{a} Q(s,a)) \text{ Eq. 13}$$

where σ is the softmax function that turns action values into probabilities.

We should keep in mind that agent 2 only illustrates the prediction part of wanting not the dopaminergic reactivity part. However, the dopaminergic state part that amplifies such prediction can be influenced by any brain state that may or may not be a need state as defined in this manuscript (see Berridge, 2004; Zhang et al., 2009).

In sum, we implemented two agents that differ in their policy selection mechanism. Agent 1, which exemplifies the ideas proposed in this article, selects actions based on a prior over policies $P(\pi)$ determined by prior preferences reflecting the pervasiveness of need states as illustrated in Eq 3. Agent 2, which is used as a comparison, selects actions based on reward prediction, in the

form of action values $Q^{*}(s, a)$, as illustrated in Eq 11, reward prediction in itself does not reflect the pervasiveness of need states. The goal of Simulation 1, illustrated below, is to assess the effects of increasing need states on the action selection mechanisms of the two agents.

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

In this simulation, we consider how increasing the costs of the eight "costly" states of Figure 1 from -1, to -2 and -5 (reflecting the fact that the agent can be in a low, mild or severe need state) affects action selection in Agent 1 (left panels of Fig. 2) and Agent 2 (right panels of Fig. 2). For this, we calculate for each level of cost (i.e. -1, -2, -5) the prior policies $P(\pi)$ that allows transition towards 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 Figure 2 show that increasing the cost of the need state, when pervasiveness occurs, significantly increases the probability assigned to policies that reach the rewarding state in 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 need state without (need induced) pervasiveness does not affect reward prediction as shown in Agent 2. This is evident when considering that the softmax of state-action (Q) values assigned by Agent 2 to the rewarding state is always relatively the same in the three right rows.

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; but this phenomenon cannot be fully captured by reward prediction alone. The main point is that need states do not change environmental statistics that predict reward delivery, but they can change the agent's belief over these statistics. This is because, in our account, expected biological costs do not control or constrain state 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 Agent 1, in which need states increase the probability to reach rewarding states. Conversely, Agent 2 correctly infers the reward of the state but it is not sensitive to the pervasiveness of the increase in "need" of the organism. Making Agent 2 sensitive to needs would require extending the basic reinforcement learning scheme with additional mechanisms that (for example) multiply the reward by a need state, learn separate representations of rewards for each need state, or represent deviations from setpoints (Keramati and Gutkin, 2014), see the Discussion.

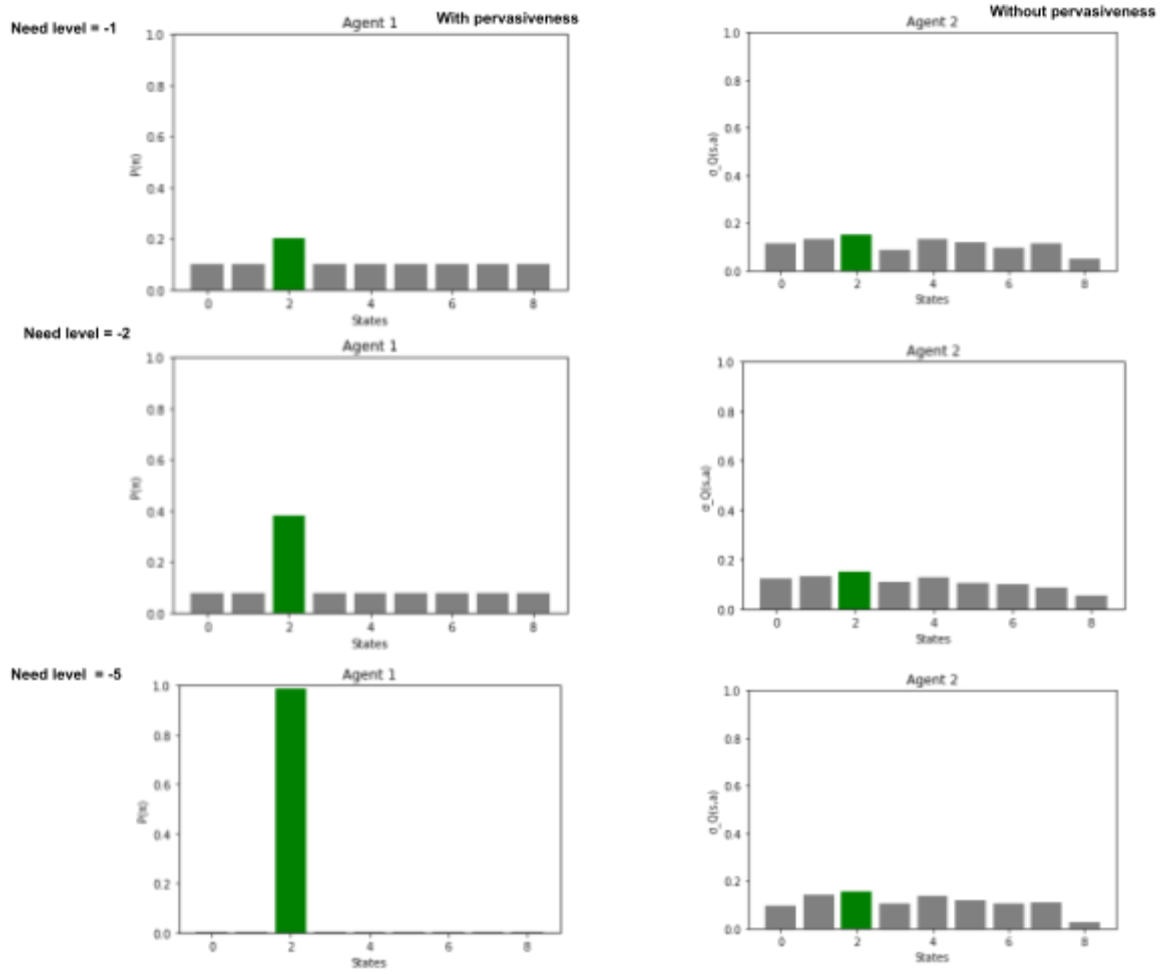

Figure 2. Effects of biological needs on policy selection under embodied prior preference (Agent 1, left panels) and computed reward prediction (Agent 2, right panels). The left and right panels show the results for Agent 1 and Agent 2, respectively. For Agent 1, which uses prior preferences, the y axis plots priors over policies $P(\pi)$ *to reach each of the states of the grid-world, whereas for Agent 2, which computes reward predictions, the y axis plots the softmax of the maximal state-action (Q) values (of the action that reaches that state). As evident when looking at the green bars, both Agent 1 and 2 assign the greater probability to the policy (or action) that reaches the rewarding state 2. The three rows show the effects of setting the costly state (state 8, see Fig. 1) to -1, -2 and -5, respectively. The results show that when need states are pervasive, 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 reward*

prediction assigned by Agent 2 to state 2. This can be apprehended by noticing that in the three panels, Agent 1 assigns different probabilities to the policies that reach state 2, whereas Agent 2 assigns the same probability to state 2. This is consistent with the idea that need states are pervasive and this allows them to (directionally) influence tendencies (i.e. probabilities) more than reward prediction. See the main text for explanation.

3.5 Simulation 2: How needing amplifies wanting

Simulation 1 showed that the action selection mechanism of Agent 1 is sensitive to the pervasiveness and the magnitude of need states. In that we looked at each state of the grid-world separately. In Simulation 2 (using the same grid-world environment), we ask if being in a state of greater need amplifies the wanting of Agent 1, when a reward is present (note that in this simulation, we do not consider Agent 2, since Simulation 1 showed that it is not sensitive to need states pervasiveness). So, we look at the effect of the need state on the system as a whole and its subsystems.

In Simulation 2, Agent 1 dwells in an environment in which the pervasive biological cost that has reached (almost) all states varies 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 entropy, i.e., the precision, which is a measure of certainty about which states to occupy in the future. We compare two conditions; namely, a condition when a reward is present (i.e. the reward state 2 is baited with a reward of 1) or absent (i.e., the reward state 2 has the same cost as all the other costly states).

The entropy is the average of all negative log of the prior preference (policies) of all states Y, and is given by the formula:

$H(Y) = E[- log(Y)]$ Eq. 14

This variable Y can indeed be a surprising state, i.e. embedded with need states, or it can be a rewarding state, i.e. a state on the path to the preferred state. Thus, we calculate the agent's "need related entropy" (or simply entropy) as follows:

$H_n(Y = S(h_n))$ Eq. 15,

when there is no reward, and

$H_{n,p}(Y = S(h_n), \pi(p))$ Eq. 16,

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 which states to occupy by the agent when such state is a path $\pi(p)$ to the preferred state (p). Alternatively, when there is no reward, the entropy is over which states $S(h_n)$ to occupy, but all states S are impacted by need's pervasiveness and 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 as they are embedded with the need. 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. 17

We calculate the precision as the inverse of the entropy:

$P_n = H_n^{-1}$ Eq. 18,

when there is no reward, and

$P_{n,p} = H_{n,p}^{-1}$ Eq. 19,

when there is a reward and hence a path to the preferred state / rewarded state. Given the inequality in Eq. 16, 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. 20

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 subsystem) or to hedonic sensation (liking subsystem), and also to simpler organisms that might not have a sophisticated brain.

In order to show how needing can amplify wanting, we connect our equation of precision to the computational model of wanting and incentive salience by Zhang and colleagues (2009). This model introduces a 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, such that the overall effect is $\overline{r}(r_t, k)$, which corresponds to the precision (neuronal gain) ascribed to reward prediction within the wanting subsystem. 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_t and k, reduces entropy. Indeed, the state value $\overline{V}(s_t)$ can be enhanced by a dopaminergic k induced by a need state (surprising state) when there is a reward, r_t, that signals the path to the preferred state (p). In this case where such amplification $\overline{r}(r_t, k)$ is due to a need state, the reward r_t is (on) the path to the preferred state $\pi(p)$ and the dopaminergic state k is induced by the need state that lead to surprise $S(h_n)$. Thus, we obtain that the function of the enhanced reward $\overline{r}(r_t, k)$, which is encoded as confidence by dopamine, is related to precision over policies leading to reward $P_{n,p}$. 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

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

The results of the simulations of entropy and precision can be appraised 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 Agent 1 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).

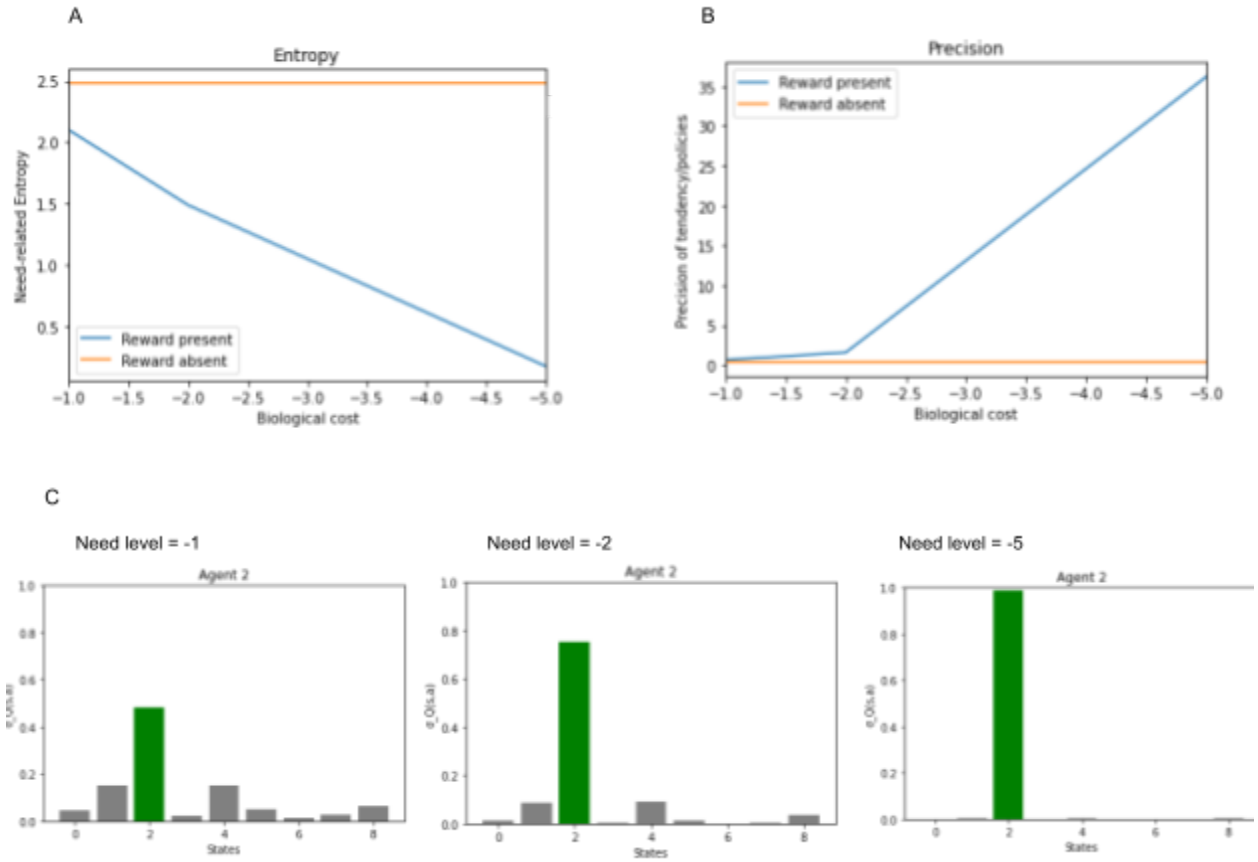

Figure 3. The impact of different need states on the entropy (left plot. A) and on its inverse, the precision (right plot. B), over which state to occupy for an agent that embodies prior preferences, in conditions in which a reward is available (blue lines) or no reward is available (orange lines). On plot C, the same effect on reward prediction, i.e. the wanting subsystem, with different levels of need states. The increase in precision corresponds to an increase in "focus", i.e. higher dopaminergic excitability, towards predictions that lead to the preferred state.

Illustrating the connection between needing and wanting

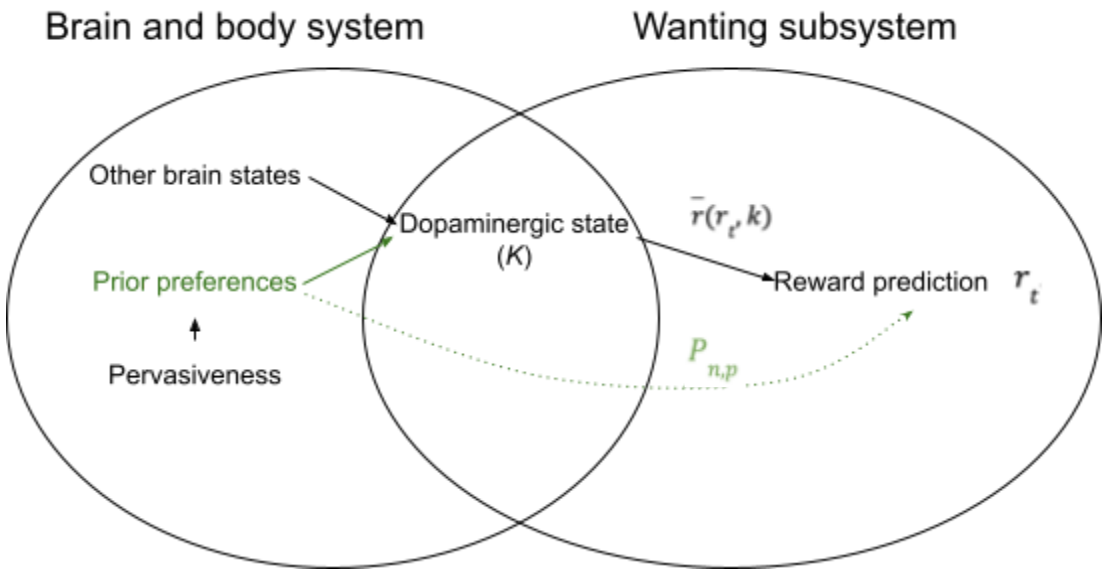

Figure 5. Illustration of the wanting subsystem which depends on the prediction of the reward r_t *under the influence of different inputs that elevate dopaminergic state* k, such that the overall effect is $\bar{r}(r_t, k)$. *If the input is the need state (in green) then* $\bar{r}(r_t, k)$ becomes $P_{n,p}$ and needing amplifies wanting.

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, the pervasiveness of need states creates an effect so that 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 relevant wanting, when there is a reward (or a cue) available. This increase in precision can happen independently on subsystems like wanting, liking, the interoceptive subsystems, etc. Due to the fact that this reduction of entropy (or increase in precision) happens to the system as a whole, it can be transmitted to any subsystem, if a reward is available, depending on the type of events/stimuli treated within that subsystem. In other words, needing can independently influence the wanting, liking, and interoceptive

subsystems, depending on the availability of a stimulus/event (treated by that subsystem) that represents the path towards, or achievement of, the relevant and preferred state.

3.5 Summary of the simulations

To summarize our results, in the first simulation, we looked at the organism in terms of its different states and in the second one we looked at the organism as a whole. To recap, the first simulation showed that need states influence the preference for the rewarding state independently from reward prediction. It also showed a dissociation between needing (which has a pervasive effect on the organism) and the wanting subsystem (which requires reward prediction, in addition to dopaminergic reactivity). The second simulation showed that needing would increase precision in the overall system when there is a stimulus/event (i.e. a state) that alleviates that need. This increase in precision at the system level means that any subsystem of the organism will assign a higher precision or response to the subsystem-specific stimulus, providing that that stimulus leads to the preferred state and alleviates the pervasive need state. So, if the stimulus is a direct contact with a reward, then the liking subsystem might assign a higher response to that stimulus. If the stimulus is a cue signaling the relevant reward, the wanting subsystem might assign higher precision/neuronal response to that cue. The same can be said for the interoceptive system, and so on.

3.6 Mapping from theory to mathematical formulation for both simulations

First we recall that, due to their pervasiveness, need states act upon the entire system, whereas wanting is related to a subsystem, namely the mesolimbic dopamine subsystem. Also, although both simulations are presented separately, there are actually two different measures (or consequences) of the same process, as discussed below.

Simulation 1

This simulation looks at the directional aspect of needing separately from reward prediction. Here, we show how pervasiveness incorporated into prior preferences generates needing and preference for the state that satisfies the need. We show that this can happen in absence of reward prediction (agent 1) and that reward prediction can happen independently of that process (agent 2). Please note that agent 2 does not have the need related pervasiveness effect. (Also, not included (in the agent 2) is the dopaminergic state that amplifies reward prediction for wanting. However such dopaminergic states may be influenced by a variety of states, including, but not limited to need states.)

Talking about need states, the pervasiveness depends on the joint probability between a need state y_n and another state y_i given by $p(y_i, y_n)$ $[Need]$ Eq. 2. Prior preferences then "judges" how any state y_i should be preferred given the preferred state. e.g. satiety, C_p , in which the system expects to be in, and this is given by: $p(y_i|C_p)$. Based on that, the pervasiveness of need states make many states less probable, i.e. surprising, noted $S(h_n)$. Hence the system, expressing states in terms of surprise -$ln\,[P(y)]$ can "infer" the trajectory, i.e. the couple of states which are noted $\pi(p)$, that lead to the preferred state, by choosing the actions from which one expects minimization of such surprise: $-\,E_{Q(y|\pi)} ln\,[P(y|C)]$.

This is different from wanting subsystem where the value or "judgment" of a state or an action $Q^{\pi}(s, a)$ does not necessarily depend on a pervasive effect and vice versa, but is rather based on the expectation of the reward r_t. This expectation is only one part of the wanting subsystem. Indeed, such value or "judgment" depends on r_t, but is modulated by a dopaminergic state/reactivity k, the second part of the wanting subsystem, such that the overall effect is $\bar{r}(r_t, k)$, which corresponds to the precision (neuronal gain) ascribed to reward prediction within the wanting subsystem. However, dopaminergic state/reactivity k can be amplified by many brain states that may or may not be need states, i.e. may or may not act like need states as described above (see Berridge, 2004; Zhang et al., 2009).

Simulation 2

Here, we show how needing's pervasiveness can increase precision in presence of reward cues, and how such precision would result in an amplified wanting if there is a reward cue. We use the process included in agent 1, i.e. pervasiveness (as defined above) and predefined rewards (as prior preference) and apply that in agent 2 (reward prediction system). We find the same increase in precision, which translates to an increase in the action value of the rewarding policy.

Nevertheless, need states can amplify dopaminergic reactivity and thus assign a high precision $P_{n,p}$ to reward prediction that leads to the preferred state. Indeed, by following policies $-\,E_{Q(y|\pi)} ln\,[P(y|C)]$ towards the preferred state and away from surprising states $S(h_n)$, when there is a possibility/path to reach such preferred state, the states on that trajectory $\pi(p)$ become more and more probable, and this decreases entropy over which state to occupy. In other words when one is in need (surprising) state that becomes pervasive, and there is a path to the preferred state, the entropy $H_{n,p}(Y = S(h_n),\ \pi(p)\ )\ =\ H_{n,p}$ decreases by following the policies

$-\ E_{Q(y|\pi)} ln\ [P(y|C)]$. Thus, the inverse of that entropy $H_{n,p}^{-1}$, i.e. the precision (or neuronal gain) $P_{n,p}$ assigned to states (stimuli/events/reward cues) $\pi(p)$ that leads to the preferred state is increased. This increase in precision happens within any subsystem (wanting, liking, interoception, etc.) if such $\pi(p)$ states (stimuli/events/reward cues) happen to be processed by that subsystem. So, the interaction between needing and wanting happens when needing enhances mesolimbic dopamine reactivity which assigns higher precision to reward cues that satisfy the need. The enhancement of dopamine reactivity is the k which amplifies the reward r_t and generates 'wanting' as expressed by $\bar{r}(r_t, k)$ defined by Zhang et al. (2009), and when need states induce dopamine reactivity (k) this $\bar{r}(r_t, k)$ is equal to $P_{n,p}$ and needing amplifies (or induces) wanting towards r_t.

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 the dopamine subsystem and can usually exert a stronger influence on (somato-motor) behavior. However, the respective roles of needing and wanting 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 normative perspective of active inference, and the hypothesis of (embeodied) pervasiveness of need states. .

We firstly defined a need state in terms of surprise. Here, the term "surprise" is used in the sense prescribed by theories like predictive coding and active inference. The key idea is that 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 need state is intrinsically costly (where the cost is related to surprise); and thus a state of need 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 the driving influences of needing on behavior, even without wanting. Secondly, we hypothesized that need states exert a pervasive effect on the system as a whole, making all states, except the relevant rewaring state, biologically costly. Hence it is the embodiment of that pervasiveness into prior preferences that allows needing to activate prior policies that direct the system towards the relevant rewaring state.

Furthermore, we defined wanting as a subsystem (one among others), and associated it to the precision of policies that achieve preferred (goal or reward) states in active inference, 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 acquire incentive salience (Berridge 2004) and determine "wanting". Needing and wanting 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 state space, represented by a grid-world, and implemented the pervasive effect of a need, i.e. most states became associated with biological costs (to mimic a state of need of the organism), except the one state associated with a reward that alleviates that need. Simulation 1 illustrates the possible functioning of the needing system, by focusing on the action selection mechanisms of two agents: Agent 1 exemplifies our hypotheses and bases action selection on prior preferences about states and policies, using a simplified version of active inference. Agent 2 based action selection on reward predictions and action values, learned through a standard form of reinforcement learning (Q learning). This simulation shows that only Agent 1 which embodied its states into its prior preferences shows a sensitivity to need states during action selection. When Agent 1 is in more severe states of need (i.e., when the non-rewarded states are affected by the pervasive effect of increasing need/biological cost), it assigns a greater probability to the policies that lead to the rewarding state. Hence, a state of need can have a directional influence on behavior, leading the Agent 1 towards the preferred states and away from costly or surprising states. Rather, the Agent 2, that uses reward prediction, correctly estimates the reward provided by the rewarding state, but this estimate was not sensitive to the agent's need state. This results shows that the functioning of the need system cannot be reduced to reward prediction per se.

In sum, this simulation illustrates that for an agent that uses prior preferences to embody the pervasiness effect, need states can have directional effects. Needing 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 and 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

Whereas simulation 1 was focused on the probability of choosing states, simulation 2 focuses on changes at the system level. The simulation shows that the presence (versus the absence) of (a path to) a reward decreases the entropy of the states that Agent 1 plans to occupy and increases the associated precision, or the confidence about occupying these states (and about the policies, not shown in our results). This increase in precision can happen independently on the wanting, the liking, the interoceptive subsystems, etc. Since it happens at the level of the entire system, it can be transmitted to any of its subsystems, depending on the availability of a subsystem-specific stimulus/event that represents the path towards, or the achievement of, a preferred (rewarding) state. Thus this simulation illustrates how the simultaneous presence of a state of need and of a reward cue can amplify "wanting" in Agent 1, which based action selection on its prior preferences. 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. In this study, what we added is the impact of pervasive effect of need states on the increase of such precision.

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 Agent 1 is in a more severe state of need (i.e., when the costs of the non-rewarded states 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/or the liking) of stimuli: by reducing entropy and making the agent more confident about what stimulus, response or 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, within the wanting subsystem, would neurophysiologically correspond to increased dopaminergic activity.

4.3 Relations between our proposal and previous computational studies

In this study, we based our analysis on an active inference model that embodies the pervasiveness of need states, which nicely permits modeling the interactions between needing and wanting. However, please note that our simulations are not primarily meant to compare active inference and reinforcement learning. One advantage of using active inference is that its notion of prior preference permits to accommodate quite naturally the notion of pervasiveness of a need state into any other state of the organism. Rather, the choice of using a RL model is motivated by the fact that reward prediction and needing can happen independently, not to prove or disprove the framework of RL. Indeed, there are extensions of reinforcement learning models that can reproduce the influences of need on reward value or on incentive salience.

In summary, our simulations show that active inference provides a natural way to model the distinction between needing and wanting. Many of the models previously proposed often assume an automatic association between needing and wanting (see Keramati and Gutkin, 2014; Zhang et al., 2009). The model proposed here provides a view where needing and wanting are both associated and dissociated. This may help explain a broader range of experimental results, both at the neural and behavioral level. Indeed, the increase in precision which, in active inference, is interpreted as neuronal gain (Friston, 2010) could possibly be viewed as the (need induced) salience that need states ascribe to needed stimuli (see Chen et al., 2016). In this sense, if we interpret the increase in precision (in our results) as a need induced neuronal gain toward stimuli/events that are treated within some subsystem, it becomes clear how needing and wanting can be both dissociated and associated (and also how needing can influence other subsystems such as the liking or the interoceptive ones). For instance, regarding their dissociation, needing can increase precision of stimuli/events treated within other subsystems, such as the liking or interoceptive ones (Cabanac, 2017; Berridge & Kringelbach, 2015; Becker et al., 2019, Bosulu et al., 2022), even in absence of a cue that triggers the wanting subsystem (Wassum et al., 2011 Salamone et al., 2018), and this fits well with some recent meta-analytic results of human fMRI data distinguishing between needing and wanting (see Bosulu et al., 2022). Regarding their association, a specific need can increase wanting when there is the relevant pavlovian reward cue, possibly through such increase in precision, and this can happen even without learning the consequence of that reward in that specific need state (Berridge, 2007; Zhang et al., 2009).

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, and the assumption that the effect of need states is not (always) learned, we modeled the setpoint as a prior and each deviation from it as a surprise, as explained in the description of Agent 1. Also, as shown in simulation 2, there is a close correspondence between our Agent 1 and the model by Zhang and colleagues (2009). This model introduces a variable, k, representing dopaminergic state/reactivity, so that the amplification of reward r_t is a function of the reward itself and of k, that is: $\bar{r}(r_t, k)$. We showed that when need states are responsible for the elevation of dopaminergic reactivity k, the $\bar{r}(r_t, k)$ is equal to the increase in precision $P_{n,p}$ within the warning subsystem. However, our treatment based on active inference helps expressing needing, wanting and their interactions more easily in terms of information-theoretic quantities, namely, information and uncertainty.

4.3 Summary and future developments

Our study adds to a large number of studies showing 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 Agent 2 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.

Note that while in our simulations we focused on the precision of future states occupied by the agent, there are different precisions associated with different subsystems (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. needed 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. Indeed, 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 dopaminergic states that are influenced by other things than the need states, such as some emotions, stress, drug states, etc. (Berridge, 1996; Berridge, 2004). This might mean that those non-need related internal states can amplify wanting-related precision. Hence, 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 that embodies prior preferences 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 could be more related to internal (prior) policies that are more automatic and wanting subsystem could be more in control of 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 can act synergistically. For example, need states increased precision on the overall system when there is a reward leading to the preferred state. In that sense, if a reward which is relevant for the need state, i.e. which is pursued by the prior (internal) policy, is the same as the reward predicted by a cue, i.e. detected by the wanting subsystem specifies the external policy, then the precision will be enhanced and needing and wanting will be synchronized.

Can the ideas discussed here be extended to social needs and wants? Wanting, as defined in the incentive salience hypothesis, has been proposed to possibly be extended to more general, psychological matters in humans (see Pool et al., 2016.; Stussi & Pool, 2022). In humans, needs can also go beyond homeostatic needs and include, for instance, social needs such as belongingness (Maslow, 1943; Baumeister & Leary, 1995; Maner et al., 2007). We argue that social need states may follow the same principles of pervasiveness and embodiment into priors as described here. Indeed, preferred states can also be conceptualized as "modes", both in the statistical and lay sense of the term, and one person may get used to being a certain way (i.e. in a certain state). This could be any kind of non physiological, i.e. any kind of “social state”. This can become incorporated into his/her prior preference, so much that not being that way (i.e. in that social state) now generates a need in that person. This idea would also apply to drugs where it would be important to ask the question whether the person consumes out of need, i.e. the drug state has become embodied as a prior preference over internal states (see Turel & Bechara, 2016), or out of wanting, i.e. excessive dopamine sensitization of policies leading to the drug (Berridge & Robinson, 2016; also see O’Brien et al., 2006). Depending on the case, it is possible to target different brain regions and behavioral therapy techniques.

In sum, this study aimed at providing a conceptual model that defines needing and wanting systems and their interactions; and to demonstrate some of their peculiarities with the aid of simulations based on the normative framework of active inference and the hypothesis of embodied pervasiveness. 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. Also, we would like to acknowledge that, although the use of simulations illustrate the benefits and predictions of the proposed active-inference framework of needing and wanting, this framework remains to be tested using empirical data. 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. This research is supported by funding from the European Union’s Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement

No. 945539 (Human Brain Project SGA3) to GP and No. 952215 (TAILOR) to GP, and the European Research Council under the Grant Agreement No. 820213 (ThinkAhead) to GP.

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

Berridge, K. C., & Robinson, T. E. (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist*, *71*(8), 670.

https://doi.org/10.1037/amp0000059

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

O’Brien, C. P., Volkow, N., & Li, T. K. (2006). What’s in a word? Addiction versus dependence in DSM-V. *American Journal of Psychiatry*, *163*(5), 764-765.
https://doi.org/10.1176/ajp.2006.163.11.2014a

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

Pool, E., Sennwald, V., Delplanque, S., Brosch, T., & Sander, D. (2016). Measuring wanting and liking from animals to humans: A systematic review. *Neuroscience & Biobehavioral Reviews*, *63*, 124-142.

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.

Stussi, Y., & Pool, E. R. (2022). Multicomponential affective processes modulating food-seeking behaviors. *Current Opinion in Behavioral Sciences*, *48*, 101226.

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

Turel, O., & Bechara, A. (2016). A triadic reflective-impulsive-interoceptive awareness model of general and impulsive information system use: Behavioral tests of neuro-cognitive theory. *Frontiers in Psychology*, *7*, 601.
https://doi.org/10.3389/fpsyg.2016.00601

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