

A Computational View on the Nature of Reward and Value in Anhedonia

Quentin J.M. Huys^{1,2}, Michael Browning^{3,4}

1 Division of Psychiatry and Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, UK 2 Camden and Islington NHS Foundation Trust, London, UK 3 Department of Psychiatry, University of Oxford, Oxford, UK 4 Oxford Health NHS Trust, Oxford, UK

Corresponding Information

Quentin JM Huys
Max-Planck UCL Centre for Computational Psychiatry and Ageing Research
Russell Square House, 10-12 Russell Square, WC1B 5EH, London, UK
Email: q.huys@ucl.ac.uk

Michael Browning
Department of Psychiatry,
Warneford Hospital,
Oxford, OX3 7JX, UK

Email: michael.browning@psych.ox.ac.uk

Word count

Abstract: 195
Total: 5362

Figures & Tables: 4
References: 66

ABSTRACT

Anhedonia—a common feature of depression—encompasses a reduction in the subjective experience and anticipation of rewarding events, and a reduction in the motivation to seek out such events. The presence of anhedonia often predicts or accompanies treatment resistance, and as such better interventions and treatments are important. Yet the mechanisms giving rise to anhedonia are not well-understood. In this chapter, we briefly review existing computational conceptualisations of anhedonia. We argue that they are mostly descriptive and fail to provide an explanatory account of why anhedonia may occur. Working within the framework of reinforcement learning, we examine two potential computational mechanisms that could give rise to anhedonic phenomena. First, we show how anhedonia can arise in multidimensional drive reduction settings through a trade-off between different rewards or needs. We then generalize this in terms of model-based value inference and identify a key role for associational belief structure. We close with a brief discussion of treatment implications of both of these conceptualisations. In summary, computational accounts of anhedonia have provided a useful descriptive framework. Recent advances in reinforcement learning suggest promising avenues by which the mechanisms underlying anhedonia may be teased apart, potentially motivating novel approaches to treatment.

CONTENTS

1 Introduction	2	5 Rewards on graphs	8
2 The Components of Anhedonia and Patient Experience	2	5.1 How cognitive distortions can reduce hedonic experience	9
3 Existing Computational Research	3	5.2 How the meaning of events may arise from associational graphs	10
4 Reward as Distance	5	6 Discussion	11
4.1 A General Effect of Unsatisfied Drives on Reward	7	6.1 Acknowledgements	12
4.2 The Geometry of Drive Space Influences Preference for Multidimensional Rewards	7	6.2 Disclosures	12
		References	12

1 INTRODUCTION

Anhedonia, broadly defined as a loss of interest or pleasure in rewarding activities (American Psychiatric Association, 2013), is a common feature of many psychiatric presentations (Husain and Roiser, 2018; Treadway and Zald, 2011; Trøstheim et al, 2020). Its association with adverse outcomes generally (Spjiker et al, 2001), and poor response to treatment specifically (Uher et al, 2012; McMakin et al, 2012), has motivated interest in the mechanistic processes that underlie it, with the ultimate aim of developing novel and more effective interventions.

While there have been many lines of enquiry investigating the mechanisms of anhedonia, a particularly popular approach has been to frame it as arising from abnormalities in reinforcement learning (RL) processes. Reinforcement learning considers how organisms use experiences of reward and punishment to motivate future actions and so has clear face validity when considering anhedonic experiences in which participants report reduced pleasure from, and a decreased motivation to complete, daily activities. Further, the development and characterisation of a set of formal models with demonstrable explanatory validity at a behavioural and neurobiological level suggests that RL is well placed to identify causally important targets for novel interventions.

In this chapter we will describe the clinical symptoms and experience of anhedonia and provide a basic overview of the logic of RL models and the processes they describe. We will then summarise the literature suggesting an association between abnormalities in RL processes and anhedonia, critically appraising what it does, and does not tell us about mechanism. To summarise our argument, we will suggest that previous work utilising an RL framework has provided consistent evidence that anhedonic patients behave "as if" rewarding outcomes were less valuable, but have provided little evidence as to why this might have occurred. We argue that, in order to understand why this occurs, it will be necessary to move beyond the largely descriptive RL models employed to date and develop mechanistic models with deeper explanatory ability. We then summarise recent work demonstrating that humans represent reward and derive value estimates in a more complex manner than suggested by the commonly used descriptive RL models. Lastly, we briefly consider some of the implications of this work on how anhedonia might be treated.

2 THE COMPONENTS OF ANHEDONIA AND PATIENT EXPERIENCE

The diagnostic manuals provide a broad definition of anhedonia as diminished interest or pleasure in activities (American Psychiatric Association, 2013; WHO, 1992). As has been noted previously (Husain and Roiser, 2018; Treadway and Zald, 2011), this definition encompasses a number of distinct components:

1. The subjective experience of pleasurable events (e.g. how much I enjoy an experience at the time)
2. The anticipation of future pleasurable events (e.g. how much I look forward to an enjoyable experience)
3. The motivation to engage in pleasurable activity (e.g. whether I act and expend effort to increase the likelihood of experiencing enjoyable experiences)

While a number of instruments have been developed to measure these components individually (Rizvi et al, 2016), they are clearly not independent (e.g. your motivation to engage in an activity will be influenced by how pleasurable you find it), suggesting that similar anhedonic presentations may arise from distinct mechanisms. To illustrate this point and motivate the formal description of RL models, which can be used to link the components together, Table 1 describes a variety of different mechanisms that might lead to the same anhedonic behaviour– not attending a party.

Before introducing the RL framework used to investigate anhedonia, it is important to highlight that, belying the simple descriptions of the components provided above, patients' experience of anhedonia

Mechanism	Mechanism in Terms of RL Model	Narrative
Reduced hedonic value	smaller r_{reward}	A woman is invited to a party. She attends the party and does not enjoy it. She does not attend the next party she is invited to as she does not want to repeat an unenjoyable experience.
Increased effort sensitivity	more negative r_{effort}	A man is invited to a party. He attends the party and enjoys it, but finds the experience exhausting. He does not have the energy to attend the next party he is invited to.
Reduced learning from past experience	Smaller model-free learning rate α	A woman attends a party and, against her expectation, enjoys it. When asked if she wants to attend another party, she declines, still feeling she will not enjoy it.
Asymmetric reward function	$\frac{\partial r}{\partial \text{work}} > \frac{\partial r}{\partial \text{social}}$ see Section 4	A man is invited to a party. He thinks he may enjoy it but performing well at work is more important to him at present. He declines the invitation, as it will not improve his work performance.
Reduced constructed value	low $\mathcal{V}(s)$ see Section 5	A woman is invited to a party. Although she likes the idea of a party, she thinks she will experience rejection and panic attacks if she attends and hence declines.

TABLE 1: Distinct causes of a single anhedonic behaviour.

(and of pleasure generally) are often far from simple. For example, a qualitative study of adolescents with depression (Watson et al, 2020) who were asked to describe their experience of anhedonia identified general themes of reduced joy and motivation, but also of a loss of connection with others and a questioning of their self and their purpose. This indicates that the "value" relevant to patients is often multi-faceted and future orientated, for example relating to an individual's social roles and their ultimate ability to fulfill these. As a consequence, any account of anhedonia that seeks to explain why patients treat potentially rewarding experiences as less valuable must be able to capture the influence of these complex valuation processes. We now provide a brief overview of the basics of RL models and how they have been applied to study mechanistic processes in anhedonia.

3 EXISTING COMPUTATIONAL RESEARCH

The motivation to engage in an activity rests on the judgement that it is worth the effort in the longer run. This depends on the rewards produced by the activity, its costs in terms of effort or associated punishments, and what the long term consequences of the activity are (e.g. will it lead to more or less reward in the future). The field of reinforcement learning (RL) is concerned with identifying optimal solutions to this difficult problem.

To be a little more precise, imagine that an individual starts in state, s , one of the set of all possible states, \mathcal{S} , each of which is associated with some reward r_s . The individual is able to perform one action, a , from those available in that state, \mathcal{A}_s , with the action resulting in a change of the state to s' and the associated reward $r_{s'}$. RL algorithms approach the problem of what action, a , to take by estimating the value, $\mathcal{V}(s)$, of the states– the long-term sum of future rewards that will be experienced if the state is visited (Sutton and Barto, 2017).

One of the simplest algorithms (Rescorla and Wagner, 1972) recursively estimates the value of states simply through experienced outcomes. A variant of this algorithm frequently used in studies of anhedonia is:

$$\mathcal{V}_{t+1}(s) = \mathcal{V}_t(s) + \alpha(\rho r_t - \mathcal{V}_t(s)) \quad (1)$$

Here, value is updated by the reward prediction error, the difference between the experienced reward, r_t (NB, which may include a negative punishment/effort term as well as a positive reward term), and the estimated value, \mathcal{V}_t . Two parameters influence the algorithm’s behaviour and allow it be fitted to participant data; a learning rate term, α , which controls how quickly the algorithm updates its value, and a reward sensitivity term, ρ , which changes the effective magnitude of the experienced rewards. While this algorithm provides a straightforward account of reward learning that links experienced rewards to estimated values, it is important to note that it does not account for a number of important processes that are likely to influence human learning. In particular:

- a The algorithm has no principled way of allowing the internal state of an individual to influence reward magnitude. For example, drinking water is generally more rewarding when one is thirsty than when one is not. The only way this can be accommodated in the algorithm is by engineering separate states for "thirsty" and "not thirsty" and simply telling the algorithm that the reward for drinking water is higher in the first state. In other words, the algorithm can describe how learning might proceed if the reward associated with drinking was reduced, but not explain how or why it occurs.
- b The algorithm learns the association between immediate reward and a state, but ignores the values of all subsequent states. For example, as exercise is effortful it would find no reason to do it, even if it leads to the valuable future state of being healthy. As described above, the value of a state is defined as the long-term reward expected from visiting that state. This algorithm ignores all future reward.

The most common behavioural observation from empirical studies which have employed this class of algorithm in depressed and anhedonic individuals is a cross-sectional association between anhedonia and reduced reward sensitivity ρ or equivalently noisiness of exploratory choice (Steele et al, 2007; Chase et al, 2010; Kunisato et al, 2012; Blanco et al, 2013; Huys et al, 2013; Robinson and Chase, 2017), with some evidence of increased punishment sensitivity (Beevers et al, 2013; Herzallah et al, 2013). An association between anhedonia and behavioural estimates of learning rate, α , has been reported (Chase et al, 2010), although much less commonly.

As the ρ parameter in Equation 1 acts to scale the magnitude of experienced reward, one interpretation of this finding is that anhedonic individuals have a generalised reduced hedonic response to the sensory experience of rewarding events. If correct, this would predict that such individuals should provide a lower rating of the hedonic value of all primary rewards in the absence of learning. This question has been addressed in a number of studies, most of which have utilised olfactory or gustatory stimuli. Overall, this literature consistently finds that symptoms of depression and anhedonia are associated with a decreased ability to discriminate between different sensory stimuli (Colle et al, 2020; Atanasova et al, 2010; Pause et al, 2001; Lombion-Pouthier et al, 2006; Kohli et al, 2016; Amsterdam et al, 1987; Berlin et al, 1998), but have not consistently found that the hedonic value of the stimuli are attenuated (Colle et al, 2020; Atanasova et al, 2010; Pause et al, 2001; Naudin et al, 2012; Amsterdam et al, 1987; Berlin et al, 1998; Clepce et al, 2010; Swiecicki et al, 2009; Dichter et al, 2010).

A related prediction is that anhedonic individuals should show a reduced neural response to the experience of rewarding stimuli, again in the absence of learning. Although this effect has been observed in depressed patients (Pizzagalli et al, 2009), the relationship is more frequently found with anticipation of stimuli, i.e. with the value or the prediction error (Kumar et al, 2008; Knutson et al, 2008; Geugies et al, 2019; Greenberg et al, 2015; Eckstrand et al, 2019; Ruppachter et al, 2020; Pizzagalli et al, 2020). However, one large study using a task with no learning did not find an effect of either prediction error or outcome on neural activity (Rutledge et al, 2017).

Overall, the work described above suggests that it is difficult to account for the lower ρ parameter observed in RL studies of anhedonic individuals simply as the result of a reduced hedonic response to outcomes. What then might be causing this effect? One possibility is that the radically simple RL model described in Equation 1 is mis-specified, it ignores important processes that lead to anhedonia and can only account for the effect of these processes in its ρ parameter. In other words, anhedonic individuals behave "as if" rewarding events are of lower magnitude, but to understand why they behave this way we must consider in more detail how rewards are constructed and used to estimate value. Indeed, a more reliable reduction of the hedonic impact of both positive and negative items can be observed in response to more complex stimuli, such as visual images or films (Bylsma et al, 2008), where any 'reward' has to be constructed or inferred. In the next sections we discuss ways in which these processes may be incorporated into RL models and describe how their perturbations may account for anhedonic behaviour.

4 REWARD AS DISTANCE

As we have seen above, a number of empirical studies have suggested that, when learning about the values of actions, people with anhedonia seem to treat rewarding outcomes as if they were of a smaller magnitude than people without anhedonia. When thinking about why this might occur, one important question is what determines how rewarding an event actually is (or should be). In this section we consider the implications of two basic observations relevant to this question:

- a The reward experienced following an event is influenced by your current state– drinking water is more rewarding if you are thirsty than if you are not.
- b Distinct aspects of your current state can independently influence the reward associated with an event– If I am both hungry and thirsty, eating will reduce my hunger, and will also reduce how rewarding future food is, but will not influence how rewarding water is.

These observations underpin "drive-reduction" theories of motivation (Figure 1) which suggest that the reward experienced following an event is a function of the degree to which the event moves you towards a homeostatic "set point" (Hull, 1943). While a number of distinct drive-reduction theories have been proposed (Berridge, 2004; Bolles, 1980; Juechems and Summerfield, 2019; Keramati and Gutkin, 2014; Berridge, 2012), their common element is that a multi-dimensional "drive space" can be defined with distinct dimensions representing different basic drives (e.g. hunger, thirst). An individual's current state is represented by a position in this space (i.e. with a thirst of "x" and a hunger of "y"). The space also includes a homeostatic set-point, the point at which all of the individual's drives are fulfilled. Different events (e.g. drinking water) move the individual from their current position to a new position in drive space, with the reward experienced by the individual following the event being equivalent to the reduction in distance between their location and the set point. When first conceived, these drive-reduction theories were motivated by physiological homeostasis (i.e. regulating temperature, hydration etc.; Hull 1943), although more recent iterations have suggested that a similar approach might be applied to more abstract goals, such as maintaining social status, with clear relevance to subjective experiences such as anhedonia (Juechems and Summerfield, 2019; Keramati and Gutkin, 2014).

Formally (Keramati and Gutkin, 2014), if an individual's current position at time t in an N dimensional drive space is defined as $H_{(t)} = (h_{1,t}, h_{2,t}, h_{3,t}, \dots, h_{N,t})$ and the position of their set point is defined as $H^* = (h_1^*, h_2^*, h_3^*, \dots, h_N^*)$, then the distance from the current position can be described as:

$$D(H_t) = \left(\sum_{n=1}^N |h_n^* - h_{n,t}|^q \right)^{1/p} \quad (2)$$

The parameters p and q influence the geometry of the drive space, with $p = q = 2$ defining a Euclidean

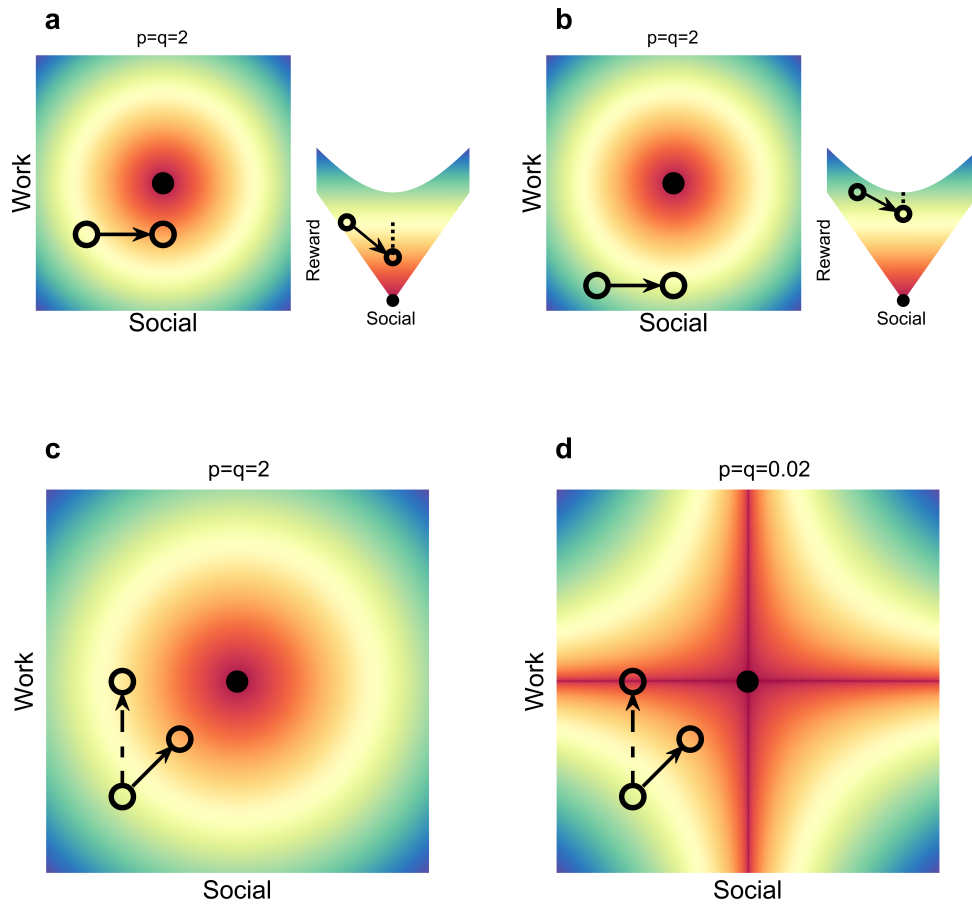


FIGURE 1: Reward as a reduction in drive. Each panel illustrates a two dimensional drive space, with social function being represented along the x axis, and work performance along the y axis. The black circle in the centre of each plot is the set-point at which all drives are satisfied. Distance from the set-point is represented by the colour of the plot and is determined by the geometry of the drive space (equation 2). Events experienced by an individual will change their position in drive space, for example the arrows connecting the two circles in panels **a** and **b** illustrate the effect of attending a party which specifically changes an individual's position on the social dimension. The reward associated with an event is defined as the reduction in distance to the set-point produced by a change in position (Keramati and Gutkin, 2014). Two situations in which the reward experienced by attending a party may be reduced are illustrated. The first situation is shown in panels **a** and **b** which are identical other than the starting position of the points on the y-axis– in panel **b** the individual is less content with his work situation than in panel **a**. This has the effect of reducing the reward experienced from attending a party (the insets of both panels show the same drive space rotated by 90° around the x axis so that the reward associated with attending the party is represented by the dotted line). The second situation is illustrated in panels **c** and **d**. These panels illustrate the relative effect of an event that improves both work and social drive (solid lines, e.g. attending a work party, which moves +1 in both axes) and an event that only satisfies the work drive (dashed line, e.g. working longer, which moves +2 along the work dimension). Panel **c** illustrates the change when the drive space is Euclidean ($p = q = 2$) and the work party produces higher reward. In panel **d** the drive space is warped ($p = q = 0.02$) resulting in longer work producing more reward. In general as p and q increase, multi-dimensional reductions in drives produce increasingly greater rewards than uni-dimensional.

space (some implications of different values of these parameters are discussed below). If an event occurs to change the individual’s position in drive space, the reward experienced as a result is defined as:

$$r_{t+1} = D(H_{t+1}) - D(H_t) \quad (3)$$

The important property of this formulation for our argument is that, by providing an account of how reward is generated, it illustrates situations when r , the reward associated with an event, should be relatively higher or lower. In other words, it describes situations in which apparently anhedonic behaviour is expected and thus potentially provides insights into why individuals may exhibit this behaviour. We describe two specific examples below.

4.1 A GENERAL EFFECT OF UNSATISFIED DRIVES ON REWARD

It is more difficult to enjoy a party if you are constantly worrying about your job than if you are not. At first glance, this observation seems trivial, however if the party has nothing to do with your job, then why should work worry make it less enjoyable?

Figure 1a-b illustrates why this might occur (Keramati and Gutkin, 2014). It shows an example two dimensional drive space in which the x dimension encodes social satisfaction and the y dimension encodes work functioning. The filled black circle in the centre represents the set point (the ideal level of both the work and social dimensions). The distance to this set point is represented by the colour of the plot (for this example the space is Euclidean, with the parameters of equation 2 being $p = q = 2$). In this setting, attending a party with friends moves one along the x axis of drive space– it will make one feel less lonely, but wouldn’t improve work performance. The effect of attending a party for someone who is a bit lonely and not very worried about work is illustrated in Figure 1a by the arrow moving the position to the right. The reward associated with attending the party is the reduction in distance to the set point from the start to the end of the arrow and is illustrated by the change in colour (and by the dotted line in the inset). Figure 1b illustrates exactly the same situation for an individual who is more worried about their work (the points are lower on the y axis). As can be seen, even though this individual starts out just as lonely before the party, and the party has an identical effect of reducing this loneliness, the increased worry about work results in a smaller reward associated with attending the party, an effect that is present for all drive geometries that satisfy $p = q > 1$.

4.2 THE GEOMETRY OF DRIVE SPACE INFLUENCES PREFERENCE FOR MULTIDIMENSIONAL REWARDS

In reality, a single event will often be associated with different types of reward (and/or cost), such as a work party that improves both your work functioning and social satisfaction (at least to some extent). Drive reduction accounts of reward are able to capture this sort of multidimensional effect well (Juechems and Summerfield, 2019) and to compare it to simpler effects which change just one dimension of drive. Interestingly, the geometry of the drive space influences this comparison in a way that may be relevant to anhedonia.

Figure 1c-d illustrates the situation where an individual is equally far from their set-point on both the social and work axes. Imagine they have to decide whether to attend a work party (which will improve both dimensions by +1, illustrated by the solid arrow in both panels) or just stay at work longer (which will improve work by +2 but not influence social satisfaction, illustrated by the dashed arrows). The reward associated with these choices is, as per equation 3, the reduction of distance to the set-point. Figure 1c illustrates the case when the space is Euclidean, with $p = q = 2$. Here the multi-dimensional work party brings the individual closer to the set-point than the uni-dimensional effect of working longer, and thus results in a larger reward. However, as both the p and q parameters reduce, the relative advantage of the multi-dimensional option drops, with Figure 1d illustrating an extreme case where $p = q = 0.02$ and

staying at work is substantially more rewarding than attending the party. Overall, as p and q increase, a multi-dimensional "bundle" of effects on different drives will increasingly produce larger rewards than uni-dimensional effects, with $p = q = 1$ being the indifference point at which the two effects produce equivalent reward.

Drive reduction effects may be incorporated into RL accounts by substituting the experienced reward term, r , from algorithms such as that described by Equation 1 with the function described by Equation 3 (Keramati and Gutkin, 2014; Juechems and Summerfield, 2019; Zhang et al, 2009). However, as described above, a second process ignored by Equation 1 is the value of future states. In the next section we consider a more general framework that incorporates the impact of estimated future states on experienced reward.

5 REWARDS ON GRAPHS

The geometrical view of rewards as distances is particularly useful for homeostatic situations. However, distances are harder to define in more general, abstract spaces, where the geometry is not necessarily pre-defined by physiologically existing imperatives. Furthermore, the criticism directed against the Rescorla-Wagner model, that it is largely descriptive rather than explanatory, is not completely resolved by simply replacing the r term with the output from a multi-dimensional drive space. For example, this output depends on the geometry of the drive space, but what determines this geometry? We now turn to RL on graphs as one way to define distances in more complex spaces and when considering more general goals. Briefly, RL provides a potentially general framework for thinking about how anhedonia may arise when important goals cannot be achieved by providing algorithms to measure progress towards goals that are sensitive to individuals' beliefs and priorities.

At the heart of these models is the value $\mathcal{V}(s)$ of states. It is akin to the distance function in equation 2, and can be used to analogously quantify the progress towards a goal in terms of a change in value when going from one state to another:

$$\delta_{\mathcal{V}} = \mathcal{V}(s') - \mathcal{V}(s) \quad (4)$$

We note that $\delta_{\mathcal{V}}$ takes on a role similar to the reward r in equation 3. Readers familiar with RL may note that it is similar to the temporal difference prediction error δ , but is missing the r term. We omitted it here to emphasize that actions in most states do not lead to immediate primary rewards r , but instead lead to a change in state. Hence, the prediction errors are mostly driven by differences in value. We now examine whether the term $\delta_{\mathcal{V}}$ may provide fruitful insights into the origins of anhedonia.

We first illustrate how $\mathcal{V}(s)$ can capture effects akin to the distance in equation 2, and then discuss how it is more general and encompasses a broader set of findings. $\mathcal{V}(s)$ is defined as the long term reward expected from visiting the state. In the grid in Fig. 2A, the goal is the red circle, providing an actual reward r , whereas no actual reward is obtained anywhere else. In this case, the value $\mathcal{V}(s)$ is related to how rapidly the goal will usually be reached from each state s to the goal state. Factors reducing this are when transitions are not fully controllable and hence at least partially random, and when only unidirectional moves along the axes are possible (black vs grey connections). In most situations, as in Fig. 1, moving along the social dimension results in a steeper increase in value (a higher $\delta_{\mathcal{V}}$) for the green path than the yellow path (Fig. 2B,C,E, bottom panels). Hence, assumptions about how feasible it is to move between states affect the perceived distance and hence value of each state. If hedonic experience is related to $\delta_{\mathcal{V}}$, this qualitatively replicates the examples illustrated in Figure 1, whereby an event may be more or less 'rewarding' depending on the presence of other competing priorities or goals.

However, this formulation of values on a graph goes beyond the reward as distance view in two ways. First, it allows us to formalize complex relationships to goals. For instance, it may not always be possible to simply choose actions that bring us straight towards a goal. If certain moves in the grid are not allowed (Fig. 2F,G), a distance as in equation 2 may no longer be a good guide. RL allows us to measure distance in more complex graphs and to integrate constraints in how individuals believe they can move in this homeostatic or goal-related space.

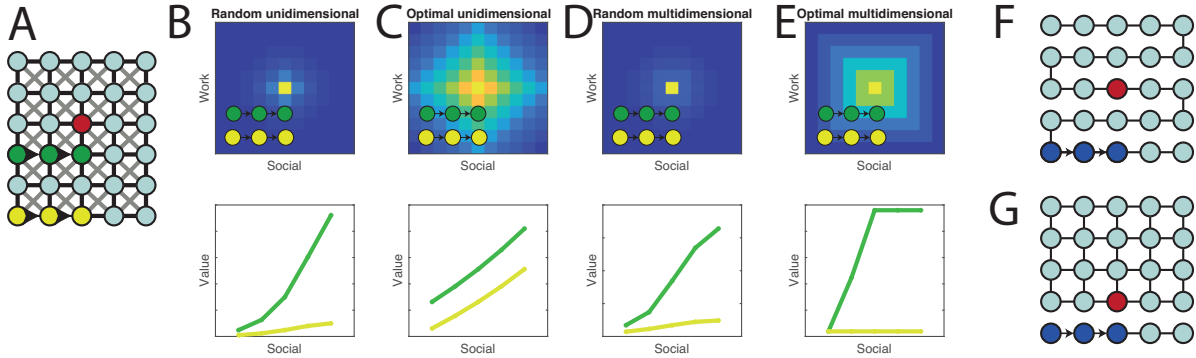


FIGURE 2: Goal distance in RL. **A:** A grid of states. The agent considers moving along the green or yellow path. The goal, or reward, can be retrieved in the red state. The green and yellow moves are equivalent to those in Figs. 1a and b. Note this does not define a Euclidian distance as in Fig. 1, but this can be approximated by setting the transition probabilities accordingly. **B-E, top panels:** Heatmaps showing the value, the long-term expected reward, for being in each of the states as derived from the Bellman equation (Bellman, 1957). The effects of different allowed transitions between states, i.e. when movement is only allowed along the axes (“unidimensional”, dark transitions in A), or also diagonally (“multidimensional”, grey transitions in A), and different action policies (i.e. “random”, in which the agent selects an action at random and “optimal” in which the agent always selects the action that increases value the most) are shown. For illustration purposes, the space is more finely quantized than suggested in A. **B-E, bottom panels:** The value of states along the green and yellow trajectories. The value generally increases more rapidly for the green than for the yellow trajectory (shown in green and yellow, respectively). The experienced reward is the change in value, represented by the slope of the lines. As can be seen, the relative advantage of the green trajectory is influenced by the available transitions and the agent’s policy (note that similar effects were produced by changes in the geometry of the drive-space in section 4). **F:** If not all state transitions are believed to be available, a previously rewarding action may result in reduced value and thus be punishing. Here, for instance the lack of transitions mean that improving social functioning by moving to the right actually takes the agent further from the goal and will result in reduced value (and thus punishment). **G:** In the extreme case, where there is no path to the goal at all, the agent is helpless and no action will produce reward.

Second, estimates of the long-term value $\mathcal{V}(s)$ depend on two terms: the immediate reward experienced in the current state, and all future (discounted) rewards from subsequently visited states. The future component has to be estimated. Model-free algorithms estimate it from past experienced rewards, while model-based algorithms derive it from a type of internal simulation according to internally held beliefs about what might happen in the future (Daw et al, 2005; Sutton and Barto, 2017). The latter, model-based process, allows individual beliefs to influence the value estimates. Equation 4 formally describes how beliefs about future events will impact the hedonic impact of current experience. An individual’s belief structure can be viewed as a graph, with reward being obtained at the goal state. In this graph, beliefs about the links between future states are represented as the edges (connections). These edges control the long term reward expected from, and hence the hedonic impact of visiting a state. We now illustrate this with two examples.

5.1 HOW COGNITIVE DISTORTIONS CAN REDUCE HEDONIC EXPERIENCE

Hopelessness and helplessness theories of depression (Seligman and Maier, 1967; Maier and Watkins, 2005; Alloy et al, 1999) emphasize the importance of beliefs about the inachievability of important goals and a perceived lack of control. Such beliefs can be captured formally in terms of either graph structure,

or the ability to navigate effectively within the graph structure (Huys and Dayan, 2009). Consider first Fig. 2F. In this graph, the person believes that there is no path from their state (blue) to the goal (red). Here, engaging in a social activity would bring the individual 'closer' to the goal in a simple distance sense, but it would have zero hedonic impact because there is no path to the goal: none of the states the person believes they can visit are linked to the goal. A related effect is obtained if the individual doubts their own ability to select the best action, or in the effect of that action. In this case, the selection of an action will not reliably cause transition to a state closer to the goal, which effectively flattens out the value function, and thereby reduces the hedonic δ_V signal.

Hence, helpless or hopeless structures of an individual's belief graph can result in profound changes to the perceived hedonic impact δ_V of events.

5.2 HOW THE MEANING OF EVENTS MAY ARISE FROM ASSOCIATIONAL GRAPHS

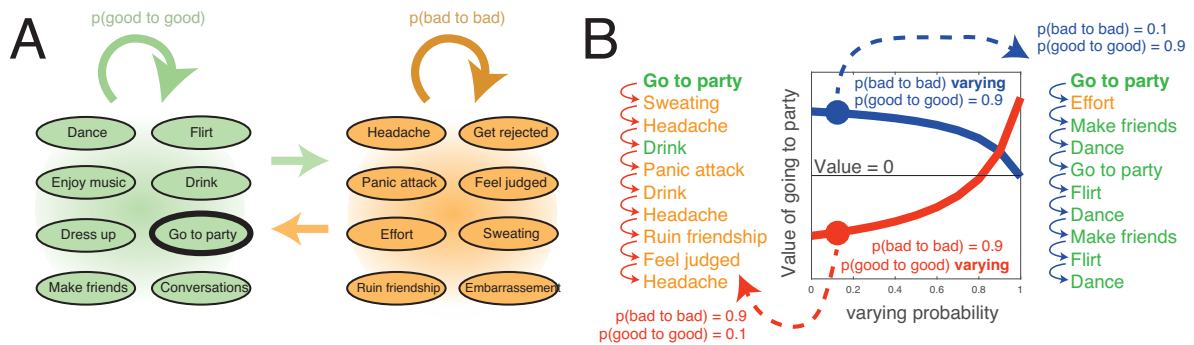


FIGURE 3: Values in thought association space. **A**: Grid of semantically associated states. One subset of states is roughly positive (green), the other is roughly negative (orange). We parametrize individual's internal belief models in terms of two parameters. The first parameter $p(\text{good to good})$ is the probability of going from one positive to another positive state, i.e. to sample a positive thought after having sampled a positive thought. The second parameter is the probability of sticking with bad thoughts ($p(\text{bad to bad})$). The probability of transitioning from from good to bad is simply $1 - p(\text{good to good})$, and $1 - p(\text{bad to bad})$ for bad to good. **B**: The sequence of words shows two random sequences drawn from this associative model. For a high bad-to-bad transition probability and a low good-to-good transition probability, a sequence starting in the 'go to party' state rapidly transitions to the bad states and stays there most of the time. The converse is true for a high good-to-good and low bad-to-bad transition probability (sequence of words on the right). The graph in the middle shows the value assigned to the 'go to party' state as a function of varying transition probabilities. We assume that the rewards associated with the green and orange states are random positive and negative numbers, respectively. The value is then the average over sequences experienced when starting from the 'go to party' state. Critically, although the reward associated with the 'go to party' thought is always positive, its value can vary widely, from strongly positive to strongly negative, depending on the nodes it is associated with. The red line in the middle panel shows the value of 'go to party' for a high bad-to-bad probability. Here, the value is only positive if the good-to-good probability is even higher. This would be a striking state, where positive and negative thoughts strongly cue each other, with little possibility of sampling a thought from the opposite valence. The blue line shows the value when the good-to-good probability is high. The value is reduced as the bad-to-bad probability increases. Note that the fact that it does not cross the zero value line here is due to random sampling of the rewards - on other simulations where the rewards were sampled less large the line did cross into negative territory.

Instruments that assess anhedonia require individuals to report how much pleasure or enjoyment they

would have derived from certain events or activities, or how much they anticipate or engage in them. This requires individuals to internally instantiate or simulate the events, and then engage in a process that allows them to estimate the required quantities such as pleasure or anticipation.

We now consider this process in terms of reinforcement learning on a graph.

Graphs provide a way to formulate how individuals believe events are related. This in turn allows events to be valued in relationship to other valued events. Consider a question akin to those typically asked in an anhedonia questionnaire "How much would you enjoy going to a party?". When thinking about this, one person might imagine getting dressed, looking good, going to dance, enjoying the music, flirting, conversing and making friends. Going to a party is hence not only potentially appetitive in itself, but is also associated with positive events, and so the person might respond that they would quite enjoy going to the party. A different person might instead imagine the effort associated with going to a party, having panic attacks on the way there, feeling rejected, sweating and even ruining any existing relationships. Even if, the actual experience of the party was enjoyable, this person might still respond that they would not enjoy the prospect of the party. Hence, as illustrated in Figure 3, even if both individuals would enjoy the actual party equally when they were there, the different associations between the party and other events would result in different assessments of its hedonic value and different motivations to attend. In other words, how nice it is to go to a party is different for different individuals because it is related, in their minds, to other events. This may be one approach to capturing how the 'meaning' of events differs for individuals as a function of their beliefs (c.f. Jackson et al 2019).

A few points about this are worth noting. First, the estimate of a value $\mathcal{V}(s)$ can be arrived at in different ways. The literature review above discussed prediction error learning from experience as one way to arrive at it. Alternatively, the conceptually simplest algorithm to estimate $\mathcal{V}(s)$ is to explicitly consider ('think through') all possible future consequences, weighing them by their probability, and computing the expectation. This requires access to an internal model of what is likely to happen. This is often very demanding because the future holds many possibilities (Daw et al, 2005). As a result, not all possible scenarios can be evaluated and thus the outcome of the evaluation is highly sensitive to the subset of scenarios selected for evaluation (Huys et al, 2012; Huys and Renz, 2017). An approximation to an exhaustive evaluation can be arrived at by sampling from the network according to the transition probabilities, e.g. generating lots of sequences as in Fig. 3B, and then averaging over their values. There is substantial behavioural evidence that humans do engage in such sample-based evaluation, and that it is modulated by attentional processes (Krajbich, 2019). A bias in the sampling process can hence effectively mimic an alteration in the underlying model samples are drawn from (Huys and Renz, 2017), and even with high bad to bad transition probabilities, and low good to good transition probabilities, sometimes the sampled associations may be positive. Such an iterative sequential sampling process may be slow, whereas hedonic assessments can often be very rapid. An interesting possibility comes from Successor representations (Dayan, 1993; Russek et al, 2017), where estimates of *what* is likely to happen, i.e. which states are likely to succeed which other states, are constructed and stored. Values can then be derived rapidly and without sampling by weighing these predictions about the future by how rewarding they are. Such successor representations could allow for very rapid evaluations without having to consider all options repeatedly or sampling.

6 DISCUSSION

We first argued that existing behavioural data suggests that anhedonia is associated with a reduction in reward sensitivity ρ . However, the simple RL models employed in this work can only accommodate changed reward sensitivities by postulating a change in the underlying state. In the absence of a cogent theory linking rewards to particular states, these models therefore provide a purely descriptive account of anhedonia. Second, we argued that the reduction of physiological drives is one area where there is a cogent theory that links specific states to reward changes. Indeed, this provides some interesting insights into how anhedonia-like phenomena may arise not from abnormalities in one reward system, but from a

trade-off or balance between different types of rewards or needs. Finally, we argued that reinforcement learning on graphs provides a useful generalization of the drive reduction account to complex spaces. We argued that the underlying graph is related to individuals' belief structure which defines the distances between states, and thereby determines the estimated long term value of visiting a state. We have briefly outlined how this may be related to hopelessness, future expectations and meaning.

These considerations have a number of implications regarding the treatment of anhedonia. The framing of reward as a drive-reduction process indicates that the belief (correct or otherwise) that one is doing particularly badly in one aspect of life will reduce the experienced reward from all other aspects. As a consequence, when attempting to treat anhedonia it may be important not just to focus on the particular activity an individual is not enjoying (e.g. going to a party), but rather on those aspects of life in general that they feel are least satisfactory (e.g. work). Similarly, the effect of the geometry of an individual's drive space on preference for multi-, relative to uni-, dimensional drive reduction effects may lead certain individuals to forgo rewards that occur in multi-dimensional bundles. Identifying, and potentially modifying, such "black and white" propensities during treatment may help patients alter maladaptive behavioural policies and increase the range of rewarding activities they engage with. More generally, it will be important to understand how existing pharmacological, psychological and social treatments influence the trade-off between rewards. Characterising the impact of these treatments on measures of choice preference for uni- relative to multi-dimensional rewards (Juechems et al, 2019) may be a useful initial step in this task. An understanding of hedonic experience as arising, at least in part, from expectations of future reward conditioned on belief about the associative structure of the world suggests that anhedonic symptoms may be seen as expected consequences of cognitive distortions, rather than as causes of those distortions. This highlights the importance of recent innovations in cognitive therapy to increase the focus on rewarding experience (Craske et al, 2019; Dunn et al, 2019; Geschwind et al, 2019).

A notable omission from our discussion is an explicit consideration of effort, which has been linked to response vigour in average reward RL models (Niv et al, 2007) and associated with symptoms of apathy (Nair et al, 2020). While space limitations prevent an extensive discussion we note that viewing effort as negative reward or cost (Table 1), suggests that it may be incorporated in the drive-reduction view as movement away from the set-point and, as we illustrate in section 5, it may also affect event value through associational belief networks (Fig. 3).

In summary, we suggest that it is important to move beyond descriptive to explanatory models of anhedonia and have argued that reinforcement learning provides a framework for understanding how different types of reward might interact and how beliefs can determine individual hedonic experiences.

6.1 ACKNOWLEDGEMENTS

QH is supported by the UCLH Biomedical Research Centre. MB is supported by the NIHR Oxford Health Biomedical Research Centre and the NIHR Oxford cognitive health Clinical Research Facility. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the UK Department of Health.

6.2 DISCLOSURES

QH has received options and consultancy fees from Aya Health, and a research grant from Koa Health. MB has acted as a consultant for Janssen Research, P1vital Ltd and CHDR, owns shares in P1vital Products Ltd and was previously a paid employee of P1vital Ltd.

REFERENCES

Alloy LB, Abramson LY, Whitehouse WG, Hogan ME, Tashman NA, Steinberg DL, Rose DT, Donovan P (1999) Depressogenic cognitive styles: predictive validity, information processing and personality

- characteristics, and developmental origins. *Behav Res Ther* 37(6):503–531
- American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders: DSM-5*, 5th edn. Autor, Washington, DC
- Amsterdam JD, Settle RG, Doty RL, Abelman E, Winokur A (1987) Taste and smell perception in depression. *Biol Psychiatry* 22(12):1481–1485
- Atanasova B, El-Hage W, Chabanet C, Gaillard P, Belzung C, Camus V (2010) Olfactory anhedonia and negative olfactory alliesthesia in depressed patients. *Psychiatry Research* 176(2-3):190–196, DOI 10.1016/j.psychres.2008.11.016
- Beevers CG, Worthy DA, Gorlick MA, Nix B, Chotibut T, Maddox WT (2013) Influence of depression symptoms on history-independent reward and punishment processing. *Psychiatry research* 207:53–60, DOI 10.1016/j.psychres.2012.09.054
- Bellman RE (1957) *Dynamic Programming*. Princeton University Press
- Berlin I, Givry-Steiner L, Lecrubier Y, Puech AJ (1998) Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects. *Eur Psychiatry* 13(6):303–309, DOI 10.1016/S0924-9338(98)80048-5, URL [http://dx.doi.org/10.1016/S0924-9338\(98\)80048-5](http://dx.doi.org/10.1016/S0924-9338(98)80048-5)
- Berridge KC (2004) Motivation concepts in behavioral neuroscience. *Physiology & Behavior* 81(2):179–209, DOI 10.1016/j.physbeh.2004.02.004, URL <https://www.sciencedirect.com/science/article/pii/S0031938404000435>
- Berridge KC (2012) From prediction error to incentive salience: mesolimbic computation of reward motivation. *The European Journal of Neuroscience* 35(7):1124–1143, DOI 10.1111/j.1460-9568.2012.07990.x
- Blanco NJ, Otto AR, Maddox WT, Beevers CG, Love BC (2013) The influence of depression symptoms on exploratory decision-making. *Cognition* 129:563–568, DOI 10.1016/j.cognition.2013.08.018
- Bolles RC (1980) Some functionalist thoughts about regulation. In: *Analysis of motivational processes*, NY: Academic Press, New York, pp 63–75
- Bylsma LM, Morris BH, Rottenberg J (2008) A meta-analysis of emotional reactivity in major depressive disorder. *Clin Psychol Rev* 28(4):676–691, DOI 10.1016/j.cpr.2007.10.001, URL <http://dx.doi.org/10.1016/j.cpr.2007.10.001>
- Chase HW, Michael A, Bullmore ET, Sahakian BJ, Robbins TW (2010) Paradoxical enhancement of choice reaction time performance in patients with major depression. *J Psychopharmacol* 24(4):471–479, DOI 10.1177/0269881109104883, URL <http://dx.doi.org/10.1177/0269881109104883>
- Clepece M, Gossler A, Reich K, Kornhuber J, Thuerauf N (2010) The relation between depression, anhedonia and olfactory hedonic estimates—a pilot study in major depression. *Neurosci Lett* 471(3):139–143, DOI 10.1016/j.neulet.2010.01.027, URL <http://dx.doi.org/10.1016/j.neulet.2010.01.027>
- Colle R, El Asmar K, Verstuyft C, Lledo PM, Lazarini F, Chappell K, Deflesselle E, Ait Tayeb AEK, Falissard B, Duron E, Rotenberg S, Costemale-Lacoste JF, David DJ, Gressier F, Gardier AM, Hummel T, Becquemont L, Corruble E (2020) The olfactory deficits of depressed patients are restored after remission with venlafaxine treatment. *Psychological Medicine* pp 1–9, DOI 10.1017/S0033291720003918
- Craske MG, Meuret AE, Ritz T, Treanor M, Dour H, Rosenfield D (2019) Positive affect treatment for depression and anxiety: A randomized clinical trial for a core feature of anhedonia. *Journal of Consulting and Clinical Psychology* 87(5):457–471, DOI 10.1037/ccp0000396, place: US Publisher: American Psychological Association
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8(12):1704–1711, DOI 10.1038/nn1560, URL <http://dx.doi.org/10.1038/nn1560>
- Dayan P (1993) Improving generalization for temporal difference learning: The successor representation. *Neural Computation* 5(4):613–624
- Dichter GS, Smoski MJ, Kampov-Polevoy AB, Gallop R, Garbutt JC (2010) Unipolar depression does not moderate responses to the sweet taste test. *Depress Anxiety* 27(9):859–863, DOI 10.1002/da.20690, URL <http://dx.doi.org/10.1002/da.20690>
- Dunn BD, Widnall E, Reed N, Owens C, Campbell J, Kuyken W (2019) Bringing light into darkness: A multiple baseline mixed methods case series evaluation of Augmented Depression Therapy (ADePT). *Behaviour Research and Therapy* 120:103,418, DOI 10.1016/j.brat.2019.103418

- Eckstrand KL, Forbes EE, Bertocci MA, Chase HW, Greenberg T, Lockovich J, Stiffler R, Aslam HA, Graur S, Bebko G, Phillips ML (2019) Anhedonia reduction and the association between left ventral striatal reward response and 6-month improvement in life satisfaction among young adults. *JAMA Psychiatry* 76:958–965, DOI 10.1001/jamapsychiatry.2019.0864
- Geschwind N, Arntz A, Bannink F, Peeters F (2019) Positive cognitive behavior therapy in the treatment of depression: A randomized order within-subject comparison with traditional cognitive behavior therapy. *Behaviour Research and Therapy* 116:119–130, DOI 10.1016/j.brat.2019.03.005, URL <https://www.sciencedirect.com/science/article/pii/S0005796719300464>
- Geugies H, Mocking RJT, Figueroa CA, Groot PFC, Marsman JBC, Servaas MN, Steele JD, Schene AH, Ruhé HG (2019) Impaired reward-related learning signals in remitted unmedicated patients with recurrent depression. *Brain : a journal of neurology* 142:2510–2522, DOI 10.1093/brain/awz167
- Greenberg T, Chase HW, Almeida JR, Stiffler R, Zevallos CR, Aslam HA, Deckersbach T, Weyandt S, Cooper C, Toups M, Carmody T, Kurian B, Peltier S, Adams P, McInnis MG, Oquendo MA, McGrath PJ, Fava M, Weissman M, Parsey R, Trivedi MH, Phillips ML (2015) Moderation of the relationship between reward expectancy and prediction error-related ventral striatal reactivity by anhedonia in unmedicated major depressive disorder: Findings from the embarc study. *American Journal of Psychiatry* 172:881–891, DOI 10.1176/appi.ajp.2015.14050594
- Herzallah MM, Moustafa AA, Natsheh JY, Abdellatif SM, Taha MB, Tayem YI, Sehwal MA, Amleh I, Petrides G, Myers CE, Gluck MA (2013) Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Frontiers in Integrative Neuroscience* 7, DOI 10.3389/fnint.2013.00067, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3779792/>
- Hull CL (1943) Principles of behavior: an introduction to behavior theory. Principles of behavior: an introduction to behavior theory, Appleton-Century, Oxford, England, pages: x, 422
- Husain M, Roiser JP (2018) Neuroscience of apathy and anhedonia: a transdiagnostic approach. *Nature Reviews Neuroscience* 19(8):470–484, DOI 10.1038/s41583-018-0029-9, URL <https://www.nature.com/articles/s41583-018-0029-9>, number: 8 Publisher: Nature Publishing Group
- Huys QJM, Dayan P (2009) A Bayesian formulation of behavioral control. *Cognition* 113(3):314–328, DOI 10.1016/j.cognition.2009.01.008, URL <http://dx.doi.org/10.1016/j.cognition.2009.01.008>
- Huys QJM, Renz D (2017) A formal valuation framework for emotions and their control. *Biological Psychiatry* 82:413–420, DOI 10.1016/j.biopsych.2017.07.003
- Huys QJM, Eshel N, O’Nions E, Sheridan L, Dayan P, Roiser JP (2012) Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput Biol* 8(3):e1002410, DOI 10.1371/journal.pcbi.1002410, URL <http://dx.doi.org/10.1371/journal.pcbi.1002410>
- Huys QJM, Pizzagalli DA, Bogdan R, Dayan P (2013) Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biol Mood Anx Dis* 3(1):12, DOI 10.1186/2045-5380-3-12, URL <http://dx.doi.org/10.1186/2045-5380-3-12>
- Jackson JC, Watts J, Henry TR, List JM, Forkel R, Mucha PJ, Greenhill SJ, Gray RD, Lindquist KA (2019) Emotion semantics show both cultural variation and universal structure. *Science* 366(6472):1517–1522, DOI 10.1126/science.aaw8160
- Juechems K, Summerfield C (2019) Where Does Value Come From? *Trends in Cognitive Sciences* 23(10):836–850, DOI 10.1016/j.tics.2019.07.012
- Juechems K, Balaguer J, Herce Castañón S, Ruz M, O’Reilly JX, Summerfield C (2019) A Network for Computing Value Equilibrium in the Human Medial Prefrontal Cortex. *Neuron* 101(5):977–987.e3, DOI 10.1016/j.neuron.2018.12.029, URL <https://www.sciencedirect.com/science/article/pii/S0896627318311498>
- Keramati M, Gutkin B (2014) Homeostatic reinforcement learning for integrating reward collection and physiological stability. *eLife* 3:e04811, DOI 10.7554/eLife.04811, URL <https://doi.org/10.7554/eLife.04811>, publisher: eLife Sciences Publications, Ltd
- Knutson B, Bhanji JP, Cooney RE, Atlas LY, Gotlib IH (2008) Neural responses to monetary incentives in major depression. *Biol Psychiatry* 63(7):686–692, DOI 10.1016/j.biopsych.2007.07.023, URL <http://dx.doi.org/10.1016/j.biopsych.2007.07.023>
- Kohli P, Soler ZM, Nguyen SA, Muus JS, Schlosser RJ (2016) The Association Between Olfaction and Depression: A Systematic Review. *Chemical Senses* 41(6):479–486, DOI 10.1093/chemse/bjw061,

- URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4918728/>
- Krajebich I (2019) Accounting for attention in sequential sampling models of decision making. *Current Opinion in Psychology* 29:6–11, DOI 10.1016/j.copsyc.2018.10.008
- Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele JD (2008) Abnormal temporal difference reward-learning signals in major depression. *Brain* 131(Pt 8):2084–2093, DOI 10.1093/brain/awn136, URL <http://dx.doi.org/10.1093/brain/awn136>
- Kunisato Y, Okamoto Y, Ueda K, Onoda K, Okada G, Yoshimura S, Suzuki Si, Samejima K, Yamawaki S (2012) Effects of depression on reward-based decision making and variability of action in probabilistic learning. *Journal of behavior therapy and experimental psychiatry* 43:1088–1094, DOI 10.1016/j.jbtep.2012.05.007
- Lombion-Pouthier S, Vandel P, Nezelof S, Haffen E, Millot JL (2006) Odor perception in patients with mood disorders. *Journal of affective disorders* 90:187–191, DOI 10.1016/j.jad.2005.11.012
- Maier SF, Watkins LR (2005) Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neurosci Biobehav Rev* 29(4-5):829–41, URL <http://dx.doi.org/10.1016/j.neubiorev.2005.03.021>
- McMakin DL, Olinio TM, Porta G, Dietz LJ, Emslie G, Clarke G, Wagner KD, Asarnow JR, Ryan ND, Birmaher B, Shamseddeen W, Mayes T, Kennard B, Spirito A, Keller M, Lynch FL, Dickerson JF, Brent DA (2012) Anhedonia predicts poorer recovery among youth with selective serotonin reuptake inhibitor treatment-resistant depression. *Journal of the American Academy of Child and Adolescent Psychiatry* 51(4):404–411, DOI 10.1016/j.jaac.2012.01.011
- Nair A, Niyogi RK, Shang F, Tabrizi SJ, Rees G, Rutledge R (2020) Opportunity cost determines action initiation latency and predicts apathy. *psyArxiv* psyarxiv.com/d9pgz, URL <https://psyarxiv.com/d9pgz>
- Naudin M, El-Hage W, Gomes M, Gaillard P, Belzung C, Atanasova B (2012) State and Trait Olfactory Markers of Major Depression. *PLOS ONE* 7(10):e46,938, DOI 10.1371/journal.pone.0046938, URL <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0046938>, publisher: Public Library of Science
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191(3):507–520, DOI 10.1007/s00213-006-0502-4, URL <http://dx.doi.org/10.1007/s00213-006-0502-4>
- Pause BM, Miranda A, Göder R, Aldenhoff JB, Ferstl R (2001) Reduced olfactory performance in patients with major depression. *Journal of psychiatric research* 35:271–277, DOI 10.1016/s0022-3956(01)00029-2
- Pizzagalli DA, Holmes AJ, Dillon DG, Goetz EL, Birk JL, Bogdan R, Dougherty DD, Iosifescu DV, Rauch SL, Fava M (2009) Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *Am J Psychiatry* 166(6):702–710, DOI 10.1176/appi.ajp.2008.08081201, URL <http://dx.doi.org/10.1176/appi.ajp.2008.08081201>
- Pizzagalli DA, Smoski M, Ang YS, Whitton AE, Sanacora G, Mathew SJ, Nurnberger J, Lisanby SH, Iosifescu DV, Murrough JW, Yang H, Weiner RD, Calabrese JR, Goodman W, Potter WZ, Krystal AD (2020) Selective kappa-opioid antagonism ameliorates anhedonic behavior: evidence from the fast-fail trial in mood and anxiety spectrum disorders (fast-mas). *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 45(10):1656–1663, DOI 10.1038/s41386-020-0738-4
- Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory* pp 64–99
- Rizvi SJ, Pizzagalli DA, Sproule BA, Kennedy SH (2016) Assessing anhedonia in depression: Potentials and pitfalls. *Neuroscience and biobehavioral reviews* 65:21–35, DOI 10.1016/j.neubiorev.2016.03.004, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4856554/>
- Robinson OJ, Chase HW (2017) Learning and choice in mood disorders: Searching for the computational parameters of anhedonia. *Computational psychiatry (Cambridge, Mass)* 1:208–233, DOI 10.1162/CPSY_a_00009
- Rupprechter S, Stankevicius A, Huys QJM, Series P, Steele JD (2020) Abnormal reward valuation and event-related connectivity in unmedicated major depressive disorder. *Psychological Medicine* pp 1–9, DOI 10.1017/s0033291719003799
- Russek EM, Momennejad I, Botvinick MM, Gershman SJ, Daw ND (2017) Predictive representations

- can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology* 13:e1005768, DOI 10.1371/journal.pcbi.1005768
- Rutledge RB, Moutoussis M, Smittenaar P, Zeidman P, Taylor T, Hrynkiewicz L, Lam J, Skandali N, Siegel JZ, Ousdal OT, Prabhu G, Dayan P, Fonagy P, Dolan RJ (2017) Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry* 74:790–797, DOI 10.1001/jamapsychiatry.2017.1713
- Seligman ME, Maier SF (1967) Failure to escape traumatic shock. *J Exp Psychol* 74(1):1–9
- Spijker J, Bijl RV, de Graaf R, Nolen WA (2001) Determinants of poor 1-year outcome of DSM-III-R major depression in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica* 103(2):122–130, DOI 10.1034/j.1600-0447.2001.103002122.x
- Steele JD, Kumar P, Ebmeier KP (2007) Blunted response to feedback information in depressive illness. *Brain* 130(Pt 9):2367–2374, DOI 10.1093/brain/awm150, URL <http://dx.doi.org/10.1093/brain/awm150>
- Sutton RS, Barto AG (2017) *Reinforcement Learning*, 2nd edn. MIT Press, Cambridge, MA, URL <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
- Swiecicki L, Zatorski P, Bzinkowska D, Sienkiewicz-Jarosz H, Szyndler J, Scinska A (2009) Gustatory and olfactory function in patients with unipolar and bipolar depression. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 33(5):827–834, DOI 10.1016/j.pnpbp.2009.03.030
- Treadway MT, Zald DH (2011) Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience and Biobehavioral Reviews* 35(3):537–555, DOI 10.1016/j.neubiorev.2010.06.006
- Trøstheim M, Eikemo M, Meir R, Hansen I, Paul E, Kroll SL, Garland EL, Leknes S (2020) Assessment of Anhedonia in Adults With and Without Mental Illness: A Systematic Review and Meta-analysis. *JAMA network open* 3(8):e2013233, DOI 10.1001/jamanetworkopen.2020.13233
- Uher R, Perlis RH, Henigsberg N, Zobel A, Rietschel M, Mors O, Hauser J, Dernovsek MZ, Souery D, Bajs M, Maier W, Aitchison KJ, Farmer A, McGuffin P (2012) Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychological medicine* 42(5):967–980, DOI 10.1017/S0033291711001905, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3787526/>
- Watson R, Harvey K, McCabe C, Reynolds S (2020) Understanding anhedonia: a qualitative study exploring loss of interest and pleasure in adolescent depression. *European Child & Adolescent Psychiatry* 29(4):489–499, DOI 10.1007/s00787-019-01364-y, URL <https://doi.org/10.1007/s00787-019-01364-y>
- WHO (1992) *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization, Geneva
- Zhang J, Berridge KC, Tindell AJ, Smith KS, Aldridge JW (2009) A Neural Computational Model of Incentive Salience. *PLOS Computational Biology* 5(7):e1000437, DOI 10.1371/journal.pcbi.1000437, URL <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000437>, publisher: Public Library of Science