The Evaluative Mind

Julia Haas || Ethics Research Team

1. Introduction

In his original introduction to *Mind Design*, John Haugeland observes that "an 'experiment' in mind design is more often an effort to *build* something and make it work, than to observe and analyze what already exists" (Chapter 2, this volume). But what happens when such an experiment in mind design succeeds in a way and to a degree that few could have predicted? And more to the point, how do we take on the implications of such successes when they challenge central features of how we understand the mind?

These are the circumstances that we find ourselves in with respect to developments in reinforcement learning. Over the past twenty-five years, reinforcement learning has had a tremendous impact on the development of artificial intelligence and has been a major driver in advancements in the so-called 'decision sciences'— computational neuroscience, neuroscience, psychology, psychiatry, and economics. But even as we continue to advance the notion of reward maximization as a general solution to the problem of artificial intelligence (Silver, Singh, Precup, & Sutton, 2021), we have not yet embraced the full implications of reinforcement learning, together with the accompanying reward-prediction hypothesis, for our conceptions of the mind. That is, we continue to think of the mind as some form of a thinking machine ("thinking, intellect," (Haugeland, Chapter 2, this volume), where such thinking is best understood as some type of computation— ecumenically including neural networks, deep learning, genetic algorithms, and so on.

I propose that the successes and contributions of reinforcement learning urge us to see the mind in a new light, namely, to recognize that the mind is fundamentally *evaluative* in nature. There are weaker and stronger versions of this thesis.

The weaker version, which I commit to here, proposes that the mind is, at a fundamental level, in the business of evaluating states of affairs as better or worse. This version is additive in nature: it says that, *in addition to* performing computations over representations of descriptive matters of fact, the mind *also* performs computations over representations of those facts as better or worse.[1] But even merely recognizing this heretofore missing piece of the puzzle transforms our understanding of many central aspects of our cognitive experience.

The stronger version, which I explore but ultimately don't subscribe to, makes a revisionary rather than an additive claim: it proposes that the mind is *at bottom* evaluative in nature. This is to say that the mind's evaluative processes are conceptually *prior* to its perceptual, cognitive, or motor processes. In this sense, the strong thesis is a type of grand unifying theory for understanding the mind. Notably, the strong version is related

[1] Of course, many approaches in the philosophy of mind and cognitive science posit what we might call 'compound states,' such as desires, that may be similarly evaluative. But it's consistent with such views that evaluative compound states are outliers - that "other stuff" - and overshadowed by traditional descriptive computations and belief-like states and processes. The weaker thesis makes a stronger claim, in that it posits widespread evaluative processing at a *fundamental* level and, notably, where evaluative processing modulates even belief-like states and processes. Thanks to Murray Shanahan for pressing me on this point.

to but distinct from the so-called 'reward is enough' hypothesis, which suggests that reward maximization is sufficient to "drive behavior that exhibits most if not all abilities that are studied in natural and artificial intelligence" (Silver, Singh, Precup, & Sutton, 2021, 1).

Even without the stronger version, reinforcement learning points us to the idea that, as living organisms, we not only continually experience the world, but experience it *as better and worse*. As Haugeland (1979, 619) puts it, the problem with classical computers is that they "don't give a damn." Montague (2006, 19) similarly suggests that the central difference between computers (as we have more traditionally conceived of them) and brains is that the latter use evolved, efficient computations that "care— or more precisely, [that] have a way to care." In my view, these notions of 'giving a damn' or 'caring' are basically right: minds *assess with respect to some goals,* i.e., they 'care' about how things are going with respect to those goals, be they as central as survival or as mundane as getting coffee.

Still, we need a much more systematic way of working out of what this actually means. Moreover, if we do in fact experience the world in this way— that is, *evaluatively*— then this will have important implications for understanding how many of our cognitive capacities function, e.g., why perception and attention select as they do, and, equally, why these capacities break down as they do, e.g., how Major Depressive Disorder may involve *both* a reduction in the primary sensitivity to rewards *and* an individual's reduced ability to *learn* from reward (Huys, Pizzagalli, Bogdan, & Dayan, 2013). This is the work I aim to do here.

I build my argument out over stages. For precision, I make several assumptions about the nature of reinforcement learning and its instantiation in minds like ours. I sketch these assumptions, together with their relationship to other versions of reinforcement learning, in Section 2. I then offer a brief survey of some of the empirical evidence suggesting that the reinforcement learning paradigm captures something important about biological minds like ours.

In Section 3, I get more specific about what that 'something important' is. I do so by characterizing the nature of valuation in the mind, defending the function of valuation as guiding selection and providing evidence for the ubiquity of valuation as selection across a wide range of 'low-' and 'high-level' human psychological capacities.

In Section 4, I defend the weak version of the evaluative thesis. I sketch what we might expect from a strictly 'thinking' mind on the one hand, and from a thinking, evaluative mind on the other, and I suggest that we find plenty of evidence for the latter in a variety of cognitive capacities.

In Section 5, I consider the strong thesis, mapping out how an argument for it might go. I suggest it is a thesis well worth bearing in mind, particularly as we continue to make advancements in artificial intelligence. Nonetheless, I suggest that we presently lack the necessary evidence to subscribe to it wholesale and raise some challenges for securing it going forward.

In Section 6, I briefly conclude by addressing what Haugeland calls the common complaint regarding artificial intelligence. According to Haugeland, the complaint suggests that artificial intelligence "pays scant attention to feelings, emotions, ego, imagination, moods, consciousness'" (this volume, p. x). I show how by adopting an evaluative account, we can not only illuminate core aspects of minds like ours, but equally appeal to powerful, computational frameworks to design many (though not all) of the features Haugeland refers to into artificial agents.

To start, let's look at the narrow end of the argumentative wedge, namely, with a basic sketch of reinforcement learning.

2. Program, concepts, and findings[2]

2.1 Overview

We can think of reinforcement learning as a research question, as a research program, and as a set of computational tools. As a research question, sometimes called the 'learning problem,' reinforcement learning asks how an agent can optimize its behavior by learning from interactions with its environment. For example, how does a baby plover learn the contours of its environment simply by hopping around in it? Or again, how does a newcomer to London find her way around, just by using a map and a bit of trial-and-error? As a research program, reinforcement learning refers to a branch of computer science, together with associated interdisciplinary approaches, that analyzes formal versions of this question and develops computational solutions to it (Dayan & Abbott, 2001; Glimcher & Fehr, 2013). Finally, reinforcement learning methods are the suites of computational algorithms that aim to solve the aforementioned learning problem (Sutton & Barto, 2018).

Notably, as a research program, the reinforcement learning framework makes certain foundational and technical assumptions, with specific versions of the framework committing to some assumptions while suspending or relaxing others. Here, I sketch what I call the 'reinforcement learning and decision-making' (RLDM) framework, drawing on assumptions made in both machine learning and computational neuroscience.[3] Specifically, in addition to assuming many of the somewhat more basic features of the framework, this version assumes that reinforcement learning is to some degree meaningfully instantiated in the minds of biological organisms, and takes a particular if minimal view regarding the problem of specifying where rewards come from in biological systems. Throughout, it will be useful to remember that this is just one— though one perhaps particularly philosophically useful— variant of the framework among many.

2.2 RLDM

Let's start with the basic ingredients. In a reinforcement learning framework, we have an agent and an environment. The agent is the learner or decision-maker in question, and it selects different actions in its environment, where actions can be understood as "any decisions we want to learn how to make," including mental actions (Sutton & Barto, 2018, 50). The environment refers to everything 'outside' of the agent, which the agent cannot arbitrarily change but rather with which the agent interacts.[4] The agent and the environment

[2] This section is indebted to Sutton & Barto (2018) and, especially, to Neil Rabinowitz.
[3] Name adapted from Gęsiarz & Crockett (2015).
[4] For example, in many cases, even parts of the agent's body are considered to be a part of the environment.

interact in the sense that the agent is presented with sensory information from the environment, and the agent chooses different actions within the environment (sometimes called a 'state-action pair'); the environment is then affected by these actions. Notably, the agent may not be able to observe the complete environment and may have no prior knowledge of the environment's dynamics. In addition, the agent may, but by no means needs to, build a model of the environment in order to choose actions in and learn from it.
Crucially, one of the distinguishing features of the reinforcement learning framework is the role of reward. Roughly, in reinforcement learning, the agent's objective in the environment is to maximize the cumulative reward it receives over time, where rewards are passed from the environment to the agent. In their influential text, Sutton and Barto call this framing the *reward hypothesis*, specifying, "all of what we mean by [an agent's] goals and purposes can be well thought of as the maximization of expected value of the cumulative sum of a received scalar signal (called reward)" (2018, 53). That is, the agent's objective is to maximize its yield of reward as it acts in the world, and this objective is characterized by assigning a quantity of intrinsic desirability to each state (or to taking each action in each state), known as the reward.[5]

This intrinsic desirability assigned to each state (or taking each action in each state), or *reward*, can be contrasted with the notion of *value*, which captures the expected, discounted, sum of future reward associated with each state (or each action in each state), conditional on a certain policy of action. We can elucidate the distinction between reward and value further using an example adapted from Silver (2015). Imagine an agent in an environment with a door. Upon arriving at the door for the first time, the agent receives a *reward* from the environment. But this reward can also be used to assess how relatively good (valuable) individual states are expected to be to the extent that, conditional on a certain action policy, they *lead to* the door and hence the reward. Hence, an agent's ongoing interactions with its environment enable it to continually revise the value attributed to a given state or state-action pair conditional on a certain policy, upgrading or downgrading as needed. This enables the agent to *learn* the most appropriate actions in the most appropriate states to maximize cumulative reward over time, conditional on a certain policy, in spite of the fact that states (or state-action pairs) can be of high *value* without being intrinsically worthwhile (i.e., *rewarding*).[6] This partly helps explain why not every state in an environment needs to be directly rewarding in order for an agent to act appropriately within it.

As a branch of machine learning, reinforcement learning represents the foregoing conceptual features in computational terms. There are countless reinforcement learning algorithms, each with a distinctive computational profile. For example, the *temporal-difference learning algorithm* represents a computationally efficient way of making predictions about reward in the future. One way to improve predictions over time is to make a prediction about an actual outcome, compare the difference (or error) between the two, and then update that prediction. To borrow an example from Sutton (1988, 10), one can make a prediction on Monday about the weather on Saturday, wait until Saturday, and then update Monday's prediction based on the difference between Monday's prediction and Saturday's actual weather. The temporal difference approach does something

[5] Thanks to Neil Rabinowitz for this formulation.
[6] We can take the example of making and having coffee to help illustrate the difference between a state of high expected value that is nonetheless not technically rewarding. Although only drinking a cup of coffee itself may be intrinsically worthwhile (rewarding), and the grinding of the beans almost certainly is not, the state-action pair of grinding the coffee is nonetheless associated with *expected value*, as it is, conditional on a certain policy, a necessary step or state-action pair on the way to having the coffee.

a little neater by updating its predictions throughout. That is, it is more akin to one making a prediction on Monday about the weather on Saturday, but then comparing Monday's prediction to *Tuesday's* prediction about Saturday and adjusting accordingly, and so on. For instance, if Monday's prediction for Saturday is a 90% chance of rain, but Tuesday's prediction for Saturday is only a 60% chance, then the temporal difference approach is to lower Monday's prediction.[7]

Notably, given that different problem settings present different challenges, there are myriad different RL algorithms in use today. These trade off factors such as memory consumption, computation cost, data efficiency, and stability; some are useful for very small environments, and others are useful for very large environments; some for discrete action spaces, and others for continuous ones.[8] Thus, 'reinforcement learning' refers to a general learning problem and a suite of computational algorithms, as well as to the branch of computer science devoted to studying them, rather than to any token solution to the problem.

The RLDM version of reinforcement learning adds two assumptions to the basic reinforcement learning framework. First, it assumes a relationship between reinforcement learning and the minds of biological creatures like us. This assumption is by no means universally held: researchers in machine learning can pursue decades of research and remain entirely agnostic regarding the role of reinforcement learning in biological agents. Similarly, cognitive and comparative psychologists can study the nature of learning and behavior without any appeals to the reinforcement learning framework. However, RLDM follows computational neuroscientists and other decision scientists who suspect that reinforcement learning does, in fact, capture something special about minds like ours. As Dayan and Niv (2008, p. 1) put it, reinforcement learning appears to offer

> More than just a computational, 'approximate ideal learner' theory for affective decision-making. [Reinforcement learning] algorithms, such as the temporal difference (TD) learning rule, appear to be directly instantiated in neural mechanisms, such as the phasic activity of dopamine neurons. That [reinforcement learning] appears to be so transparently embedded has made it possible to use it in a much more immediate way to make hypotheses about, and retrodictive and predictive interpretations of, a wealth of behavioral and neural data collected in a huge range of paradigms and systems.

Notably, we are free to relax the condition that reinforcement learning is directly *instantiated* in the workings of the brain. It is sufficient to say that reinforcement learning provides remarkably useful frameworks for thinking about decision-making and selection in the mind.

RLDM's second assumption has to do with the *subjective* nature of reward. As noted above, in the basic reinforcement learning framework, rewards are passed from the environment to the agent when an agent enters certain states of the environment, or when the agent takes certain appropriate actions in certain appropriate states. This external nature of reward is unproblematic in the context of machine learning because the reward is simply designed by the researcher as a means of communicating what the researcher wants the artificial agent to achieve. But things get thornier when we get to biological organisms, since it's not clear where rewards would

[7] For a more detailed discussion, see Sutton and Barto (2018, Chapter 6, and especially Example 6.1.).
[8] Thanks to Neil Rabinowitz for this formulation.

then come from. This question regarding the origin of reward in biology generates what Juechems and Summerfield call the *paradox of reward*. The issue is paradoxical, the authors contend, because,

> No external entity exists that can directly quantify the consequences of each action, like the points that are awarded in a video game for completing levels or shooting monsters. Nor is it obvious that biological systems have a dedicated channel for receipt of external rewards that is distinct from the classical senses. Rather, rewards and punishments are sensory observations— the taste of an apple, the warmth of an embrace— and so stimulus value must be inferred by the agent, not conferred by the world. In other words, rewards must be intrinsic, not extrinsic" (2019, 837-838).

Exactly how this conversion between sensory observations and assignments of intrinsic rewards occurs— assuming that it occurs at all— remains the subject of lively theoretical debate. One possible explanation is that minds like ours have evolved specific mechanisms that convert sensory observations into hedonic signals (e.g., see Schultz, 2015). Another, complementary possibility is that, in addition to the evolved mechanisms for basic rewards (e.g., food and water), human beings develop cognitive setpoints, akin to homeostatic setpoints, on which reward amounts to a by-product of computing the distance to self-defined goals (e.g., such as getting married or going to graduate school) (Juechems & Summerfield, 2019). Here, RLDM again takes a minimal approach, and merely assumes *that* minds like ours subpersonally assign subjective rewards to, *e.g.*, sensory observations, albeit indirectly; it remains provisionally agnostic about how this assignment takes place.

2.3 Substantiating the first assumption

Finally, let's explore the first assumption in more depth. In what sense does RLDM provide a distinctive, interpretive lens for cognitive neuroscientific evidence?

As gestured at above, arguably the most significant connection is between RLDM and the reward system in the mammalian brain. In the mid-1990s, theoretical and empirical work showed that the firing of dopamine neurons is closely described by the temporal difference learning algorithm (for narrative accounts of the discovery, see Montague, 2006; Redish, 2013; see also, Colombo, 2014). That is, dopamine neurons fire when an organism experiences a higher- or lower-than-expected value in association with a given state (Schultz, Dayan, & Montague, 1997). This discovery provides the foundation for the so-called *reward prediction error hypothesis of dopamine neuron activity,* which holds that "one of the functions of the phasic activity of dopamine-producing neurons in mammals is to deliver an error signal between an old and a new estimate of expected future reward to target areas through the brain" (Sutton & Barto, 2018, 381).

This seminal finding in turn led to the use of reinforcement learning methods to study the neuroscience of vision (Hayhoe & Ballard, 2005; Hikosaka, 2000; Hickey et al., 2010), attention (Della Libera & Chelazzi, 2009; Chelazzi et al., 2014; Anderson & Kim 2018), memory (Patil et al., 2017; Ergo, De Loof, & Verguts, 2020), prospective memory (Krishnan & Shapiro, 1999; Katai et al., 2003; Kliegel et al., 2005; Walter & Meier, 2014), cognitive control (Savine & Braver, 2010; Botvinick & Braver, 2014; Chiew & Braver, 2014; Cubillo, Makwana, & Hare, 2019) , and above all, decision-making (Sutton and Barto, 2018; Dayan and Niv 2008; Rangel, Camerer, & Montague, 2008; Dayan 2011; Glimcher & Fehr, 2013).

For example, a systematic body of evidence now indicates that the reward system guides visual fixation and saccadic eye movement, i.e., what we look at, when, and in what order (Liao & Anderson, 2020). Similarly, reward guides what we do or don't attend to more precisely than do either location or salience (Anderson & Kim, 2018). Conversely, deficits and disruptions (e.g., by addictive substances) to the reward system are not only implicated in diseases such as Parkinson's and Tourette's, but also in a range of psychiatric disorders, including depression (Huys, Daw, & Dayan, 2015) and addiction (Hyman, 2005; Redish, Jensen, & Johnson, 2008; Redish, 2013). Arguably, methods from reinforcement learning thus represent an important and, to date, under-utilized framework for elucidating the nature and mechanisms underlying selection between competing states of affairs across a range of 'low'- as well as 'high-level' kinds of cognitive processing.

When we say there's something special about RLDM, then, proponents tend to point to one or both of the following considerations. First, reinforcement learning algorithms successfully predict and characterize the workings of the reward system; by contrast, other approaches, including predictive processing (see Clark, this volume), often provide merely retrodictive explanations of known phenomena.

Second, the reward system appears to play an outsized role in a range of cognitive capacities, from sensation through to economic choice. The question is, what's the best way of characterizing this role of reward and value in the mind, from a philosophical point of view?

3. Valuation

3.1 Overview

In principle, the role of reward can be characterized at multiple levels of explanation and across multiple, co-dependent theoretical domains, including in computational terms, cellular and systems neuroscientific terms, cognitive neuroscientific and neuroeconomic terms, and psychological and behavioral terms (Hochstein, 2016). For instance, as discussed above, we can capture the role of reward and value in computational terms using methods from reinforcement learning (for an overview, see Sutton & Barto, 2018; though see also hybrid approaches, such as that put forward in Gershman, 2015). Or again, following the Schultz, Dayan, & Montague (1997) discovery, we can characterize reward and value in cellular and system neuroscientific terms, both in terms of dopaminergic functioning as well as in terms of the more general, system-level neural analyses of the reward system in the brain. At a 'higher' level still, we can characterize reward and value in cognitive neuroscientific and neuroeconomics terms, drawing on behavioral experiments and fMRI data, and using constructs such as 'decision-making,' 'motivation,' and 'willingness-to-pay.' And so on.

In what follows, I characterize the role of reward and value in the mind at roughly a 'conceptual' level of explanation, i.e., at a coarseness of grain typical in the philosophy of mind. Accordingly, my argument also broadens out at this stage, moving from the specifics of RLDM and associated empirical evidence to a more traditional, philosophical characterization— namely, to characterize a cognitive process I'll call *valuation*. This is essential for future work in the philosophy of mind, e.g., to enable us to distinguish and understand the relationship between, say, valuation and the philosophical folk psychological notion of *desire* (for work in this spirit, see Schroeder, 2004; Arpaly & Schroeder, 2014), or again, to enable us distinguish and understand the relationship between, valuation and the various notions of *affect*, *mood*, and *emotion* (for a philosophical discussion of emotion see, e.g., Scarantino & de Sousa, 2021).

In this way, the resulting characterization of valuation in some cases *complements* and in some cases *revises* the traditional conceptual machinery used to describe and understand the mind and minds like ours.

3.2 Characterising valuation

Recall from the previous section that in basic reinforcement learning, reward is some quantity assigned to represent the intrinsic desirability to each state (or to taking each action in each state). Further, this intrinsic desirability assigned to each state (or taking each action in each state), or reward, can be contrasted with the notion of value, which captures the expected, discounted, sum of future reward associated with each state (or each action in each state), conditional on a certain policy of action. So, while coffee is intrinsically rewarding for me in the morning, grinding coffee or getting milk is not— but these latter states are nonetheless valuable to the degree that, conditional on my action policy, they lead me to my cup of coffee.

Recall in addition that, according to RLDM, the reward hypothesis captures something special about the mind, namely, the substantial role of the reward system in the mammalian brain, where the reward system is itself implicated in a wide range of 'low-' and 'high-level' cognitive capacities.

I argue that if both of these claims are right, then we can use RLDM and the corresponding empirical evidence to revise our philosophical understanding of what the mind is doing, how it is going about it, and what this kind of processing is for.

Let's start with the 'what.' Very simply, I argue, the mind engages in *valuation*. Informally, I take this to mean that the mind continually attributes reward and value to a range of sensations, perceptions, actions and so on - essentially forming a kind of evaluative layer over the features of its experience.

In more technical terms, I argue that valuation refers to the subpersonal attribution of goal- and context-dependent subjective reward and value to internal and external stimuli. Valuation is *subpersonal* in the sense that it demarcates a causal rather than an intentional mechanism (Dennett, 1969; Drayson, 2014). This is key: the mind routinely, mechanistically assess states of affairs as better or worse.[9] Further, it is *goal- and context-dependen*t in the sense that what is rewarding or valuable depends on what the agent is trying to do, and when and where the agent is trying to do it. For example, if my goal is to wake up and have a productive day, then drinking a cup of coffee first thing in the morning is valuable. But if my goal is to rest and get a good night's sleep, then drinking a cup of coffee late at night is not. It is *subjective* in the sense that what is considered rewarding and/or valuable is agent-relative; while this author finds coffee rewarding, many individuals do not. And the term *stimuli* here is intended as a broad catchall: reward and value can be attributed to external objects

[9] This subpersonal process very likely plays a role in our personal-level experiences of 'value,' 'valuing,' and 'values,' e.g., see foregoing discussion of willingness-to-pay. But the focus throughout the remainder of this paper will be on the nature and workings of the subpersonal process.

(commodities), states, state-action pairs, and action policies, but also to internal states of affairs, such as experiences, feelings, and moods.

In terms of the 'how,' valuation is realized in a number of complementary ways. One important way is through the retroactive attribution of value to states that lead to reward in subsequent states. Recall the task of walking to a nearby door in the previous section. Upon arriving at the door for the first time and therefore receiving or experiencing the *reward*, there occurs a subpersonal, retroactive attribution of *value* to the antecedent states that then led to the reward. That is, there occurs a subpersonal, retroactive attribution of value to the penultimate state, derived from the reward associated with arriving at the 'ultimate' state, i.e., the door. This retroactive attribution in turn continues to feed backwards, i.e., there occurs the subpersonal, retroactive attribution to the antepenultimate state, and so on. In this way, ongoing interactions continue to revise the value attributed to a given state or state-action pair, upgrading or downgrading as needed. For instance, if the baby plover finds a new trove of bugs, the value of a certain path leading to the beach can increase. But values can also be computed 'on the fly' (Balleine & Dickinson, 1998; see more recently Langdon, Sharpe, Schoenbaum, & Niv, 2018), relative to features of context (e.g., Hunter & Daw, 2021), and with respect to imagined or expected future states (Gagne & Dayan, 2021; Russek, Momennejad, Botvinick, Gershman, & Daw, 2021). For instance, if the newcomer to London is traveling from Green Park to Russell Square and Holborn Station is under renovation, the value of taking the blue line decreases.

Here, the main idea is that the mind continually assesses and reassesses states of affairs as better or worse, constructing and casting, to put things in fairly figurative terms, a kind of evaluative fabric over its states and experiences.

3.3 Valuation as selection

But it's the 'what for' of valuation that is of most interest (as these things tend to go).

The *function* of valuation in minds like ours, I argue, is to solve for what I call the *selection problem*, or the problem of selecting between one or more competing alternatives. The selection problem can be described in general terms, insofar as the mind must continually select what to *compute,* what to *sense*, what to *perceive*, what to *attend to*, what to *choose* (as an action in the world), and so on. Technically characterized examples of the selection problem include selecting between *multiple action controllers* (Daw et al., 2005), the problem of *perceptual decision-making* (Gold & Shadlen, 2007), and the problem of *action-based decision-making* (Glimcher, 2011). Crucially, as the span of these examples should illustrate, the selection problem occurs ubiquitously in the mind. It occurs at every major stage of mental processing, from sensation and computation to action, and at every level of description of mental processing, from the sub-personal to the personal.

A central, underappreciated upshot of the RLDM's experiment in mind design, I argue, is that the mind selects between available computations, sensations, perceptions and so on *conditional on* attributions of reward and value.

To illustrate, consider the unlikely phenomenon of binocular rivalry. Binocular rivalry occurs when one stimulus is shown to one eye at the same time as a different stimulus is shown to the other. The resulting experience is of the two images alternating back and forth; *perceptual dominance* in binocular rivalry refers to one of the two images appearing first, or for a longer period of time during the overall duration of the experience of alternation. Notably, both rewarded stimuli and rewarded percepts result in perceptual dominance; that is, participants are more likely to perceive stimuli and percepts associated with a reward (Balcetis, Dunning, & Granot, 2012; Wilbertz, Van Slooten, & Sterzer, 2014; Marx & Einhauser, 2015; Haas, 2021). Moreover, a complementary phenomenon occurs for punished percepts: participants experience perceptual dominance for the *non*-punished percept in the pair, suggesting that the reward or punishment is *not* simply additional information taken into consideration by Bayes-like predictive processing, as a predictive processing view might suggest (Wilbertz, Van Slooten, & Sterzer, 2014). In this way, *the selection of the perceptually dominant percept is directly conditional on the attribution of reward and value in the binocular rivalry paradigm, i.e., on valuation.* Participants tend to perceive the most rewarded or valuable stimulus or percept. Hence, when it comes to the cognitive task of selecting 'what to perceive,' valuation plays a driving role.

But valuation doesn't just play a driving role in perception. Rather, when I say that the mind is fundamentally evaluative in nature, I mean that we *sense*, *perceive*, and *attend* to the features of our environment conditional on our distributions of reward and value attribution, as when we attend to rewarded rather than salient or location-based percepts (Anderson and Kim, 2018). We *remember*, *remember to remember* (remember prospectively) conditional on reward (for a useful review, see Walter & Meier, 2014). We *allocate our cognitive resources* (in cognitive control) conditional on our distributions of reward and value attributions, as shown by the expected value of control account of cognitive control (Musslick, Shenhav, Botvinick, & Cohen, 2015). And we *decide*, *choose*, and *plan* our future actions conditional on our distributions of reward and value attributions, as when prior reward experience determines a participant's willingness-to-pay in everyday economic transactions (Plassman, O'Doherty, & Rangel, 2007).

Conversely, when the reward system is impaired, for example, through cell death in the basal ganglia (Parkinson's) or due to allostatic shift (substance addiction), there are direct, corresponding deficits in selection: e.g., in motor tremors, mood disorders, and executive dysfunction in Parkinson's disease, and e.g., in cravings, impaired control, and continued use in spite of overwhelmingly negative consequences in substance addiction (for extended discussions, see Redish, 2004; Redish, Jensen, & Johnson, 2008). And so on.

To emphasize, I do not argue that selection is *synonymous* with valuation. But selection is *conditional* on valuation: we select or avoid what we learn is better or worse over a life-long course of iteration.[10] Moreover,

[10] It is worth emphasizing that valuation needn't be 'online' in order to guide selection. On the contrary, as in the foregoing example of retroactive attribution, selection can and often is informed by past reward and value attributions. And this 'carried over' feature of valuation as selection in turn has important implications for the nature of self-regulation and control, insofar as it implies that at least in many cases, we do not have direct, intrapsychic control over our motivational states (see Haas, in prep). Thanks to Neil Rabinowitz for pressing me on this point.

valuation is deployed and redeployed across a range of selection problems in the mind, including selection in sensation, perception, attention, and cognition generally.[11]

Hence, where I suggested above that the reward system "influences" or is "implicated in" a range of cognitive processing, I can now be much more specific: *valuation guides selection* across the range of mental processing that occurs in minds like ours.

4. The weaker thesis

What, then, of the evaluative thesis, or the view that the mind is fundamentally evaluative in nature?

At the outset, I suggested that on the weaker version of the view, the mind encompasses *both* thinking *and* evaluation. That is, according to the weaker thesis, the mind does something like 'see' two competing stimuli in binocular rivalry, and 'perceive' only one of those stimuli at a time, resulting in the signature perceptual experience of perceptual alternation. In a standard case, we might also say than an individual could go on to draw on this perception to form beliefs, draw inferences, and perform all the other kinds of cognitive tasks that are typically associated with, as Haugeland put it, ("thinking, intellect," (Haugeland, Chapter 2, this volume), or as others put it, "intelligence."

But, on the weaker version of the evaluative thesis, the mind *also* does something else, without which it would not be the mind it is— namely, it continually assesses things as better or worse, conditional on certain goals and aspects of the environment, in the ways described above, i.e., subpersonally, through various forms of attribution, in a two-place relation, and so on.

In this sense, the weaker thesis doesn't exactly try to unseat the traditional conception of thinking mind but rather complements it by describing a *fundamental* cognitive process that has heretofore been relatively overlooked.

I defend the weaker thesis on three grounds.

First, evidence bears out the positive features of the view. A survey of mature, textbook neuroscience suggests that the reward system is indeed implicated in *basic biophysical processes* such as eating, drinking, and reproduction; in *basic cognitive processes* such as working memory, executive functioning and time estimation; and, crucially, in *all learned behaviors*, ranging from learning-based sensory processing through planning, strategizing, and second-order preference-formation (for a concise review, see Arias-Carrión, Stamelou, Murillo-Rodríguez, Menéndez-González, & Pöppel, 2010; for extended discussions, see Glimcher & Fehr, 2013). Equally, the reward system is implicated in the kinds of *'sophisticated' cognitive processe*s that are often of interest to philosophers, including in emotional responding, social preference formation, speech and language

[11] And has been for millions of years: see, e.g., the role of reinforcement signaling in *Drosophila* (Waddell, 2013; see also Haas & Klein, 2020). Though this is beyond the scope of the current paper, valuation appears to be a highly conserved cognitive process.

processing (see especially Simonyan, Horwitz, & Jarvis, 2012; and also Ripolles et al., 2014; McNamara & Durso, 2018), and generalization.

Second, *predictions* made by the weaker thesis are better supported than predictions made by competing theoretical accounts, e.g., by accounts in the predictive processing space or accounts emphasizing the role of emotions in our cognitive processes. Returning to the example of binocular rivalry offers a good example of the former comparison. The weaker thesis predicts that rewards (and negative rewards, i.e., punishments) should influence perceptual dominance in binocular rivalry; predictive processing accounts make no such prediction, and in fact struggle to explain this type of finding post hoc. But as noted above, reward modulates perceptual dominance in binocular rivalry (Haas, 2021).

An example of the latter type of comparison might involve competing explanations of psychopathy. The weaker thesis proposes that psychopathy is a disorder of valuation, perhaps involving an inability to predict negative outcomes, and/or an inability to update appropriately following negative experiences (E.g., see Oba, Katahira, & Ohira, 2021). By contrast, on an account of psychopathy emphasizing emotions, individuals with psychopathic traits fundamentally suffer from a disorder of empathy, or the ability to respond appropriately to emotional stimuli (Hare, 1998; Soderstrom, 2003; Blair, 2007; Brook & Kosson, 2013; Domes, Hollerbach, Vohs, Mokros, & Habermeyer, 2013; Blair, 2018). Accordingly, the former but not the latter account predicts that individuals with psychopathic traits will exhibit deficits in basic economic decision-making. Here, some evidence seems to bear out the weaker thesis: controlling for other deficits, psychopaths appear to perform significantly worse on the Iowa Gambling Task (Mahmut, Homewood, & Stevenson, 2008) as well as on other types of risky decision-making (e.g., Takahashi, Takagishi, Nishinaka, Makin, & Fukui, 2014).

Third, *deficits* in the reward system corroborate the view. Here, standard cases again emerge in the computational and cognitive neuroscientific literature, including regarding the aforementioned Parkinson's and Tourette's diseases, as well as diseases such as Major Depressive Disorder and different categories of substance addiction. Take the case of prospective memory, or the ability to 'remember to remember.' I suggested above that, like so many of our cognitive capacities, prospective memory is conditional on valuation; we are more likely to 'remember to remember' something in the future when it's associated with a reward. For example, participants show higher prospective memory performance for tasks that were associated with a monetary reward as compared to those that were not (Krishnan and Shapiro, 1999). By extension, consistent with the weaker thesis, we would expect to see *deficits* on prospective memory tasks among individuals with Parkinson's disease. The reasoning goes like this: prospective memory is conditional on valuation, valuation by realized in the reward system in the brain, and the reward system is compromised in Parkison's disease. Hence, we should expect deficits on prospective memory tasks among individuals with Parkinson's.

And this is indeed what we find. Individuals with Parkinson's exhibit impairment in several core stages of prospective memory, most notably when it comes to the phases of intention formation and intention initiation (Katai et al., 2003; Kliegel et al., 2005; Kliegel, Altgassen, Hering, & Rose, 2011; Pirogovsky, Woods, Filoteo, & Gilbert, 2012; Ramanan & Kumar, 2013; D'Iorio, et al., 2019; Coundouris et al., 2020; though see Zabberoni, Carlesimo, Peppe, Caltagirone, & Costa, 2017; Kinsella, Pike, Cavuoto, & Lee, 2018). Analogous arguments propose that impaired reward valuation, i.e., the dysfunctional underestimation, downgrading, or failure to update regarding rewards in individual with Major Depressive Disorder (Takamura et al., 2017; Rupprechter, Stankevicius, Huys, Steele, & Seriès, 2018; Rupprechter, Stankevicius, Huys, Seriès, & Steele, 2021) may explain why this demographic also exhibit systematic deficits in prospective memory tasks (Altgassen, Kliegel, & Martin, 2009; Chen, Zhou, Cui, & Chen, 2013; Li, Weinborn, Loft, & Maybery, 2013; McFarland & Vasterling, 2018). And so on. The basic structure of this second kind of argument, then, is to identify a cognitive capacity modulated by valuation; identify a disease that either upregulates or downregulates valuation (via the reward system) and then determine whether, as predicted by the weaker thesis, individuals with the relevant disorder also exhibit deficits on the corresponding cognitive capacity.

Finally, a common feature of all three sets of reasons should by this stage have become clear: namely, that kind of evidence is inductive in nature, which is to say that each piece catalogs a *confirming* instance of the weaker thesis. Saying 'valuation is ubiquitous in the mind' is akin to saying 'lots and lots of swans are white.' This means that the weaker thesis can be *disconfirmed* - namely, by uncovering a meaningful number of instances where cognitive selection is clearly not, at least in part, underwritten by valuational processes. But this is in fact precisely why I defend the weaker thesis. The normative principles originating in RLDM, together with evidence from the decision sciences, enable us to make a principled but nonetheless fundamentally *empirical* claim about a certain process in the mind - where this claim already brings with it significant high-level implications for understanding the workings of the mind.

By contrast, these same principles and evidence, to my mind, will struggle to bear out something conceptually stronger, including a universal claim regarding the role of valuation in the mind, which I discuss next.

5. The stronger thesis

Whereas the weaker thesis holds that valuation is empirically ubiquitous in the mind, the stronger thesis proposes that the mind is *at bottom* evaluative in nature.

There are a couple of ways of understanding the stronger thesis. It can mean that valuation as selection guides *all cognitive selection* in the mind, such that valuation amounts to grand unifying theory for exploring the nature of the mind. This is the stronger, universal version of the weaker thesis. And it can mean that valuation is ontologically prior *to* and thus conceptually necessary for understanding the mind's perceptual, cognitive, and motor processes. We can call the former claim the *scope commitment* and the latter the *priority commitment*.

Prima facie, one might assume that a proponent of both RLDM and valuation would by extension directly subscribe to one or both of these commitments. As we will see, they may hold some theoretical advantages over the weaker thesis. They are also nominally more in line with the prominent 'Reward is Enough' hypothesis (Silver, Singh, Precup, & Sutton, 2021). Nonetheless, I don't commit to either.

So, why not go whole hog and defend the stronger version of the evaluative view? Let's start with the scope commitment.

5.1 The scope commitment

As its name suggests, the scope commitment is a supercharged version of the weaker thesis. Whereas the weaker thesis holds that valuation is ubiquitous in the mind, the scope commitment holds that valuation lies at the heart of *all* cognitive capacities. Hence, where the weaker thesis suggests that 'lots of swans are white,' the scope commitment rounds up to claim that 'all swans are white,' period.

So formulated, the central challenge with the scope commitment should quickly become obvious: the scope commitment requires defending a universal claim, and no amount of evidence will get us there, as there's always the possibility of an untested counterexample somewhere.[12] The scope commitment is just too easily falsified.

Moreover, it simply doesn't strike me as *likely* that valuation underwrites everything of interest in the mind. The evolved mind is a messy artefact, and at a bare minimum, we can expect 'spandrel' capacities that don't rely on valuation in any interesting sense. I can get plenty of mileage out of the weaker thesis without needing to extend it to the logical limit.

This leaves us with the priority commitment.

5.2 The priority commitment

The priority commitment is trickier to deal with. The priority commitment makes an ontological claim about the mind, analogous to action-first theories in cognitive science, namely, suggesting that our 'thinking' processes are *conditional* on our evaluative processes (for a review of action-first theories, see Briscoe & Grush, 2020). That is, we have the memories, beliefs and so on that we do *in virtue* of our assessments of better or worse.

To take a concrete example of this kind of theorizing, one might argue that the normative function of episodic memory is not to encode a past event 'as it actually happened,' but rather to encode a past event in light of what it might be useful for an agent to remember - and by extension, *do* - in the future.

Adopting the priority commitment enables us to make top down rather than inductive predictions regarding the workings of various cognitive capacities. For instance, to continue with the case of episodic memory, adopting the scope commitment can help us make predictions about what will and won't be remembered, or why individuals experience flashbulb memories (if indeed they do). On the priority commitment, flashbulb memories may contain such an impressive level of detail because, following a traumatic event, it is not clear which features of the preceding event are most relevant to future action, such that 'all' of them are carried forward for future learning. This interpretation draws a close connection between flashbulb memories and the more general credit assignment problem in reinforcement learning, or the problem of determining which actions lead or led to a given outcome (Minsky, 1961; Sutton & Barto, 2018).

[12] Thanks to Carl Craver for helping me drill down on this point.

This kind of hypothesis generation is certainly appealing. It's also pretty tempting to defend the priority of valuation as a way of counteracting the standard emphases placed on computation (and predictive processing!) in the philosophical and cognitive scientific literatures. Still, I stop short of doing so, for two reasons.

First, where the scope commitment is too easily falsified, the priority commitment is, conversely, unfalsifiable. If I can describe any cognitive or behavioral phenomenon of interest in terms of the maximization of reward, it becomes more difficult to test the hypothesis.

Second, 'grand unifying' theories of mind encourage us to recast broad swathes of empirical evidence into a single explanatory framework. However, the resulting explanations are sometimes less than illuminating. Moreover, surely some explanatory richness is lost if everything about the mind is ultimately, say, 'imagination,' 'attention,' or 'prediction-error minimization.' In some cases, these kinds of theories even run the risk of discounting evidence that is at odds with their theoretical commitments (Haas, 2021).

There's no reason to expect that the priority commitment would avoid such a fate. To try and keep to a fine-grained and falsifiable view, I thus stick with the weaker thesis.

5.3 'Reward is enough'[13]

Finally, let me draw out a few points of comparison between the evaluative thesis explored in this paper and the prominent and somewhat controversial 'reward is enough' (RIE) hypothesis (Silver, Singh, Precup, & Sutton, 2021). RIE holds that reward maximization is enough to "drive behavior that exhibits *most if not all* abilities that are studied in natural and artificial intelligence" (Silver, Singh, Precup, & Sutton, 2021, 1, added emphasis mine). Here, reward is understood in the sense put forward by the basic reinforcement learning framework introduced in Section 2.

Like the stronger thesis, RIE can be understood as involving a couple of different claims. First, RIE can be understood as the *epistemological* claim that reward maximization is enough to understand many - if not all - features of intelligence. Implicit in this claim is that reward maximization provides better and richer explanations than other rival scientific theories do. Second, RIE makes the *ontological* claim that intelligent processes *just are* reward maximization processes, where "intelligence, and its associated abilities, can be understood as subserving the maximization of reward by an agent acting in its environment" (Silver, Singh, Precup, & Sutton, 2021, 5). And third, RIE makes that *causal* claim that reward maximization is sufficient to *drive* the kinds of abilities we associate with behavior, such as gathering nuts or playing Go. According to this last claim, the forms of intelligence "implicitly emerge" through and as a direct result of the process of reward maximization. By extension, the authors contend, "a good reward-maximizing agent, in the service of achieving its goal, could implicitly yield all the abilities associated with intelligence that have been considered in natural and artificial intelligence" (Silver, Singh, Precup, & Sutton, 2021, 5).

[13] Thanks to Sean and Legassick and Hado van Hasselt for helpful discussions of the REI thesis.

What is the relationship between the evaluative thesis and RIE? At least on their face, the stronger thesis's priority commitment and RIE's epistemological claim appear consistent: the role of reward provides a unified and valuable way of understanding the mind and the nature of intelligence.

But the evaluative thesis and REI come apart on the ontological and causal fronts. At the end of the day, even the stronger thesis amounts to a pair of claims about the function and scope of a *cognitive proces*s in the mind. By contrast, RIE suggests that *all intelligence processing* is an expression or *byproduct* of reward maximization where, at bottom, the pursuit of reward drives the emergence of all other kinds of intelligence. These start to look like two very different kinds of arguments.

This being said, one softening feature of RIE is that it makes a pragmatic bet regarding the role of reward maximization in generating diverse forms of intelligence in artificial agents. That is, the authors of RIE propose that pure reinforcement learning frameworks will be sufficient to arrive at artificial general intelligence, without the need for handcrafting or pre-training. The authors acknowledge,

> We do not offer any *theoretical* guarantee on the sample efficiency of the reinforcement learning agent. Indeed, the rate at and degree to which abilities emerge will depend upon the specific environment, learning algorithm, and inductive biases; furthermore one may construct artificial environments in which learning will fail. Instead, we conjecture that the solution strategy of learning to maximize reward via interaction will be 'enough' for intelligence, and its associated abilities, to emerge in practice (Silver, Singh, Precup, & Sutton, 2021, 10).

In this sense, by adopting a kind of maker's approach (Craver, 2021), RIE is at least indirectly falsifiable through efforts to leverage reward maximization to design artificial intelligence.[14]

6. Conclusion

At the outset of this chapter, I proposed that RLDM is an instance of mind design so successful that we have not quite figured out what to do with it yet. I further argued that, in light of this success, we should move beyond characterizing the mind as exhaustively constituted by "thinking, intellect," as Haugeland originally put it, and begin to recognize its fundamentally evaluative nature. At the same time, I've sought to distinguish my view, which some philosophers may take to be remarkably strong, from even stronger views, which are more in line with views held by some in the machine learning and reinforcement learning literatures.

By way of conclusion, I want to briefly address what Haugeland called the common complaint about artificial intelligence, namely, that it cannot or may never achieve the rich interiority of everyday life, including "feelings, emotions, ego, imagination, moods, consciousness - the whole 'phenomenology' of an inner life. No matter how smart the machines become, there's still 'nobody home'" (this volume, p. x). Haugeland's characterisation is reminiscent of the traditional dichotomized conception of the mind: namely, of understanding the mind in

[14] Thanks to Neil Rabinowitz for pressing me on this point.

terms of ‘thinking’ and, well, ‘everything else’ - even if the ‘everything else’ includes a lot of the important processes.

The notion of valuation - normatively rich, empirically substantiated - allows us to put pressure on this type of traditional, dichotomized view. At a minimum, it challenges the idea that we can in good scientific conscience continue to group together phenomena as disparate as emotions, consciousness, and ego under the heading of ‘phenomenology.’ As noted above, with a notion of valuation in place, we can, for instance, start to work out the relationship and differences between valuation and the various philosophical theories of emotion, or the role of valuation in driving instances of imagination (Gershman, Zhou, & Kommers, 2017). Moreover, without in any way diminishing the ‘thinking’ or ‘computational’ mind, valuation brings with it new avenues for revising our extant philosophical and psychological cognitive taxonomies (Janssen, Klein, & Slors, 2017).

More broadly, the notion of valuation challenges our assumptions regarding which aspects of mind can or cannot be quantified - and thereby understood in properly scientific terms. For example, in their discussion of “intelligence” and “intelligent” processes, Silver and colleagues (2021) largely appeal to features of the conventionally thinking mind such as perception, language, and generalization. But what the foregoing discussion should show is that we can also appeal to the normative principles of RLDM to better decompose and understand those allegedly more ‘qualitative’ aspects of the mind such as valuation - and, by extension, our personal-level capacities such motivation, cognitive control, choice, and moral cognition.

We should also carry these insights forward into our ongoing efforts at mind design. That is, as we make advancements toward more sophisticated artificial and artificial general intelligence, we can enrich our understanding of the kinds of mental capacities that we can and should include in these efforts - and we should move past the idea of designing only ‘thinking’ machines in the traditional sense.

References

Altgassen, M., Kliegel, M., & Martin, M. (2009). Event-based prospective memory in depression: The impact of cue focality. *Cognition and Emotion*, *23*(6), 1041-1055.

Anderson, B. A., & Kim, H. (2018). Mechanisms of value-learning in the guidance of spatial attention. *Cognition*, *178*, 26-36.

Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M., & Pöppel, E. (2010). Dopaminergic reward system: a short integrative review. *International archives of medicine*, *3*, 24. https://doi.org/10.1186/1755-7682-3-24.

Arpaly, N., & Schroeder, T. (2014). *In praise of desire.* Oxford University Press.

Balcetis, E., Dunning, D., & Granot, Y. (2012). Subjective value determines initial dominance in binocular rivalry. *Journal of Experimental Social Psychology*, *48*(1), 122-129.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5), 407-419.

Blair, R. J. R. (2007). Empathic dysfunction in psychopathic individuals. *Empathy in mental illness*, *1*, 3-16.

Blair, R. J. R. (2018). Traits of empathy and anger: implications for psychopathy and other disorders associated with aggression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1744), 20170155.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual review of psychology*, *66*.

Briscoe, R., & Grush, R. (Summer 2020 Edition). Action-based theories of perception. *Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.), Available at https://plato.stanford.edu/archives/sum2020/entries/action-perception

Brook, M., & Kosson, D. S. (2013). Impaired cognitive empathy in criminal psychopathy: evidence from a laboratory measure of empathic accuracy. *Journal of abnormal psychology*, *122*(1), 156.

Chelazzi, L., Eštočinová, J., Calletti, R., Gerfo, E. L., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience*, *34*(25), 8594-8604.

Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 509-529.

Chen, S., Zhou, R., Cui, H., & Chen, X. (2013). Deficits in cue detection underlie event-based prospective memory impairment in major depression: an eye tracking study. *Psychiatry research*, *209*(3), 453-458.

Colombo, M. (2014). Deep and beautiful. The reward prediction error hypothesis of dopamine. *Studies in history and philosophy of science part C: Studies in history and philosophy of biological and biomedical sciences*, 45, 57-67.

Coundouris, S. P., Terrett, G., Laakso, L., Schweitzer, D., Kneebone, A., Rendell, P. G., & Henry, J. D. (2020). A meta-analytic review of prospection deficits in Parkinson's disease. *Neuroscience & Biobehavioral Reviews*, 108, 34-47.

Craver, C. (2021). Towards an epistemology of intervention: Optogenetics and maker's knowledge. *Tool Use in Neuroscience*, John Bickle, Carl F. Craver, & Ann-Sophie Barwich (Eds). Routledge.

Cubillo, A., Makwana, A. B., & Hare, T. A. (2019). Differential modulation of cognitive control networks by monetary reward and punishment. *Social cognitive and affective neuroscience*, *14*(3), 305-317.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, 8(12), 1704-1711.

Dayan, P. (2011). *Interactions Between Model-Free and Model-Based Reinforcement Learning.* Seminar Series from the Machine Learning Research Group. University of Sheffield, Sheffield. Lecture recording. <http://ml.dcs.shef.ac.uk/>.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems.* Computational Neuroscience Series.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology*, *18*(2), 185-196.

Dennett, D. C. (1969) *Content and Consciousness*. Routledge.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological science*, *20*(6), 778-784.

D'Iorio, A., Maggi, G., Vitale, C., Amboni, M., Di Meglio, D., Trojano, L., & Santangelo, G. (2019). Prospective memory in Parkinson's disease: the role of the motor subtypes. *Journal of neurology*, *266*(10), 2505-2511.

Domes, G., Hollerbach, P., Vohs, K., Mokros, A., & Habermeyer, E. (2013). Emotional empathy and psychopathy in offenders: An experimental study. *Journal of personality disorders*, *27*(1), 67-84.

Drayson, Z. (2014). The personal/subpersonal distinction. *Philosophy Compass*, *9*(5), 338-346.

Ergo, K., De Loof, E., & Verguts, T. (2020). Reward prediction error and declarative memory. *Trends in cognitive sciences*, *24*(5), 388-397.

Gagne, C., & Dayan, P. (2021). Peril, Prudence and Planning as Risk, Avoidance and Worry.

Gershman, S. J. (2015). A unifying probabilistic view of associative learning. *PLoS computational biology*, 11(11), e1004567.

Gershman, S. J., Zhou, J., & Kommers, C. (2017). Imaginative reinforcement learning: Computational principles and neural mechanisms. *Journal of cognitive neuroscience*, *29*(12), 2103-2113.

Glimcher, P. W. (2011). *Foundations of neuroeconomic analysis*. OUP USA.

Glimcher, P. W., & Fehr, E. (Eds.). (2013). *Neuroeconomics: Decision making and the brain*. Academic Press.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, 30, 535-574.

Haas, J. (2021). Can hierarchical predictive coding explain binocular rivalry?. *Philosophical Psychology*, *34*(3), 424-444.

Haas, J., & Klein, C. (2020). Holistic resource-rational analysis. *Behavioral and Brain Sciences*, *43*.

Hare, R. D. (1998). Psychopathy, affect and behavior. In *Psychopathy: Theory, research and implications for society* (pp. 105-137). Springer, Dordrecht.

Haugeland, J. (1979). Understanding natural language. *The Journal of Philosophy*, 76(11), 619-632.
Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in cognitive sciences*, 9(4), 188-194.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward guides vision when it's your thing: Trait reward-seeking in reward-mediated visual priming. *PloS one,* 5(11), e14087.

Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *Journal of neurophysiology*, 95(2), 567-584.

Hochstein, E. (2016). Giving up on convergence and autonomy: Why the theories of psychology and neuroscience are codependent as well as irreconcilable. *Studies in History and Philosophy of Science* Part A, 56, 135-144.

Hunter, L. E., & Daw, N. D. (2021). Context-sensitive valuation and learning. *Current Opinion in Behavioral Sciences*, 41, 122-127.

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of mood & anxiety disorders*, 3(1), 1-16.

Huys, Q. J., Daw, N. D., & Dayan, P. (2015). Depression: a decision-theoretic analysis. *Annual review of neuroscience*, *38*, 1-23.

Hyman, S. E. (2005). Addiction: a disease of learning and memory. *American Journal of Psychiatry*, *162*(8), 1414-1422.

Janssen, A., Klein, C., & Slors, M. (2017). What is a cognitive ontology, anyway? *Philosophical Explorations. 20:2, 123-128.*

Juechems, K., & Summerfield, C. (2019). Where does value come from?. *Trends in cognitive sciences*, 23(10), 836-850.

Katai, S., Maruyama, T., Hashimoto, T., & Ikeda, S. (2003). Event based and time based prospective memory in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(6), 704-709.

Kinsella, G. J., Pike, K. E., Cavuoto, M. G., & Lee, S. D. (2018). Mild cognitive impairment and prospective memory: Translating the evidence into neuropsychological practice. *The Clinical Neuropsychologist*, *32*(5), 960-980.

Kliegel, M., Phillips, L. H., Lemke, U., & Kopp, U. A. (2005). Planning and realisation of complex intentions in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(11), 1501-1505.

Kliegel, M., Altgassen, M., Hering, A., & Rose, N. S. (2011). A process-model based approach to prospective memory impairment in Parkinson's disease. *Neuropsychologia*, *49*(8), 2166-2177.

Krishnan, H. S., & Shapiro, S. (1999). Prospective and retrospective memory for intentions: A two-component approach. *Journal of Consumer Psychology*, *8*(2), 141-166.

Langdon, A. J., Sharpe, M. J., Schoenbaum, G., & Niv, Y. (2018). Model-based predictions for dopamine. *Current opinion in neurobiology*, 49, 1-7.

Li, Y. R., Weinborn, M., Loft, S., & Maybery, M. (2013). Patterns of prospective memory impairment among individuals with depression: The influence of cue type and delay interval. *Journal of the International Neuropsychological Society*, *19*(6), 718-722.

Liao, M. R., & Anderson, B. A. (2020). Reward learning biases the direction of saccades. *Cognition*, 196, 104145.

Mahmut, M. K., Homewood, J., & Stevenson, R. J. (2008). The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths?. *Journal of Research in Personality*, *42*(3), 679-692.

Marx, S., & Einhäuser, W. (2015). Reward modulates perception in binocular rivalry. *Journal of vision*, *15*(1), 11-11.

McFarland, C. P., & Vasterling, J. J. (2018). Prospective memory in depression: Review of an emerging field. *Archives of Clinical Neuropsychology*, *33*(7), 912-930.

McNamara, P., & Durso, R. (2018). The dopamine system, Parkinson's disease and language function. *Current opinion in behavioral sciences*, *21*, 1-5.

Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, *49*(1), 8-30.

Montague, R. (2006). *Why choose this book? How we make decisions*. Dutton.

Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the expected value of control. In *The 2nd multidisciplinary conference on reinforcement learning and decision making*.

Oba, T., Katahira, K., & Ohira, H. (2021). A learning mechanism shaping risk preferences and a preliminary test of its relationship with psychopathic traits. *Scientific reports*, *11*(1), 1-11.

Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & memory*, *24*(1), 65-69.

Pirogovsky, E., Woods, S. P., Filoteo, J. V., & Gilbert, P. E. (2012). Prospective memory deficits are associated with poorer everyday functioning in Parkinson's disease. *Journal of the International Neuropsychological Society*, *18*(6), 986-995.

Plassmann, H., O'doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of neuroscience*, *27*(37), 9984-9988.

Ramanan, S., & Kumar, D. (2013). Prospective memory in Parkinson's disease: a meta-analysis. *Journal of the International Neuropsychological Society*, *19*(10), 1109-1118.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews neuroscience*, 9(7), 545-556.

Redish, A. D. (2004). Addiction as a computational process gone awry. *Science*, *306*(5703), 1944-1947.

Redish, A. D. (2013). *The mind within the brain: How we make decisions and how those decisions go wrong*. Oxford University Press.

Redish, A. D., Jensen, S., & Johnson, A. (2008). A unified framework for addiction: vulnerabilities in the decision process. *The Behavioral and brain sciences*, *31*(4), 415.

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H. J., ... & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*, *24*(21), 2606-2611.

Rupprechter, S., Stankevicius, A., Huys, Q. J., Steele, J. D., & Seriès, P. (2018). Major depression impairs the use of reward values for decision-making. *Scientific reports*, *8*(1), 1-8.

Rupprechter, S., Stankevicius, A., Huys, Q. J., Seriès, P., & Steele, J. D. (2021). Abnormal reward valuation and event-related connectivity in unmedicated major depressive disorder. *Psychological medicine*, *51*(5), 795-803.

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2021). Neural evidence for the successor representation in choice evaluation. *bioRxiv*.

Savine, A. C., & Braver, T. S. (2010). Motivated cognitive control: reward incentives modulate preparatory neural activity during task-switching. *Journal of neuroscience,* 30 (31), 10294-10305.

Scarantino, A., & de Sousa, R. (2021). Emotion. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.). Available at https://plato.stanford.edu/archives/sum2021/entries/emotion

Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. *Physiological reviews*, 95(3), 853-951.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.

Schroeder, T. (2004). *Three Faces of Desire*. Oxford University Press.

Silver, D. (2015). *Reinforcement learning* [Video]. YouTube. Available at: https://www.youtube.com/watch?v=2pWv7GOvuf0

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence*, 103535.

Simonyan, K., Horwitz, B., & Jarvis, E. D. (2012). Dopamine regulation of human speech and bird song: a critical review. *Brain and language*, *122*(3), 142-150.

Soderstrom, H. (2003). Psychopathy as a disorder of empathy. *European child & adolescent psychiatry*, *12*(5), 249.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning*, 3(1), 9-44.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* MIT press.

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

Takahashi, T., Takagishi, H., Nishinaka, H., Makino, T., & Fukui, H. (2014). Neuroeconomics of psychopathy: Risk taking in probability discounting of gain and loss predicts psychopathy. *Neuroendocrinology Letters*, *35*(6), 510-517.

Takamura, M., Okamoto, Y., Okada, G., Toki, S., Yamamoto, T., Ichikawa, N., ... & Yamawaki, S. (2017). Patients with major depressive disorder exhibit reduced reward size coding in the striatum. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *79*, 317-323.

Waddell, S. (2013). Reinforcement signaling in Drosophila; dopamine does it all after all. *Current opinion in neurobiology*, *23*(3), 324-329.

Walter, S., & Meier, B. (2014). How important is importance for prospective memory? A review. *Frontiers in psychology*, *5*, 657.

Wilbertz, G., van Slooten, J., & Sterzer, P. (2014). Reinforcement of perceptual inference: Reward and punishment alter conscious visual perception during binocular rivalry. *Frontiers in psychology*, *5*, 1377.

Zabberoni, S., Carlesimo, G. A., Peppe, A., Caltagirone, C., & Costa, A. (2017). Does dopamine depletion trigger a spreader lexical-semantic activation in Parkinson's disease? evidence from a study based on word fluency tasks. *Parkinson's Disease*, *2017*.