

Review

Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making

Jacqueline Scholl*, Miriam Klein-Flügge*

Department of Experimental Psychology, University of Oxford, Tinsley Building, Mansfield Road, Oxford, OX1 3SR, United Kingdom



ARTICLE INFO

Keywords:

Reinforcement learning
Decision-making
Computational psychiatry

ABSTRACT

Recent research in cognitive neuroscience has begun to uncover the processes underlying increasingly complex voluntary behaviours, including learning and decision-making. Partly this success has been possible by progressing from simple experimental tasks to paradigms that incorporate more ecological features. More specifically, the premise is that to understand cognitions and brain functions relevant for real life, we need to introduce some of the ecological challenges that we have evolved to solve. This often entails an increase in task complexity, which can be managed by using computational models to help parse complex behaviours into specific component mechanisms. Here we propose that using computational models with tasks that capture ecologically relevant learning and decision-making processes may provide a critical advantage for capturing the mechanisms underlying symptoms of disorders in psychiatry. As a result, it may help develop mechanistic approaches towards diagnosis and treatment. We begin this review by mapping out the basic concepts and models of learning and decision-making. We then move on to consider specific challenges that emerge in realistic environments and describe how they can be captured by tasks. These include changes of context, uncertainty, reflexive/emotional biases, cost-benefit decision-making, and balancing exploration and exploitation. Where appropriate we highlight future or current links to psychiatry. We particularly draw examples from research on clinical depression, a disorder that greatly compromises motivated behaviours in real-life, but where simpler paradigms have yielded mixed results. Finally, we highlight several paradigms that could be used to help provide new insights into the mechanisms of psychiatric disorders.

1. Introduction

Recent research in cognitive neuroscience has produced an array of paradigms and computational approaches to study facets of motivation, learning, and decision-making. While our understanding of the underlying mechanisms in the healthy brain is progressing at a fast pace, the progress made in understanding psychiatric disorders has been slow in comparison. Patients and their practitioners report very striking impairments in day-to-day learning and decision-making, yet many lab studies reveal only small differences, if any, between patients and healthy controls. Several reasons could account for this disparity including patient sample sizes [1], disease heterogeneity [2,3], or that disease phenotypes cut across diagnostic criteria for different diseases [4,5]. But there is another possibility explored here, namely that commonly used paradigms may not be sufficiently sensitive to the relevant features of everyday cognitive processes.

Tasks in the laboratory can be overly simplistic and not capture the sophistication and complexity of real-life situations. Alternatively, tasks

can be framed in ways that are unnatural and for that reason fail to capture cognitions relevant for everyday life. To counteract these problems, basic neuroscience research has recently turned towards laboratory tasks that incorporate more ecological features ([6–10], Table 1 and Fig. 1). Here, the emphasis is on designing experiments that capture the types of processes that our brains have evolved to solve. More specifically, the idea is to identify the relevant cognitive process of interest, and to design a task, which requires the same process (and thus underlying brain networks) and thus mimics the computation identified as relevant to everyday learning and decision-making. Ecologically inspired designs do not by nature have to be more complex or be run in natural environments (there is still a balance to be struck between ecological validity and simplicity) but they require careful consideration of the processes relevant for a behaviour of interest and tasks that are adapted to precisely probe the underlying mechanisms. This review focuses on the question of whether paradigms that capture features of the cognitive processes required in real-life may also help us understand what is functionally changed in psychiatric disorders. In the

* Corresponding authors.

E-mail addresses: Jacqueline.scholl@cantab.net (J. Scholl), Miriam.klein-flugge@psy.ox.ac.uk (M. Klein-Flügge).

<http://dx.doi.org/10.1016/j.bbr.2017.09.050>

Received 31 March 2017; Received in revised form 24 July 2017; Accepted 27 September 2017

Available online 28 September 2017

0166-4328/© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1

A list of tasks incorporating features of learning and decision-making that are relevant in natural environments. For all of these tasks, computational models have been proposed. The reader is referred to the text and relevant papers for further details. This list is meant to highlight some important examples but it is not an exhaustive list.

Learning	
Sharot et al., <i>Nature</i> [252]	How strongly is the updating of people's beliefs biased towards positive compared to negative information ('optimism bias')?
Behrens et al., <i>Nat. Neurosci.</i> [51] 2007; Mathys et al. [52] <i>Front Hum Neurosci.</i> , 2011	How do people flexibly adjust their speed of learning to the environment (thus providing a measure of the ability to learn that is not available when learning in stable environments)?
Jocham et al. [79], <i>Neuron</i> , 2016; Lee et al. [80] <i>PLOS Biol.</i> , 2015; Leong et al. [61], <i>Neuron</i> , 2017	How do people attribute outcomes to causes and what are potential biases when making such attributions?
Scholl and Kolling et al., <i>J. Neurosci.</i> [102], 2015	How do people focus learning on what is important, ignoring irrelevant but salient reward information?
Lee et al., <i>Neuron</i> [97], 2014	How do people flexibly adjust, or arbitrate, between model-based and model-free learning systems?
Decision-making	
Klein-Flügge et al., <i>PLOS Comput. Biol.</i> [247] 2015	How do people weigh up rewards and different kinds of costs (delays or effort)?
Huys et al., <i>PLOS Comput. Biol.</i> [169] 2011; Guitart-Masip et al. [165], <i>NeuroImage</i> , 2012; Bach et al., <i>Curr. Biol.</i> [176], 2014	How do reflexive Pavlovian and reflective goal-directed valuation systems interact?
Eldar et al. [182] <i>Nat. Commun.</i> , 2015	How do emotional states induced by task-irrelevant wins or losses influence subsequent evaluations of stimuli?
Wilson et al. [224] <i>J. Exp. Psychol.</i> , 2015; Blanco et al. [29], <i>Cognition</i> , 2013	How do people adjust the degree of exploration and exploitation given environmental constraints?
Kolling et al. [119] <i>Science</i> , 2012	How do people weigh up whether to take an immediately available option or to look for alternatives elsewhere?

majority of cases, the progress made in basic neuroscience has not yet been translated to clinical populations.

We will first provide an introduction to basic concepts of learning and decision-making for readers new to this field, including basic concepts of computational models of cognition. Computational modelling is essential, especially for creating more sophisticated tasks, because models can parse and quantify the performance on different sub-processes that may be recruited within the same task. In the remainder of the review, we illustrate, for both learning and decision-making, how simple paradigms have been extended to begin to capture some of the sophistication of natural environments, and how this has provided some key insights into both behaviour and brain function that cannot be gained from simpler tasks alone. Throughout, we draw examples from clinical depression to illustrate the relevance of the ecological approach. We propose that ecological tasks may help to bridge the gap between impairments seen in real-life and those reported in the laboratory. This may allow the field to advance from symptom-based to more mechanistic and quantitative diagnoses (see also [11–15]). However, we find that, while researchers have successfully begun to apply simpler paradigms to psychiatric patients, most of the more advanced and ecologically valid paradigms that we highlight have not yet been used in the context of psychiatry. One potential way forward could be to use these tasks in larger samples of patients online [1,16,17] to help identify sub-clusters within disorders but also common symptoms across disease boundaries.

2. Learning

2.1. Basic learning processes

Learning from experience is crucial for adaptive behaviour. Many decisions we make in daily life are based on values that we have learnt from experience. For example, when deciding whether to go out to meet friends at a pub, this depends on how much you have enjoyed similar experiences in the past. Studying learning holds great promise for understanding depression because it intuitively relates to many of the typical symptoms. As an illustrative example, let's imagine a depressed patient showing social withdrawal, but on one occasion she goes out and thoroughly enjoys meeting her friends. Nevertheless, when deciding whether to go out again, she chooses not to, possibly because she did not update her belief about how enjoyable this would be. In this way, reduced learning from positive experiences could be a mechanism that maintains depression. Of course, there are other potential reasons for deciding not to go out, including for example motivational deficits, which we will consider later. But what this example

highlights is how real-life complexities can be parsed into separate measurable component processes. In the following, we will first consider simple learning processes, originating from a behaviourist framework, before considering more ecological types of learning from a cognitive and computational perspective.

The simplest learning scenario, studied extensively in psychiatric research, involves the learning of associations between a single stimulus (e.g. an abstract shape) and an outcome (e.g. monetary reward). This kind of scenario directly relates to behaviourist views on behaviour and therapy [18] which have focused on different forms of conditioning (Pavlovian or operant); e.g. symptoms of anxiety are seen as a maladaptive learnt response to a particular stimulus and the treatment involves learning new responses [19–21].

More specifically, in a typical experiment (Fig. 2A), learning is measured by presenting repeated pairings (trials) of the stimulus followed by the reward or no reward; the probability of a stimulus being followed by a reward is set by the experimenter. The measure of (clinical) interest is how participants learn these associations between stimuli and rewards, as indicated, for example, by how quickly they prefer stimuli that more likely lead to reward. To capture this learning we can use a computational model, i.e. an algorithm that simulates the processes going on in participants' brains as they gradually learn across trials (Fig. 2B). In this particular model, learning is driven by how unexpected the outcome is (i.e., by the prediction error (PE): the difference between outcome and prediction). A large range of studies have identified PE signals in several brain areas [22–26], especially in regions receiving strong dopaminergic inputs such as the striatum. How much the PE is used (by the model or participants' brains) to update their beliefs is determined by a free parameter in the model, called the learning rate (α). The higher the learning rate, the faster a person updates their beliefs (Fig. 2C + E). This mathematical description of the learning process as one that simply depends on how frequently a stimulus or action is paired with reward has also been coined as 'model-free' learning (see e.g. [27] for a review). Unlike other types of learning, it does not rely on a model of the world. However, some key features that determine whether an agent is indeed behaving in a model-free way cannot be examined by the simple task design described above. Specifically, model-free learning is not flexible. For example, you might find yourself taking your usual way to work even though you actually intended to go somewhere else today. Furthermore there is evidence that even in simple tasks, humans use other brain mechanisms, such as working memory, to supplement the model-free learning mechanism [28]. Therefore, to avoid confusion we will in the following use the

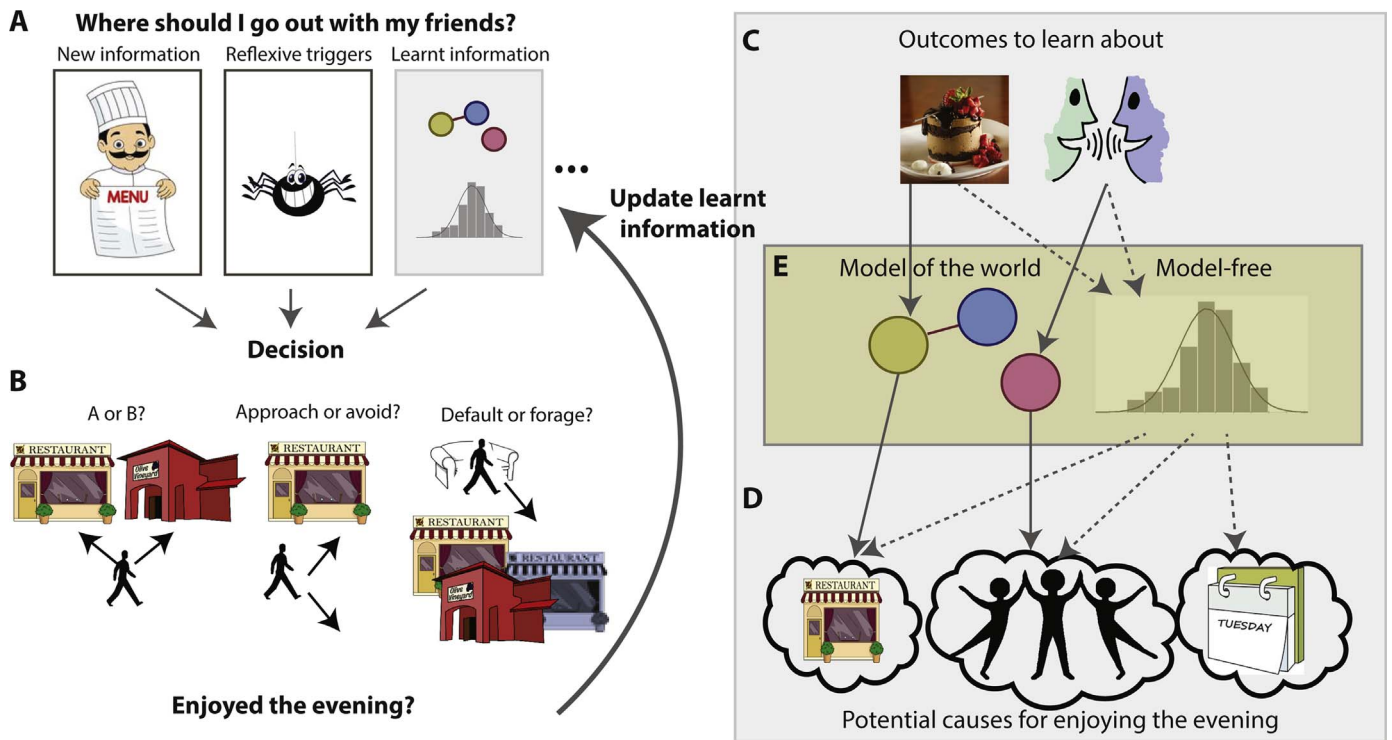


Fig. 1. Overview of processes relevant for learning and decision-making.

Many real-life situations use cognitions that can be grouped under the umbrella terms of learning and decision-making. However, examining the real-life situations more closely, we can see that there are many distinct component processes that rely on different neural substrates and can therefore be differentially affected by psychiatric disorders. A) When making decisions, we take into account different kinds of information, such as different types of rewards (e.g. money, food) or costs (e.g. delay or effort). This can be new information only available explicitly at the time of choice (e.g. reading a menu in a restaurant) or it can be learnt and recalled from past experience. Beyond these types of information that we want to take into account, other stimuli may not be relevant to the decision at hand, yet reflexively affect our judgment (e.g. seeing a spider crawl across the menu you are reading). Integrating across different kinds of information ultimately enables us to make decisions. B) In real-life situations, we make different types of decisions. Sometimes we are presented with concrete options amongst which to choose ('A or B?'). Sometimes we have to decide whether to approach something or avoid it. Lastly, sometimes we are engaged with a behaviour (e.g. relaxing on the sofa) and need to decide whether to continue with this default or go and look for something else (e.g. decide to go into town to look for a restaurant (forage)). C) As a result of our decisions, outcomes happen (e.g. eating nice food, having a good conversation). D) However, in addition to the plausible causes for outcomes (good restaurant -> good food or friends -> enjoyable conversations), there can also be other causes present (e.g. day of the week) that are less likely to have caused the outcome. E) This multiplicity of causes and outcomes poses an attribution problem (i.e. how do you know which outcome to attribute to which cause). This can be resolved using diverse mechanisms. For example, we can either use a model of the world that tells us which outcomes and causes we should learn about and how they might relate; or we can learn in a model-free way, i.e. learn about all outcomes and causes (i.e. also about the implausible or irrelevant ones) based on how often outcomes and causes occur together. Developing tasks that can capture the ecological complexity illustrated by this example is of paramount importance for understanding the psychological and neural mechanisms underlying psychiatric disorders in real-life.

term 'simple learning tasks', rather than 'model-free'.

Paradigms measuring learning of these simple, behaviourist-inspired, associations have been used extensively to study learning in depression. The original hypothesis was that depression should be related to worse learning, which should be reflected in a reduced learning rate. Below we explain why this may not be the best possible hypothesis. Indeed, the behavioural results of different studies have been mixed and not consistently demonstrated learning deficits in simple learning tasks [29–41]. Changes in neural signals, by contrast, specifically a reduced PE encoding in subcortical areas including the striatum, have been reported somewhat more consistently ([42–44], but see also [17] for intact PE coding in a non-learning context). One cause for the variability of the behavioural findings might be that each individual study only included a small number of participants and while some findings have been brought together [45], no formal meta-analysis has been performed on the complete evidence. However, a meta-analysis [46] of a subset of these studies found no evidence for a change of the learning rate in depression. Relatedly, Gillan et al. [47] recently analysed learning in a large online sample of 1400 participants and again, found no evidence for a relationship between measures of simple (model-free) learning and questionnaire measures of depression severity.

One potential explanation for the discrepancy between real-life deficits and the apparent absence (or at best subtle nature) of basic learning deficits in depression could relate to the nature of the learning processes probed by

these tasks. Using more ecological tasks could reveal learning deficits that are not apparent in simpler tasks. As we will show in more detail below (section 'Not all value is equal'), different types of values are processed in different brain regions. Therefore, it is possible that depression will only impact learning about some types of value but not others. Indeed, when measuring updating (i.e. learning) of beliefs about real-life events, two recent studies consistently found that depression alters learning [48,49]. Healthy controls updated the beliefs of how likely negative life events were to happen to them in a biased way, i.e. they updated their beliefs more when given desirable information. By contrast, depressed patients did not show this optimism bias in learning. Another reason why ecological tasks may reveal changes in learning more clearly than simpler tasks is because more sophisticated learning mechanisms need to be recruited. In the next section, we will consider precisely such situations, namely when learning needs to be adapted to match the stability of the environment or when making causal attributions in environments with many possible causes.

2.2. Environmental context and adaptive learning

As noted earlier, in the simplest learning experiment, the speed of learning as captured by the learning rate is a measure of individual differences. Many studies focus on whether patients learn faster or slower than controls, with faster learning often being equated to better

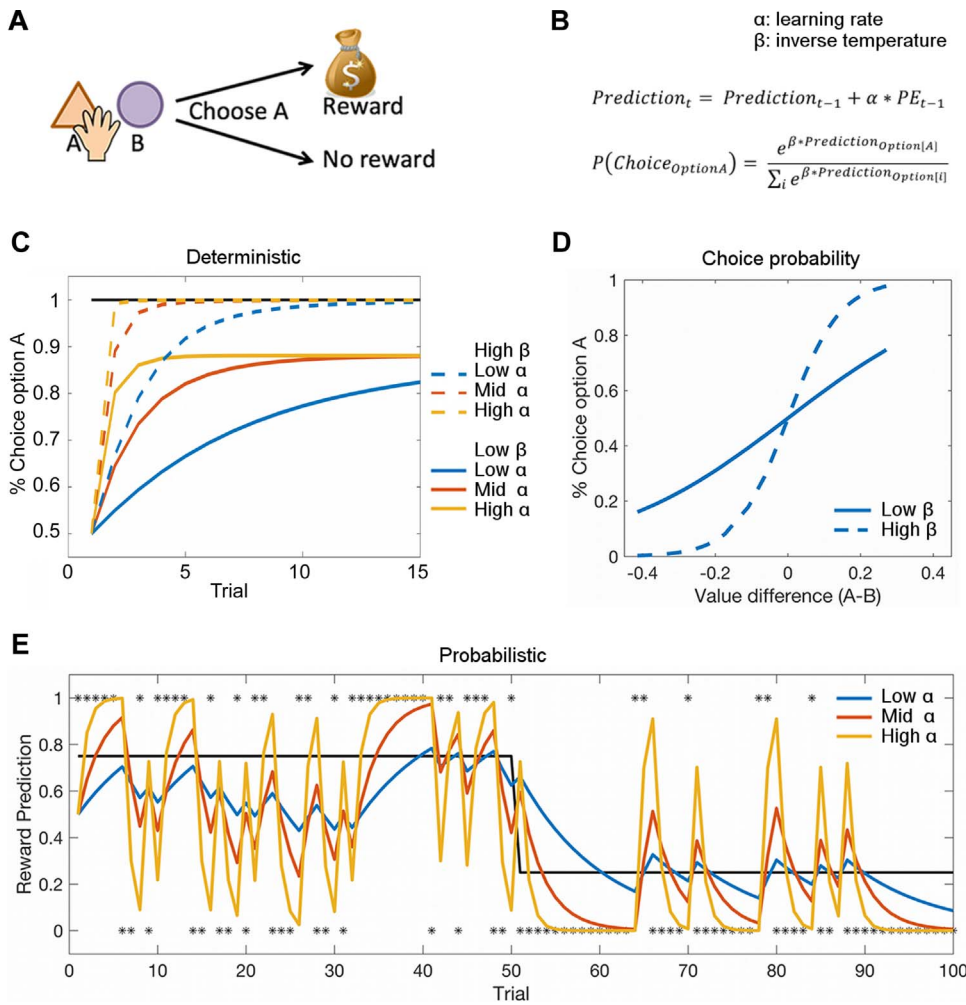


Fig. 2. Describing learning and decision-making using computational models.

A) Schematic of a task examining simple learning. On each trial, participants are presented with different options amongst which to choose. They try to choose the option that is more likely to lead to a reward. Once they have made a choice, they either receive a reward or not. Based on repeated trials, participants try to learn how likely each stimulus is to lead to reward. B) This behaviour can be described with a simple learning model. In the simplest kind of model, there are two free parameters for each person, the learning rate (α) and the inverse temperature (β): The model learns, i.e. updates its predictions, on each trial based on the difference between the outcome (reward or no reward) and the expected outcome (probability of reward), i.e. the reward prediction error (PE). The learning rate (α) determines how quickly predictions are updated based on prediction errors. Based on these predictions, the model then chooses between the options based on a softmax decision rule, i.e. the model does not always pick the option with the higher value, but only chooses that option with a certain probability (for more information see section ‘Decision-making: choice stochasticity’). How stochastic the choices are is determined by the stochasticity parameter, commonly called the ‘inverse temperature’, β (the term derives from the thermodynamics: at lower temperatures (i.e. higher ‘inverse temperature’) particles move less. Translating this to human behaviour, the higher the inverse temperature, the less random participants’ behaviour). This model can then be fitted to participants’ behaviour. This means that for each participant we determine the value for the ‘free’ parameters α and β for which the model most closely matches participants’ choices. C) The effect of different values for α and β in a deterministic task, i.e. a task in which the probability of reward is 100% for an option (and the model chooses between this option and one with a known reward probability of 50%): The higher β (dashed lines), the more likely the model is to consistently pick the option with

higher value so that eventually only the better option is chosen; in contrast, for a lower β (continuous lines), the model continues to select the lower value option from time to time even once it has learnt the value of the better option. The higher α , the faster the model starts to prefer the better option. In this deterministic task, higher α is always better. D) This effect of β can also be illustrated by plotting the probability of choosing one option over the other as a function of the value difference between the two options. β is reflected in the slope of this curve, the higher β , the steeper the slope. D) The predictions learnt using different learning rates in a probabilistic task, i.e. when the stimulus only gives a reward on some trials (75% reward probability for trials 1–50 and 25% for trial 51–100, black line). Now, having a very high α (yellow) is no longer advantageous because random reward omissions (e.g. at trial 30–31) quickly pull beliefs away from the true probability. In contrast, a lower α (blue) means that beliefs are more resistant to random reward omissions. Therefore when the environment is noisy, it is more optimal to integrate information from past outcomes over a longer period, rather than relying only on the most recent observations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

learning. However, in ecological environments, faster learning is not necessarily better learning. Instead the speed of learning should be matched to the environment ([50], Fig. 2C + E and Fig. 3A). To illustrate this, imagine you are trying to predict what mood your friend is in; if your friend generally has a stable mood, then knowing how she has felt over the last week gives you a good indication of how she feels now. In contrast, if your friend is stressed and therefore has more unstable or volatile moods, then knowing how she felt last week may not be informative. Rather, you need to find out how she felt yesterday or even an hour ago. Expressing this intuition in terms of learning rates, if the association you are trying to learn is stable, a low learning rate is advantageous but for unstable associations, i.e. those that change more quickly over time, fast updates and thus a high learning rate are more appropriate. As a computational measure of ‘the goodness’ of learning, we can then consider how well participants can adapt their learning rate to match the environment.

Tasks have been designed to measure the ability to adjust the learning rate to the environment [51,52] and they have recently also been applied to psychiatric questions [53–55]. For example, Browning

et al. [53] found that increased levels of anxiety (which is often comorbid with depression) correlated with a decreased ability to adapt to different environments. In other words, anxious individuals showed a reduced change in learning between unpredictable and changing contexts (where it makes sense to feel anxious) compared to stable contexts, which might be considered ‘safer’ (Fig. 3A). Relatedly, de Berker et al. [54] found that perceived life stress (which is a risk factor for depression) was predictive of how volatile, i.e. unstable, participants perceived an environment in a laboratory task. This relationship could suggest that life stress is the result of living in volatile or unpredictable environments. Or alternatively, the causality might be the other way around and perceiving one’s environment as more volatile than it actually is, may cause chronic stress. Of course to establish causation, future longitudinal studies are needed. Neither of those potential explanations has so far been tested in the context of depression.

The quality of learning can also be captured by parameters other than the learning rate. In environments with interfering information such as distractors or when learning about multiple things simultaneously, it might be most important to learn robustly without

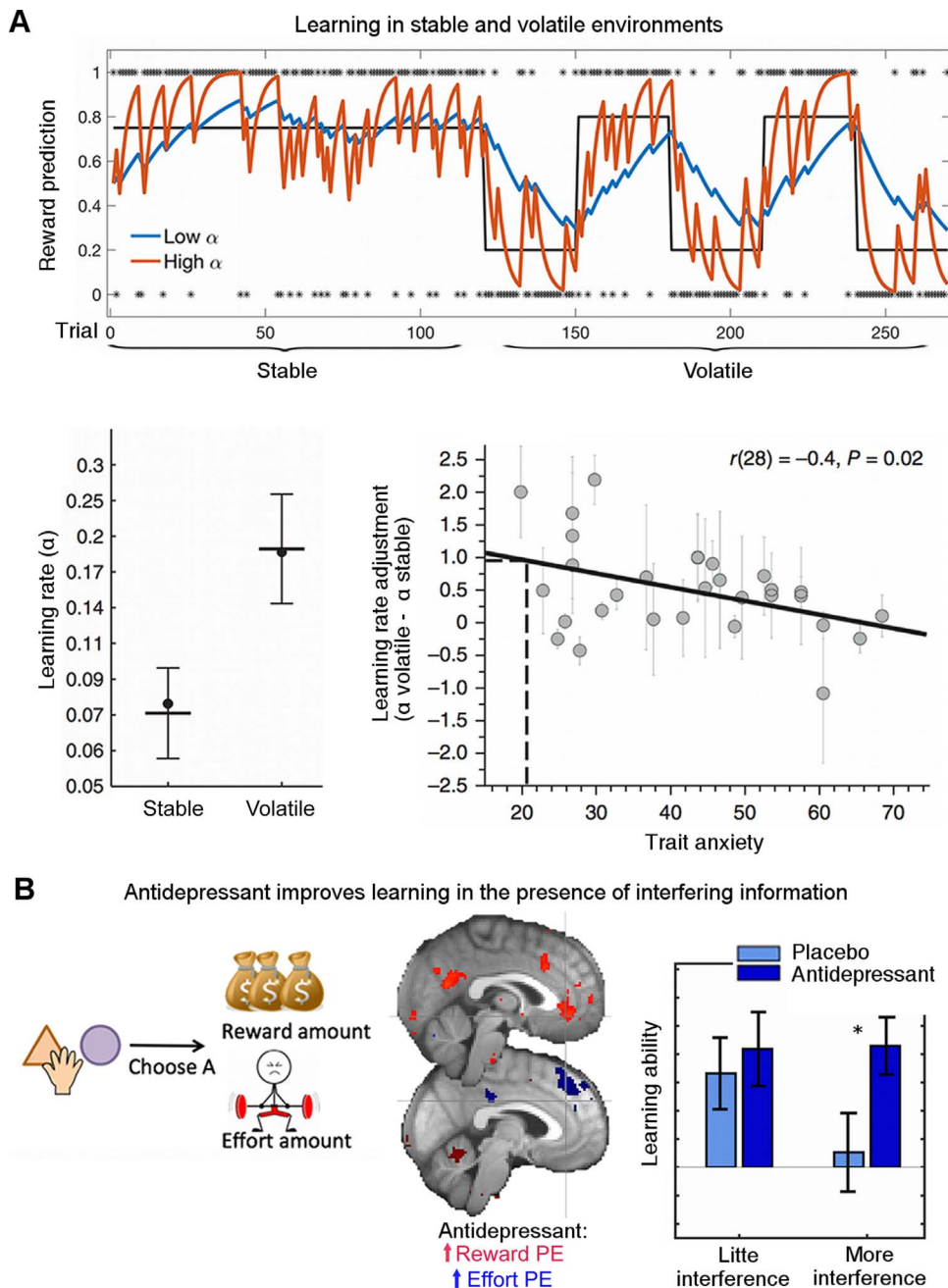


Fig. 3. Environmental context and adaptive learning.

A) Effects of different simulated learning rates (α) on learning in stable (trials 1–120) or volatile (trials 120–270) environments. In stable environments, lower α (blue) results in predictions closer to the true underlying probability of a stimulus (black) and predictions that are less affected by random reward omissions. In contrast, in unstable environments, lower α leads to predictions lagging behind the quickly changing underlying probabilities, while higher α (red) leads to predictions that track the true probabilities more closely. Behrens et al. [51] (bottom left) found that indeed human participants modulate their learning rate, learning more slowly in stable environments and faster in volatile environments. Browning et al. [53] (bottom right) found that this ability to adjust α between volatile and stable environments was related to trait anxiety with more anxious participants being less able to adjust their learning rates. B) We [56] designed an experiment to measure whether serotonergic antidepressants, that have been proposed to increase plasticity in animal models, also improve learning in humans. In the task (here simplified to the relevant features), participants repeatedly chose between two options based on their reward and effort magnitudes, which had to be learnt from experience. Neurally, the antidepressant increased learning signals, i.e. prediction errors, for both reward (left, red) and effort (right, blue). Importantly, participants had to learn different dimensions (reward and effort) simultaneously which meant that they could interfere with each other, thus making learning more difficult. Therefore, learning well could mean learning more robustly, i.e. being less affected by interference. Indeed we found that compared to placebo, antidepressants increased how well participants could learn (i.e. use prediction errors to guide future choices), when there was more interference. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

interference. To illustrate this, in the above example about learning how much you enjoy social encounters, imagine learning about how much you enjoy meeting your friends, while at the same time learning about how much you like the new pub and how much effort it was to get to the pub. It is clear that in this situation learning will be made more difficult by distractions that compete for processing resources. However, learning could be improved by increasing the neural strength (“representation”) of the relevant learning signals. This would ensure that you can learn what you are intending to learn without being distracted; in other words, learning could be improved without having to change the learning rate. We measured this ability to selectively learn from the relevant feature in a study in which participants needed to learn how much monetary reward was associated with different stimuli. While doing so, there were several sources of interfering information ([56], Fig. 3B). We found that a selective serotonin re-uptake inhibitor (SSRI, citalopram) commonly prescribed to treat depression boosted the

relevant neural learning signals and enabled participants to learn better in the face of interference. Of clinical relevance, finding increased neural learning signals supports the hypothesis that one way in which antidepressants act is by increasing learning or relatedly brain plasticity [57]. One future avenue might be to test whether early changes in plasticity can be predictive of treatment effects after weeks or months, thus helping to better tailor treatment to patients with potentially different combinations of symptoms of depression.

2.3. Attributing outcomes – the role of attention in selecting candidate causes

The learning scenarios considered so far focused on learning the strength of association between a single candidate cause and an outcome. However, in ecological situations, there is often more ambiguity about what even constitutes a potential candidate cause or more

broadly a relevant stimulus (Fig. 1). For example, if you did not enjoy yourself in the company of friends, was it because your friends were overly tired, because the pub did not serve good food, because you said something wrong that upset everyone or because you wore a pink shirt? Importantly, there are distinct brain mechanisms that enable us to selectively attribute outcomes to the most appropriate causes in such situations.

The first step in this more sophisticated learning process is to narrow down the number of possible causes (from thousands) to only a few realistic ones based on ecological heuristics. This means ignoring unrealistic causes completely (e.g. wearing a pink shirt should not influence how much you enjoy your friends' company). However, even using these heuristics, there may still be too many causes for the brain to keep track of simultaneously. Therefore, further mechanisms are needed to simplify the attribution problem. One is to selectively pay attention to only a few possible causes at a time and mentally test hypotheses about them. That means gathering evidence for or against the hypothesis that a cause predicts an outcome. This is done until the hypothesis can be confirmed or disconfirmed, and another hypothesis can be tested. There is indeed evidence that people focus attention on likely causes and that this influences how outcomes are attributed during learning [58–60]. Akaishi et al. [60] found, using computational modelling, that people paid attention to one hypothesis at a time and learnt selectively depending on whether this hypothesis was confirmed or not. Beyond leading to a categorical selection of potential causes, attention can also have more gradual effects (see the next section for a fuller discussion): Leong et al. [61] measured attention by tracking people's eye movements, and found that participants deployed attention more to causes that seemed likely, and that participants were in turn more likely to attribute outcomes to the causes in their attentional focus.

There is ample evidence of attentional biases in depression, where an increased attention to negative outcomes has been reported [62,63]. Relatedly, mood has been linked to the breadth of attention, i.e. to how attention is spread over different stimuli rather than focused on a single stimulus, with positive mood being linked to increased breadth of attention [64–66]. Interestingly, breadth of attention has also been related to noradrenaline levels, approximated by measures of pupil dilation [67]. As noradrenaline levels are known to relate to stress, this could suggest a neural substrate for how happiness or stress affect the breadth of attention. However, it remains unclear whether these attentional biases influence how patients learn to attribute outcomes to causes.

2.4. *Attributing outcomes – neural mechanisms*

Having narrowed down the number of possible causes you are considering at any one time, how does the brain correctly attribute outcomes to their underlying causes? Specific component processes rely on different specialized brain areas (see Fig. 4 for the anatomical location of the brain areas discussed in the following sections). As those component processes could lead to behavioural changes in unique ways we will consider them individually below. The concept of component processes is of particular relevance when considering the heterogeneity of psychiatric disorders because certain processes could be affected by a disorder but others left intact; and different subtypes of a disorder could affect different component processes.

Distributed neural networks can process and keep online the relevant pieces of information or relevant dimensions of an outcome that determine its value: First, the reward *identity* needs to be processed. This could be the broader category of reward (e.g., a social reinforcer or food item) or its specific characteristics (e.g., salty or sweet). For instance, when processing food rewards, outcome identity is encoded in a region of prefrontal cortex, namely the posterior and medial region of OFC [68,69]. Second, it is important to keep in *memory* what potential causes or stimuli are currently relevant, when those do not appear at

the same time as the outcome. To date, our knowledge of such representations is limited, but a recent study found evidence for stimulus-specific memory traces in the same sensory areas that initially processed the potential causes in the experiment [70]. Third, a representation of the *strength of the association* between the candidate causes and outcomes to be updated needs to be stored; there is evidence that the hippocampal memory systems [71] and more lateral OFC (lOFC) [69] keep track of this information. To make correct attributions, all these different mechanisms need to work together. Different lines of evidence suggest that integration of these different pieces of information happens in lOFC. When this area is lesioned, monkeys [72–74] and humans [75] no longer attribute outcomes to the correct causes. Additionally, fMRI studies have shown that during the learning process, lOFC is active when specific attributions are made [71,76–79]. Using a more complex design with many possible causes (see Fig. 5A for detailed explanation), Jocham et al. [79] found that people in whom lOFC is more active during learning are better at making correct attributions.

Knowing which causes and outcomes to consider generally, the next question is how much of the outcome to attribute to each cause. To return to the example of figuring out why you enjoy socializing, imagine one day there is a new person present. If everything else is the same, and you feel different, it would seem most likely that this change in your feelings is attributable to the new person. In other words, a good heuristic is that surprising outcomes are more likely to be due to causes that you are uncertain about or know less about (e.g. because they are new) than by ones you are familiar with. We can quantify this intuition in mathematical terms as uncertainty about how a cause and an outcome are related. That people indeed use uncertainty to guide attributions has been confirmed in a recent study using a novel design ([80], Fig. 5B). Participants had to learn to attribute outcomes to causes with different causal uncertainty. Neurally, the computation of uncertainty was linked to a region in the prefrontal cortex, namely the ventrolateral PFC (vlPFC). It is interesting to note the potential parallels between this gradual effect of uncertainty on learning and the gradual effects of attention [61] and environmental volatility [51,52] discussed above. These *quantitative* effects whereby a potential cause is given a stronger or weaker weight in predicting an outcome stand in contrast to more categorical selection effects where some potential causes are completely disregarded so that outcomes will not at all be attributed to them.

Interestingly, completely independently from the above mentioned line of research, faulty attributions have been proposed as a key mechanism in the development of depression [81,82]. Questionnaire-based studies have found that patients with depression or healthy people who later go on to develop depression are more likely to attribute negative events to themselves rather than external causes, compared to healthy controls [81–87]. However, results from experimental studies have been less clear-cut ([88–92] and [93] for a review). It is therefore currently not possible to conclude that the attribution of outcomes to their underlying causes is a process that is generally impaired in depression. The application of new and more ecological paradigms, such as the ones described here and in the next section, could help to shed light on how the diverse mechanisms are affected in depression.

2.5. *Attributing outcomes – different strategies for different situations*

The mechanisms for making attributions described above work well when agents have an accurate model of the world, i.e. when they know what the plausible causes are and what outcomes are important. However, in the real world, it may sometimes not be possible to have a precise model and one's model may also be entirely wrong. For example, if you do not even consider the possibility that the cause of your friends liking you is because you are a nice person (and instead only consider external causes such as them being polite), you can never learn about your niceness. A solution to this problem might be to have

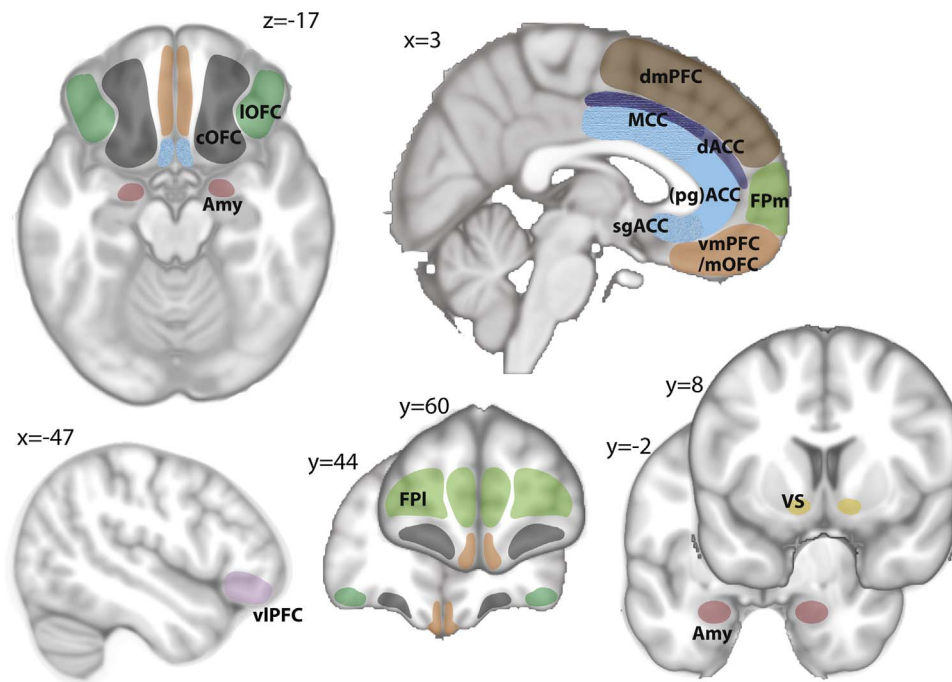


Fig. 4. Overview of brain regions.

Schematic highlighting the subcortical and prefrontal brain regions most central for the learning and decision-making processes discussed in this review. The anterior cingulate cortex (ACC) is shown in shades of blue with the dorsal sulcal portion dACC/ACCs in dark blue and the ventral gyral portion ACCg in light blue. The posterior aspect of ACCg and ACCs is also referred to as midcingulate corect (MCC) [236,237], the aspect curving around the genu of the corpus callosum as perigenual ACC (pgACC) and the most ventral portion as subgenual ACC (sgACC). The frontal pole (FP, green) spans a large area of cortex and can be subdivided into its medial (Fm) and lateral (FPI) portion. Note that these are the abbreviations used throughout the review but there is not always consistency in the literature (for example, what we refer to as central orbitofrontal cortex (cOFC) has sometimes been referred to as medial OFC (mOFC) as well). (vmPFC = ventro-medial prefrontal cortex; dmPFC = dorso-medial PFC (includes pre-SMA); IOFC = lateral OFC; vIPFC = ventrolateral prefrontal cortex; VS = ventral striatum; Amy = Amygdala). For more detailed maps see [237–241]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

concurrent learning mechanisms, some of which indiscriminately learn about all possible causes that are present, without filtering out unlikely causes.

Indeed, studies in which the IOFC – the region we described above as being important for attributing causes based on a model – is lesioned show that non-human primates can still learn, but that they use a different strategy [72,73]. Specifically, control monkeys associated outcomes most strongly with stimuli that actually caused them (e.g. a stimulus preceding an outcome in the same trial). In contrast, after lesions, monkeys associated outcomes with stimuli that were temporally proximal to the outcome even if they could not have caused the outcome (e.g. stimuli from previous or subsequent trials).

While such a strategy of learning by temporal proximity is not optimal if you have a model of the task, it is a good additional strategy in many natural environments where it is often less clear which cause is going to produce which outcome and when. Indeed, Jocham et al. [79] found evidence that such a learning mechanism exists by changing the task structure from one in which outcomes could be clearly associated with previous causes (as shown in Fig. 3B; and which humans solved using IOFC) to one in which it was not clear when in the past causes for outcomes had occurred (i.e. when participants could not use a model of the task; not shown). In the latter situation, humans changed their learning and now associated outcomes with choices at varying times in the past. This type of imprecise learning has been termed ‘spread-of-reward’ and linked to the amygdala [78,79].

2.6. Selecting the appropriate learning mechanisms

We have considered how more complex learning may depend on different kinds of brain mechanisms all operating together at the same time. But learning with several independent brain systems in parallel raises a new issue: How do you decide which one should guide choices? For example, one system could tell you to choose stimulus A, while the other one might favour stimulus B.

While this has not yet been studied for attributional learning, it has been studied in the context of model-based and model-free learning. We will not consider the task in detail – it has been covered in some excellent reviews [27,94–96]. In brief, in this task, there are also different

ways in which participants can learn, relying on a complex model of the task (‘model-based’), or on simpler mechanisms (‘model-free’). Lee et al. [97] proposed that one way to decide which system should be used to drive choices would be to monitor how reliable the knowledge of either system is. To examine this, they manipulated how well different parts of their task could be solved by each of the systems. They found that the signature of this process of assessing the reliability of each system was associated with activity in a distinct area of prefrontal cortex, namely the vIPFC and frontal pole. It may be important to note that while this specific task and related versions of it [98,99] measures one form of what is referred to as ‘model-based’ learning, there are many other ways in which learning can be ‘model-based’ or in other words, rely on a model of the task. As described above, for example, using a model of the world to know what stimuli are potentially relevant, using a model of how quickly the environment is changing to adapt the learning rate, using uncertainty to adapt the learning rate, learning from fictive feedback [100,101], or finding out which learning mechanism is appropriate (as described here). This is important because these very distinct forms of learning that are not ‘model-free’ rely on different brain mechanisms and are therefore likely affected differently by different psychiatric conditions.

Unlike situations where there is ambiguity about which learning mechanism is appropriate, there are situations where the model of the world is clear and model-free learning is not appropriate. For example, when meeting friends in the bar, it is clear that whether or not you like the music should not inform your judgment of how much you like your friends. However, if the music is emotionally salient (strongly driving the model-free learning mechanism relying on temporal proximity discussed above), it may be hard to ignore it. Model-free learning in this context means that the pleasant feeling caused by the music spills over to your judgements of your friends because the two events co-occurred in time. In contrast, if you have an accurate model of the world (i.e. you know that the enjoyment of the music cannot be caused by your friends), you should not make this misattribution. How can the brain deal with this kind of situation? We tested this in an experiment in which participants had to learn to predict how much reward was associated with different stimuli [102] (Fig. 5C). In that experiment there were also irrelevant, yet salient, rewards (like the music in the bar). We

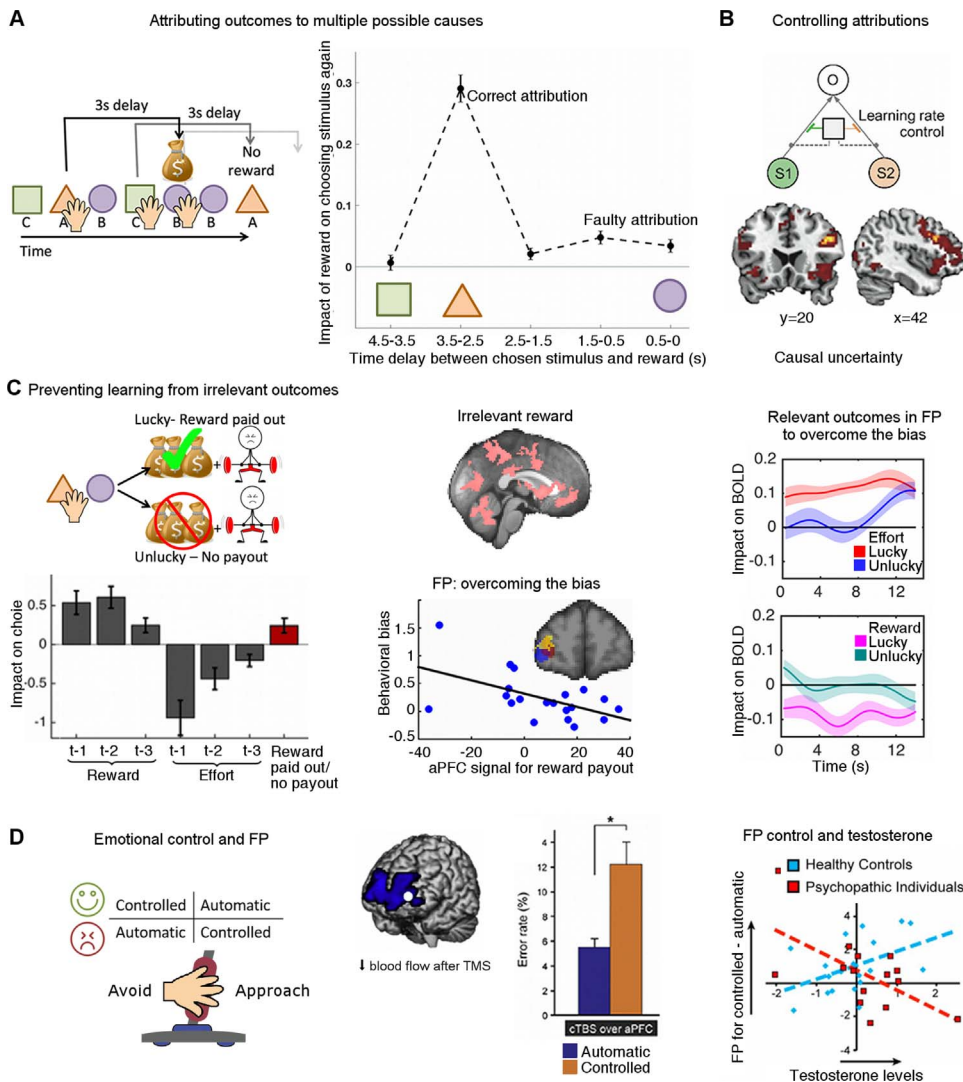


Fig. 5. Different mechanisms for making attributions.

A) Jocham et al. [79] designed an experiment to assess how well people can attribute outcomes to correct causes or instead make faulty attributions. In the task (simplified here), participants were presented with a continuous stream of stimuli, each appearing on the screen for 1.5 s. For each stimulus, participants could either choose it for a small cost (hand symbol) or ignore it. If participants selected a stimulus it led to a reward (with a certain probability) after a fixed 3 s delay. This meant that in between having selected a stimulus (e.g. orange triangle) and receiving its reward, other stimuli appeared on the screen. Thus, the reward could potentially be misattributed to other stimuli that did not cause it. Behavioural results (right) showed that participants were mostly attributing outcomes to the correct causes: they were most likely to select a stimulus again that had appeared about three seconds before the reward. However, they also misattributed outcomes to causes that occurred just before the outcome (1.5 s to 0s). B) Lee et al. [80] designed an experiment to measure to what degree outcomes were associated to different plausible causes, or in other words how high the learning rate was for each cue. They found that how the learning rate was split amongst different cues depended on the causal uncertainty: the more participants were uncertain about whether a cue had caused an outcome, the higher the learning rate for that cue. Neurally, they found that ventrolateral prefrontal cortex represented this causal uncertainty. C) We [102] designed an experiment to assess how the brain avoids misattributing salient, yet irrelevant, outcomes. In the task (simplified here), participants repeatedly chose between two symbols based on how much reward and effort was associated with each stimulus. In addition, there was also salient, but irrelevant, outcomes not dependent on participants' choices, namely whether the amount of reward from the current trial would be paid out as monetary reward (green tick) or not (red cross). Behavioural analyses (bottom left) revealed that participants' choices were mostly guided by the relevant dimensions (i.e. reward and effort amounts on the past three trials (t-1

to t-3)). However, they also misattributed the irrelevant outcomes: they were biased to choose an option again if it had led to a reward payout (red bar). This bias was potentially present because many areas in the brain were sensitive to the irrelevant reward outcome (pink, top middle column). We found several signals in frontal pole (FP) that might suggest that this area plays a role in overcoming the bias: FP carried a signal for the irrelevant reward payout and the people who had a stronger FP signal were less biased (bottom middle column). Additionally FP increased its representation of the information to be learnt (amount of reward and effort outcome) when reward was paid out thus potentially helping to overcome the bias (right column). D) Volman et al. [104] tested whether FP also plays a role for overcoming automatic emotional biases during approach-avoidance decisions in social contexts. On each trial, participants had to make an approach or avoid decision (i.e. move a joystick towards or away from themselves) after a very brief presentation (100 ms) of a happy or an angry face. Half of the actions were reflexive/automatic (i.e. approach-happy and avoid-angry) and the other half controlled (i.e. approach-angry and avoid-happy). Disruptive transcranial magnetic stimulation (TMS) over FP changed blood flow in bilateral FP. At the same time, TMS also selectively increased the rate of errors in the controlled condition. Interestingly, in a separate study [106], the same authors found that in psychopaths with particularly high testosterone levels, activity in FP in the controlled compared to the automatic condition was decreased. This suggests that maybe they are less able to control the impact of reflexive emotional information during rule-driven behaviour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

found that indeed, people were biased by these rewards and misattributed them to the stimuli, even though the optimal behaviour would have been to completely ignore them. Such a bias may exist because many brain areas are sensitive to reward [103]. On the other hand, we also found a network centred on the frontal pole which encoded signals that suggested it was trying to compensate for this bias. The frontal pole increased its representation of the relevant information when the salient information needed to be ignored, while also producing a signal driving behaviour to overcome the bias. In other words, when an irrelevant reward experience biased participants to choose one option, this brain area would produce a signal to change participants' preference towards the other option. These findings are in agreement with studies of emotional control during decision-making ([104–106], Fig. 5D) and with findings showing that lesions of this area and nearby areas in humans lead to misattributions of outcomes to irrelevant

dimensions [107]. Thus, in situations in which it is clear what brain mechanism should guide behaviour, the brain not only selects the most relevant mechanism, but actually over-writes, or counter-acts, learning from other inappropriate mechanisms. It will be interesting to investigate in the future how psychiatric disorders affect the abilities to flexibly arbitrate between different learning systems and to suppress inappropriate learning mechanisms.

2.7. Building beliefs about the world

So far, we have focused on diverse mechanisms for learning the reward value of stimuli. Another important aspect in naturalistic environments is of course learning about the structure of the world or, in other words, learning a cognitive map. More specifically, a cognitive map has information about how different states in the environment

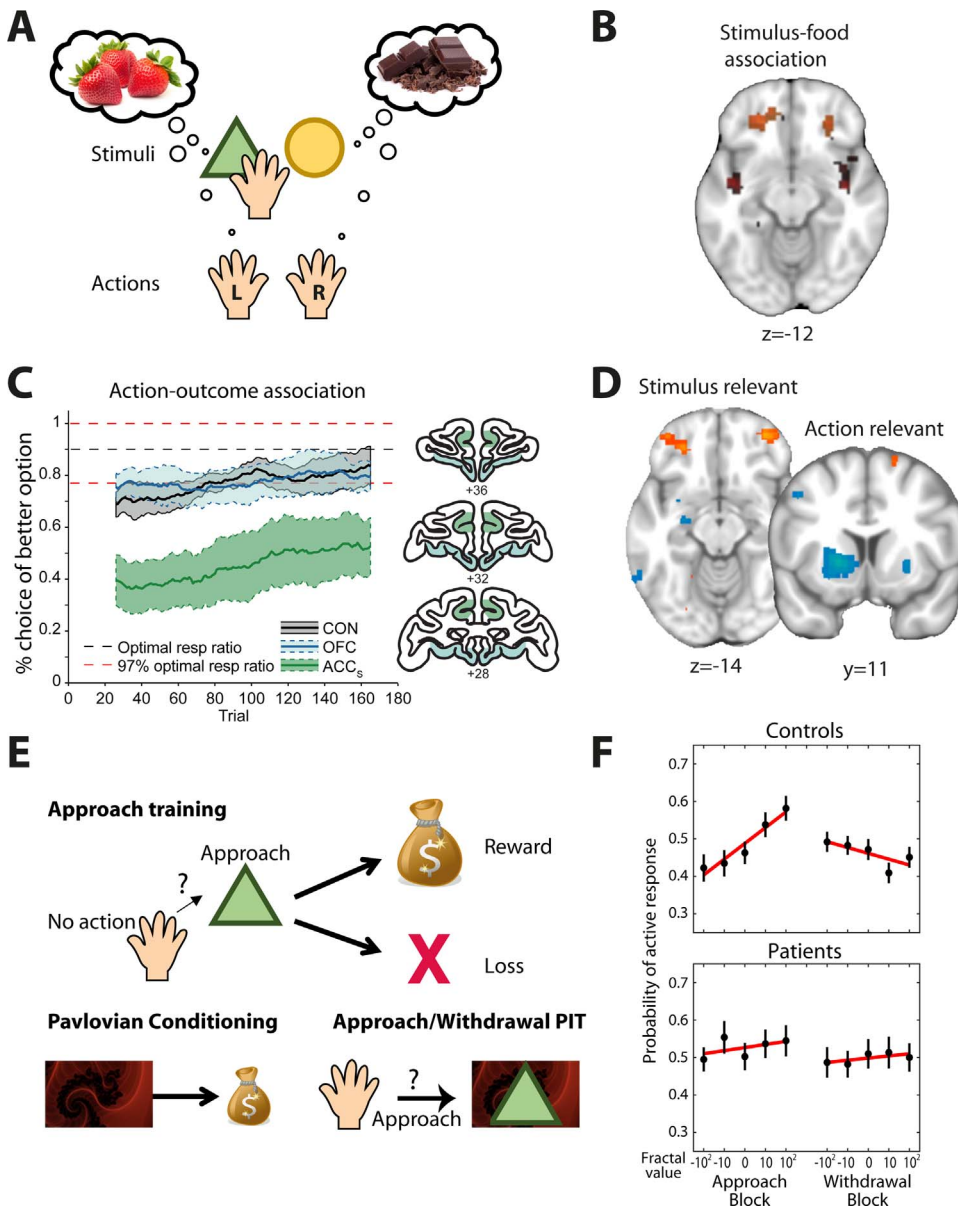


Fig. 6. Different brain valuation systems. **A)** Illustration of a simple binary choice task that assesses valuation processes when assigning value to stimuli (e.g. abstract shape, top) or actions (e.g., right or left hand, eye or joystick movement, bottom). In some paradigms information changes over time but it is also possible to vary the properties of the visual stimulus to signal the expected outcome which requires evaluating new information. For example, the colour and quantity of symbols might denote the juice type and quantity of juice as in [68]. **B)** Encoding associations between stimuli and reward relies on central orbito-frontal cortex (orange; red is irrelevant here). Here this was shown using repetition suppression [69] which provides a way to study neural representations using human neuroimaging [242]. **C)** In contrast, the anterior cingulate cortex (ACC) is critical for using action-outcome associations to guide choice as shown here using lesions in macaque monkeys (shaded areas show lesions in OFC and ACC sulcus respectively; adapted from [117] and [138]). **D)** When both actions and stimuli are relevant for guiding choices, interactions with several valuation networks take place (interactions with parietal cortex when actions are relevant (blue: striatum) or stimuli are relevant (orange: OFC)). Adapted from [148]. **E)** Sometimes, new information elicits reflexive changes in value. In the Pavlovian-Instrumental-Transfer (PIT) task [169], participants first learn whether or not to approach visual stimuli (triangle) to obtain reward and avoid a punishment (in separate blocks, participants had to learn whether or not to actively withdraw from a stimulus to get a reward and avoid punishment; not shown). In the second stage, they passively view several Pavlovian stimuli (fractal) followed by tones predictive of wins or losses (-100, -10, 0, 10, 100). In the test phase, the measure of interest is how the Pavlovian stimulus (fractal) affects decisions about whether or not to approach/withdraw from the previously learnt instrumental stimuli (triangle). No outcomes are shown in this phase. **F)** The optimal behaviour would be to ignore the incidental Pavlovian stimuli. However, the pattern usually observed in healthy controls is that in approach blocks (top left), participants are more likely to approach the instrumental stimulus if the Pavlovian stimulus in the background has a higher positive value than if it has a negative value. The opposite behaviour is seen in avoid/withdrawal blocks (top right). This rudimentary response was absent in depressed patients (bottom) (adapted from [175]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

relate to each other and how different stimuli should be processed in different states. For example, you can think of different ways to interpret a friend ignoring you, depending on your internal model: you could think 'my friend does not like me', based on a model, or a negative self-view, made up of a whole range of connected statements such as 'I am worthless' or 'people don't like me'. Or you could think 'my friend didn't see me', based on a different model that includes statements such as 'how other people act often has more to do with their state than with me'. The states, with their specific cognitive content, in this kind of model, directly relate to the concept of 'schemas' in psychiatry, which have been used to explain aspects of depression. For example, Beck's theory of depression proposes negative schemas, i.e. sets of beliefs about the self and the world and how they relate to each other, as a problem in depression [108–110]. In this theory, people possess different schemas (e.g. 'I'm great' and at the same time 'I'm a failure') and they can become active at different times. In depression for example, stressors are thought to activate negative schemas that then shape

perception and action and thus trigger a depressive episode. A person who had been depressed in the past may possess the same schemas but they would lie dormant.

Neurophysiologically, the vmPFC/mOFC system (Fig. 4) has been related to hidden states and inferential models, i.e. schemas that determine how to process stimuli, or a cognitive map that can be searched through or employed in simulations [111–115]. For example, Schuck et al. [115] designed a task in which participants needed to make decisions about stimuli based on a model of the task, i.e. knowing the different task states that they had previously learnt (e.g. is feature A or feature B of the stimulus relevant for your decision?). To make the decisions, they needed to infer which (hidden) task state was currently relevant. This information was represented in vmPFC/mOFC, thus supporting the view that this region has a mental model/map. As an aside, we also note that this is not the only type of mental model and other research has highlighted the dorsal anterior cingulate cortex in the context of learning the structure of the world and using this

knowledge to explore the environment [116–126]; but this is beyond the scope of this review.

While there are intuitive links to be drawn between mental maps of the sort studied in these experiments and the schemas invoked in theories of depression, the kind of paradigm described above has not yet been used to test whether depressed and healthy controls differ, for example, in how flexibly they can shift between different task schemas or whether they have biases towards certain kinds of schemas.

3. Valuation and decision-making

3.1. Using new information and integrating multiple sources of value

Thus far in this review, we have focused on different learning mechanisms and on how they might influence our choices. Yet, in everyday life, to guide behaviour we often need to integrate *learnt* information with *newly perceived* information (Fig. 1, left-hand side). To illustrate this, in the example about enjoying the company of your friends, imagine deciding whether to go out for dinner with them. This not only depends on how much you have recently enjoyed their company, but it may also depend on what the restaurant has on the menu today, the cost of the food or your appetite. Here we will first focus on the valuation of new information, and the integration of costs; the decision process will be covered later in this section. Studying valuation processes is of great importance for psychiatry, as patients commonly have motivational problems in specific contexts, possibly because they do not value them appropriately. For instance, lack of energy or fatigue is a core symptom of depression [127] and this may impact on choices requiring physical effort, e.g. deciding to spend energy to get to the restaurant, but not other types of choices. Furthermore, different brain regions process different types of value information. This, in turn, may mean that different groups of patients, with different underlying pathology (e.g. for depression see [128]), experience differing types and degrees of impairment depending on the nature of the valuation problem probed by a given task.

3.2. Not all value is equal

It might be tempting to think of value as being one singular currency that predicts a person's preference. However, there is now ample evidence to show that the brain contains multiple valuation systems [129,130]. Broadly speaking, the value of an outcome will be represented in the brain regions that process the information the value is constructed from (Fig. 6A–D). For example, evaluating a food item involves processing its taste, visual appearance and smell, and these sensory inputs converge in the central orbitofrontal cortex (cOFC, Fig. 4) [131], a region critical for processing the value of food items. In experiments, a reward is often predicted by a visual cue displayed on the screen (a Starbuck's sign might make you think of coffee) and evidence from lesion, recording and neuroimaging studies demonstrates cOFC's essential role in representing multiple dimensions of stimuli such as the associated food type or value (Fig. 6B) [68,69,132–140]. In contrast, an outcome can also be predicted as a result of performing a certain action, rather than from a visual or sensory cue. For example, finding out whether your best friend is coming along for dinner involves picking up the phone, and thus assigning value to the result of an action plan. Typically, in tasks investigating action value representations, different actions, e.g. eye movements or joystick movements lead to different amounts or types of reward. Evaluating rewards that are tied to actions involves a different network of brain regions to the one described above, including the dorsal striatum and anterior cingulate cortex (ACC) [141–144]. Lesions to ACC impair choices guided by action values (but not stimulus values) in macaques and human patients (Fig. 6C) [138,145]. This is consistent with the connectivity profile of ACC; it has much weaker sensory inputs than OFC but more direct projections to premotor and motor cortices [146,147]. Compellingly,

when multiple attributes (e.g., actions *and* cues) are relevant for evaluating a choice option, interactions between several of the brain regions we have discussed take place (Fig. 6D) [76,148]. Thus, *where* in the brain new information is evaluated will depend on *what* the value is assigned to (a food outcome, an abstract cue, an action etc.). This is worth bearing in mind when designing tasks for specific patient populations, as they could manifest problems of value corresponding to the functional circuit that is affected by their pathology. One type of value we have not touched upon here but which may have great relevance for psychiatric disorders is the value of social information. Many real-life situations involve several individuals and thus, it is important for tasks to manipulate social context and the type of social information that needs to be processed. This is not covered here due to space constraints (but see [149–151]).

So far, in depression, a range of valuation deficits have been reported but overall the conclusions have been quite mixed [152,153]. Basic tests such as the sucrose liking test [154,155] show no difference between healthy controls and patients suffering from depression. Similarly, ratings of cartoon images with emotional content were found to be unchanged in depression [156] and a recent study deploying a computational model to assess valuation of gambles in the absence of learning found no difference between patients suffering from depression and controls [157]. By contrast, other studies have found reactions to emotional stimuli to appear blunted in depression. For example, sad movies did not seem to trigger the same increase in sad feelings in depressed patients as in controls [158]. While it was originally thought that patients suffering from depression are overly sensitive to negative feedback and show a catastrophic response to failure [159,160], a more recent study suggested instead that there might be a diminished sensitivity to emotionally negative information which leads to less adjustments in subsequent behaviour, meaning unlike controls, patients would not correct their performance as much after error feedback [161]. Evidence from a meta-analysis including a wide range of measures (self-report, behavioural, physiological) suggest that this reduced response to emotional stimuli is present in both the positive and negative domain [162]. Still, overall results have been mixed which could be related to several reasons. These include the particular valuation system probed by a task, the large variety of biotypes of depression [128], or alternatively, the way participants 'report' value. For example, studies relying on reflective valuation processes (e.g. when participants give ratings or make deliberate choices [156,157]) may recruit different brain systems than those relying on automatic valuation processes (e.g. approaching or avoiding; see next section). All of these are open questions that will need to be addressed by future work.

3.3. Reflective versus automatic valuation

There are not only different types of value for e.g. actions and stimuli, but also differences in terms of cognitive accessibility [163]. The valuation processes discussed thus far are largely 'reflective' in the sense that they involve thinking about and imagining possible outcomes, as well as considering which cues are relevant for achieving the outcomes. The goal of this reflective process is to act in a goal-directed way. However, sometimes a new piece of information elicits a more automatic, and thus less controlled, change in value. If, for example, a spider unexpectedly walked across the restaurant table when you are out with your friends, you may not enjoy your food as much. Considering this situation from the perspective of reflective (or model-based) reasoning, what has gone wrong here is that – assuming the spider disappeared again and there is no action you currently need to take – the stimulus (the spider) is irrelevant to your current task (eating) and should therefore not influence your choices. While this situation illustrates that it is not always appropriate for values to become active and influence behaviour, these types of automatic influences on behaviour are ubiquitous, raising the question as to why this would be. As considered in depth in previous reviews [94,96,163,164],

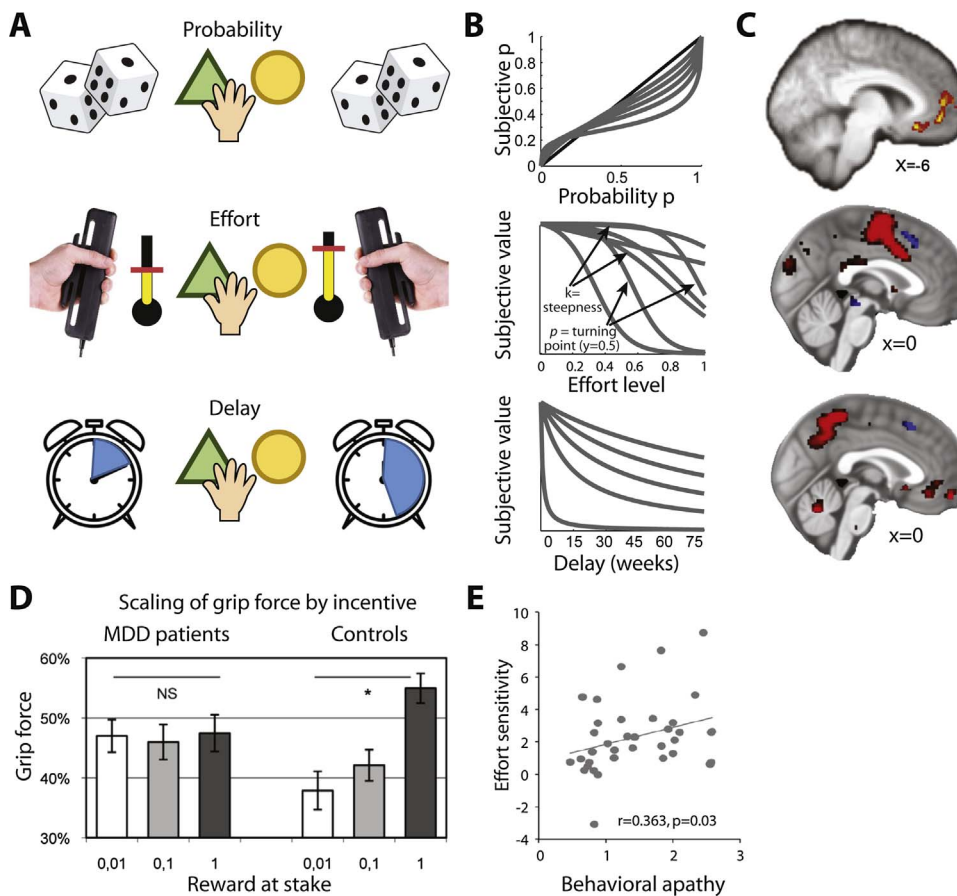


Fig. 7. Mechanisms for integrating costs into value. **A)** Illustration of binary choice tasks involving different types of costs. In all cases, choices are made between two options, A and B, associated with varying quantities of a monetary or food reward (displayed or pre-learned). Importantly, the reward comes at a cost which could be the ‘risk’ associated with winning the reward (varying probabilities, top), the physical effort that needs to be exerted to obtain the reward (e.g., grip force, middle), or the delay before which the reward will be received (bottom). **B)** To model the integration of costs and benefits into subjective value and capture an individual’s discount preference – e.g., how much the subjective value of an option with a fixed reward decreases as a function of different cost levels – simple behavioural models are used. The parameter(s) fitted for each individual explain how the value of reward decreases with increasing costs. *Top:* for probability, prospect theory provides the standard model accounting for the over-weighting of small and under-weighting of large probabilities that most individuals exhibit [243–245]; *middle:* for effort discounting, the best fit is achieved using an initially concave function, for example an inverse sigmoidal (shown here) or parabolic/quadratic function (not shown) [246,247]; *bottom:* for delay discounting, a convex hyperbolic model is appropriate [248,249]. The different shapes of the discounting functions suggest that different types of costs affect choices differently: for example the concave shape for effort discounting meant that people care not much about whether they have to make a small effort or no effort, but they care more when the choice is between a small and a medium effort. In contrast, for delays, people will care about whether a reward will be paid out immediately or only in an hour, but they care less about whether reward is received in three weeks, or three weeks

plus an hour. Having a model of subjective value provides several critical advantages. First, it provides sensitivity to inter-individual differences. Second, it can be used to capture the influence of cost in situations involving more than one relevant variable. Third, fitted parameters can be used to examine relationships with other behavioural or clinical markers. Finally, neuroimaging data can be explained in ways that would otherwise not be possible, for example by looking at the representation of value difference as a marker for choice. **C)** Three exemplar studies looking at the encoding of value (i.e. the difference in value between the chosen and unchosen option; red: activation; blue: deactivation) for the three different types of costs (top: probability; middle: effort; bottom: delay). This highlights distinct networks depending on the type of cost (red), while the inverse contrast is consistently encoded in dmPFC (blue). Adapted from [101,185,189]. **D)** The effort (here grip force) exerted to get a reward scales with expected reward (x axis) in controls but not depressed patients. Adapted from [199]. **E)** Participants performed a task where they accept or reject a reward given the required effort. A model captured each individual’s effort sensitivity, or in other words, how much weight they placed on the effort in their choices. This parameter was directly related to an individual’s apathy trait with higher apathy relating to higher effort sensitivity. Adapted from [197]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

the advantage of automatic/reflexive systems is that they are computationally efficient, and therefore available more quickly, which can be important in many real-life situations. For example, if a bear is approaching you, running away quickly out of fear is more adaptive than carefully considering how likely it is that the bear will hurt you. One way in which automatic influences on valuation can be studied is in tasks measuring approach and avoidance tendencies either without ([104], Fig. 5D) or with a learning component [165]. These approach/avoid decisions recruit a network centred on amygdala and subgenual ACC (sgACC) [166], a prefrontal region with a high density of amygdala inputs [167], for review, see [168]. Reflexive influences on behaviour can also be measured using Pavlovian-Instrumental-Transfer (PIT) tasks where an incidental appetitive or aversive Pavlovian cue that is irrelevant to the task at hand (e.g., the spider) increases the likelihood to approach or avoid another stimulus (e.g., your food), respectively [169] [e.g., 169] (Fig. 6E). In both types of tasks, there is a strong link between valuation and action and the effect of the incidental cue is measured in terms of action (approach/avoid). Interestingly, the brain networks driving the influence of Pavlovian cues on value and behaviour are primarily phylogenetically older structures, in particular striatum and amygdala [170–173]. Automatic valuation signals have also been identified in neo-cortical regions such as vmPFC (e.g. [102,174] (Fig. 5)). Altogether, this suggests that the strength of a reflexive bias (or value) is encoded separately from the action that it

biases.

Reflexive valuation mechanisms and their interaction with more reflective systems are of particular interest for psychiatry, and they might relate to aspects of cognitive therapy. Huys et al. [175] used a PIT task to examine how Pavlovian cues influence valuation in depression (Fig. 6F). In healthy participants, a positive incidental cue promoted approach and a negative incidental cue promoted avoid responses, respectively. However, this rudimentary response was absent in depressed people. As the authors pointed out, one potential consequence of this might be that in depression approaching a positive or avoiding a stressful situation might rely on more computationally expensive (‘reflective’) neural mechanisms and thus feel more effortful. If this is the case then it may begin to explain some of the symptoms of depression such as anhedonia and increased exposure to stressful life-events. Future studies could build on this work to test the idea that in depression Pavlovian approach-avoid tendencies are reduced, using for example a recent task proposed by Bach and colleagues [176] where Pavlovian impulses for approaching and avoiding produced helpful behavioural responses rather than unhelpful behavioural biases as in PIT tasks. Overall, it seems that in depression the interaction between a more automatic and a more reflective valuation system are off-balance, with a reduced influence of more automatic valuation mechanisms [175]. In line with this, lowering central serotonin levels using acute tryptophan depletion had similar effects, reducing approach/avoid

behaviour, yet only in the aversive domain [177]. And when deciding whether to approach or avoid an outcome, stimulation in a ventral zone of sgACC induced depressive phenotypes (negative biases) and altered approach/avoid type choices but not choices between two options [166]. All of these findings are particularly interesting given depression is associated with structural and functional brain changes in the amygdala, ventral striatum, and sgACC [sgACC; 43,161,178–181]. In other words, both neurophysiological and behavioural evidence point towards reduced use of reflexive value and increased reliance on reflective computations, which could together reduce motivation and lead to rumination. In related work, Eldar et al. found that people with high mood instability (a risk factor for bipolar disorder), valued monetary rewards differently as a result of an incidental event (large win or loss) that affected their current mood [182,183]. Specifically, participants had to learn how likely different slot machines (i.e. stimuli) were to give rewards. After learning about half of the slot machines, participants experienced a large win or loss in a wheel-of-fortune draw, which affected participants' mood. However, whether they won or lost was not related to the next set of slot machines that participants subsequently went on to learn about. Despite being aware of this, participants with high mood instability behaved as if the slot machines after a large reward were better than the ones after a large loss. This is another example of automatic processes interfering with goal-directed processes.

3.4. Integrating costs and benefits

So far we have focussed on the valuation of rewarding outcomes, or benefits, and aversive outcomes, such as losses. However, in natural environments, valuation processes frequently entail an integration of costs and benefits. The types of costs typically encountered include temporal delays (e.g., having to wait for your friend to pick you up), motor costs (e.g., having to walk to the bus stop), and risks (e.g., getting caught in a traffic jam). Tasks looking at how a given cost affects the evaluation of a rewarding outcome usually offer participants two options that are each associated with different levels of cost and reward (Fig. 7A). Alternatively, sometimes only one option varies in reward and cost and another option stays constant across trials ('default') (for effort costs see e.g. [184–186]). Simple behavioural models can be used to capture an individual's propensity to be influenced by a given type of cost and the most commonly used examples of such 'discounting functions' are shown in Fig. 7B.

The brain networks recruited to encode and integrate different types of costs again differ depending on the type of information that needs to be processed. For instance, the vmPFC encodes information about subjective value when probability (i.e., risk) and magnitude need to be integrated to evaluate an option (Fig. 7C, top) [e.g.,101,187]. When evaluating whether a reward is worth waiting for, a similar but sometimes slightly more dorsal and posterior perigenual ACC (pgACC/vmPFC) region encodes subjective values [188–190] (Fig. 7C, bottom). By contrast, when physical energy is necessary to obtain an outcome, more dorsal regions in ACC/MCC, with more direct projections to premotor and motor cortices, are essential for encoding the integrated cost-benefit value (Fig. 7C, centre) [102,117,185,191–197].

Valuation deficits in depression may possibly be found most consistently when physical effort is involved. In other words, when there is a need to energize behaviour through self-motivated actions and thus express how much it is worth working for an outcome. For instance, going out with your friends requires mobilizing some energy to get to the restaurant. Unlike in some of the domains of learning and valuation covered above, there is converging evidence that this process might be altered in depression, and many symptoms of depression, such as lack of energy, fatigue or decreased engagement in activities, relate to this. For instance, healthy controls produce more effort for higher reward incentives, but depressed individuals show overall reduced levels of effort production [198] and no scaling of effort with expected reward

size ([156,199], Fig. 7D). This nicely dissociates liking an outcome from being willing to produce motivated behaviour to obtain it. Changes in the willingness to exert effort for reward are also seen in patients suffering from apathy ([200,201], Fig. 7E), and apathy has been found to be strongly associated with symptoms of depression in a large online cohort [47]. Several things about the existing work on effort in depression are striking. First, it remains quite unclear where exactly the deficit lies: in the mobilization of effort *per se*, the integration of rewards and costs, or some aspect of the reward itself. Second, it is unclear how specific this deficit is to physical, as opposed to other types of costs, e.g. mental effort or temporal delay. Finally, and most importantly, the effect sizes reported in laboratory studies are small compared to the deficits observed in real-life (e.g., for energy loss see [202]). This suggests that experimental paradigms need to better capture the complexity of real-life situations, and combine this with careful computational modelling. For example, many real-life decisions are about whether to continue with a default behaviour or make an effort to look for alternatives (e.g. deciding to stop relaxing on the sofa to go out, see Fig. 1B). Instead, the situation that is often encountered in laboratory tasks is to face two concrete options (with explicit rewards and efforts) to choose between (see also Fig. 9B).

3.5. Integration of different sources of value: the role of different neurotransmitter systems

Another critical question for decision-making is how much to rely on different types of value, for instance information we have learnt over time in contrast to newly perceived information. For example, if a new cook has started to work in the restaurant that your friends are taking you to, your previous experience should influence your choice to go out with them less. In addition to asking whether different brain areas encode different types of value, there has been interest in whether different neurotransmitters and neuromodulators influence *how* this integration occurs. This is especially important for psychiatry as many treatments are drug-based and involve particular pharmacological targets.

The neurotransmitter noradrenaline (NA), is a good candidate for regulating how much to rely on learnt information (e.g., given past experience, how likely am I to enjoy the time with my friends). In rats, increasing the influence of NA on ACC activity and thus decreasing ACC activity, caused animals to abandon prior knowledge, whereas silencing NA had the opposite effect, making them more reliant on learnt information ([122], Fig. 8A). In humans, magnetic resonance spectroscopy (MRS) can be used to measure the concentration of the two main excitatory and inhibitory neurotransmitters, glutamate and GABA. Using MRS, we were able to show that the balance of excitation and inhibition (E-I balance) in dACC regulates how much choices rely on learnt versus newly perceived value information ([126], Fig. 8B). Higher levels of glutamate and lower levels of GABA were associated with both increased strength of the value to be learnt in dACC and increased use of learnt information over new information when making choices. Taken together, these results suggest that the dACC plays a crucial role in allowing information that has been learnt to influence behaviour (see also section 'Building beliefs about the world'), and that, on a molecular level, this may be realised by regulating the E-I balance which in turn is possible via the noradrenergic system. This is interesting because the E-I balance can be altered in acute stress [203] and when manipulating serotonin levels [204], as would typically be done with SSRIs in depression. Some reports also suggest that glutamate/glutamine and GABA levels, specifically in ACC, are reduced in depression [205,206] but there is mixed evidence regarding this possibility [207,208]. It is therefore plausible that such neurochemical changes could affect how different aspects of value are integrated in these patients. Indeed, we have recently made the observation that patients with dysphoria use learnt information less, compared to new information, when making decisions (unpublished).

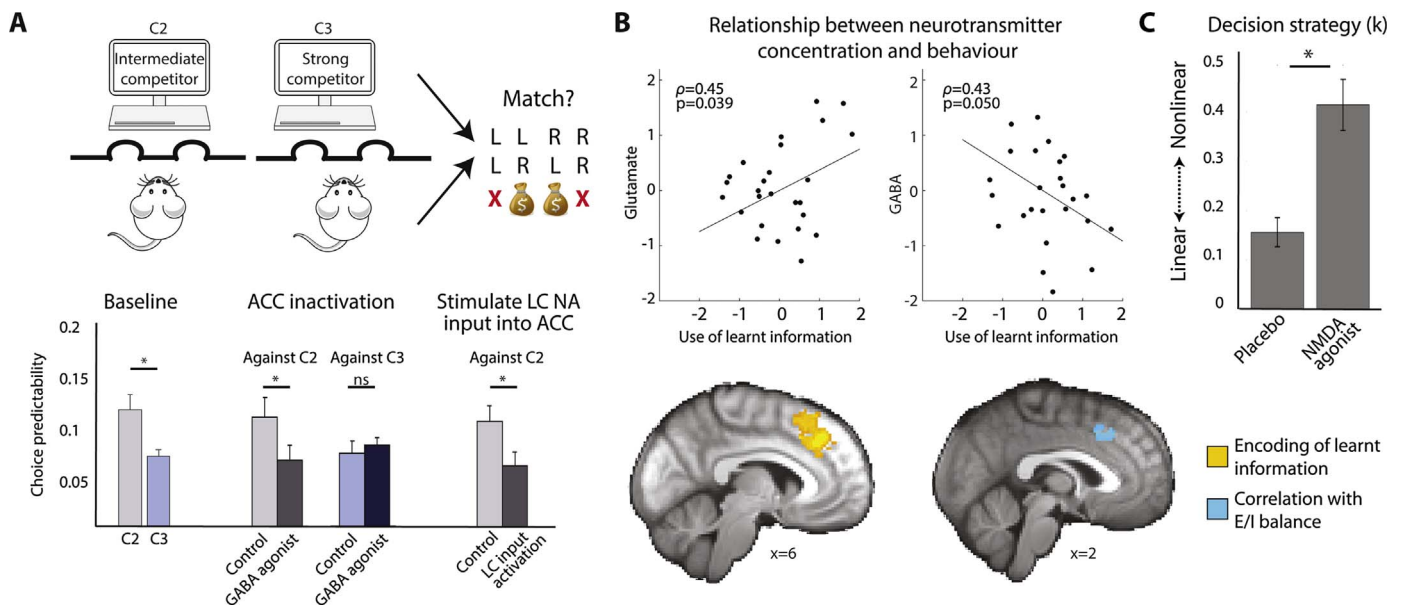


Fig. 8. Molecular mechanisms for integrating information.

A) Top: Rats are playing against either an intermediate (C2) or a strong competitor (C3) who is trying to predict their choice between two food ports. Animals are only rewarded on trials where the computer does not match their choice. The difference between the competitors is that the strong competitor is better able to detect any patterns in the animal's choices and exploit them. In other words, if there are any statistical regularities, i.e. if the rats' behaviour is predictable, the competitor will use this information to prevent the rats from getting rewards. Bottom: Playing against a stronger competitor leads to more random choices (decreased choice predictability). Inactivation of anterior cingulate cortex (ACC) using a GABA agonist produces increasingly random choices in rats playing against C2, but not in the rats playing against C3 who already exhibit strongly random choice behaviour. This effect is mediated via noradrenergic (NA) input from the locus coeruleus (LC) onto ACC because the same effect is observed when LC inputs to ACC are stimulated pharmacologically or optogenetically. This suggests that the level of NA in the ACC controls the balance between using a model (trying to predict the opponent) and random choices. Adapted from [122].

B) Magnetic resonance spectroscopy (MRS) was used to examine the balance of excitation (glutamate) and inhibition (GABA) in the dACC when choices relied on learnt and explicitly cued information. Top: A model parameter that captured how much participants relied on learnt relative to new information showed a positive relationship with Glutamate and a negative relationship with GABA. Bottom: dACC encoded the information to be learnt and this signal increased as a function of the E/I balance in the region. Thus, consistent with [122], dACC controls how much learnt information influences behaviour and this might be achieved by regulating its E/I-balance. Adapted from [126].

C) In a similar task, a partial NMDA agonist made participants rely more strongly on an optimal non-linear strategy for combining different pieces of information to calculate integrated value (here reward probability * magnitude). Choices in the placebo group were guided predominantly by the simpler linear computation (here reward probability + magnitude). Adapted from [209]. This suggests that NMDA receptors play a role in non-linear integration of information.

In addition to asking whether learnt (or other) value information should be weighed more or less strongly when making a choice, it is also important to think about how information is combined and how this might be influenced by neurotransmitter levels. In other words, how does the brain implement linear or non-linear combinations for forming an estimate of integrated subjective value from, e.g. reward and probability (see discounting curves for combining costs and benefits in Fig. 7)? We tested this in another study ([209], Fig. 8C), and quantified the degree participants relied on a linear or non-linear strategy for combining values with a computational model. We found that a partial NMDA agonist made participants rely more strongly on the optimal non-linear strategy while choices in the placebo group were guided predominantly by the simpler linear computation. This is in agreement with previous findings that NMDA receptors facilitate non-linear integration of information in e.g., multisensory integration [210]. Our data suggests that this role extends to the value domain. While in this study, the integration happened between learnt and newly cued information, it is plausible that a similar mechanism of integration might be recruited when several explicit cues need to be integrated non-linearly to compute value – e.g., the delay and size of reward or effort and size of reward.

3.6. Decision-making: choice stochasticity

The previous sections have focussed on mechanisms of learning, the valuation of new information and integration of different pieces of information. In this section, we focus on the decision or selection process itself. Indecisiveness, or a greater difficulty in making choices, is a core symptom of depression (e.g., [211]). Depressed individuals tend to avoid making decisions, show maladaptive decision-making, and

enhanced stress levels when making choices [212–216].

It seems straightforward that once the integrated values of different choice options have been computed, for example the values assigned to different options for spending your evening, the option with the highest subjective value should consistently be chosen. This is not, however, what is usually observed. For example, in a simple situation where one option gives reward 80% of the time, and a second option only 20% of the time, participant's choices generally match these probabilities in that option 1 is chosen 80% of the time. However, to maximise reward, option 1 should be chosen 100% of the time. This raises the question of where this randomness, or stochasticity, in the observed choices comes from. Many different possible sources might contribute, such as the desire to explore other options in case their value has changed; uncertainty about underlying value estimates and the optimal way of integrating different aspects of value (covered above); priors that skew values in one direction; and mistakes, for example due to distraction, tiredness, carelessness. Below we will first describe how choice stochasticity is usually accounted for in computational models before discussing why it may provide an advantage in ecological settings.

When using computational models to explain participants' choices in tasks, it is important to account for choice stochasticity. This is typically done by fitting a parameter, referred to as the softmax inverse temperature, for each participant which given the value of an option, provides an estimate of the likelihood of choosing this option (Fig. 2B–D). Work using MRS in humans has shown that the degree of stochasticity relates to the E-I balance in the ventromedial prefrontal cortex, a brain region with a role in comparing choice option values [217]. Relatedly, causal manipulation of the E-I balance in this region (using depolarizing transcranial direct current stimulation (tDCS)) leads to increasingly stochastic choices (Fig. 9A, [218]). Both of these results make sense in the context of current models

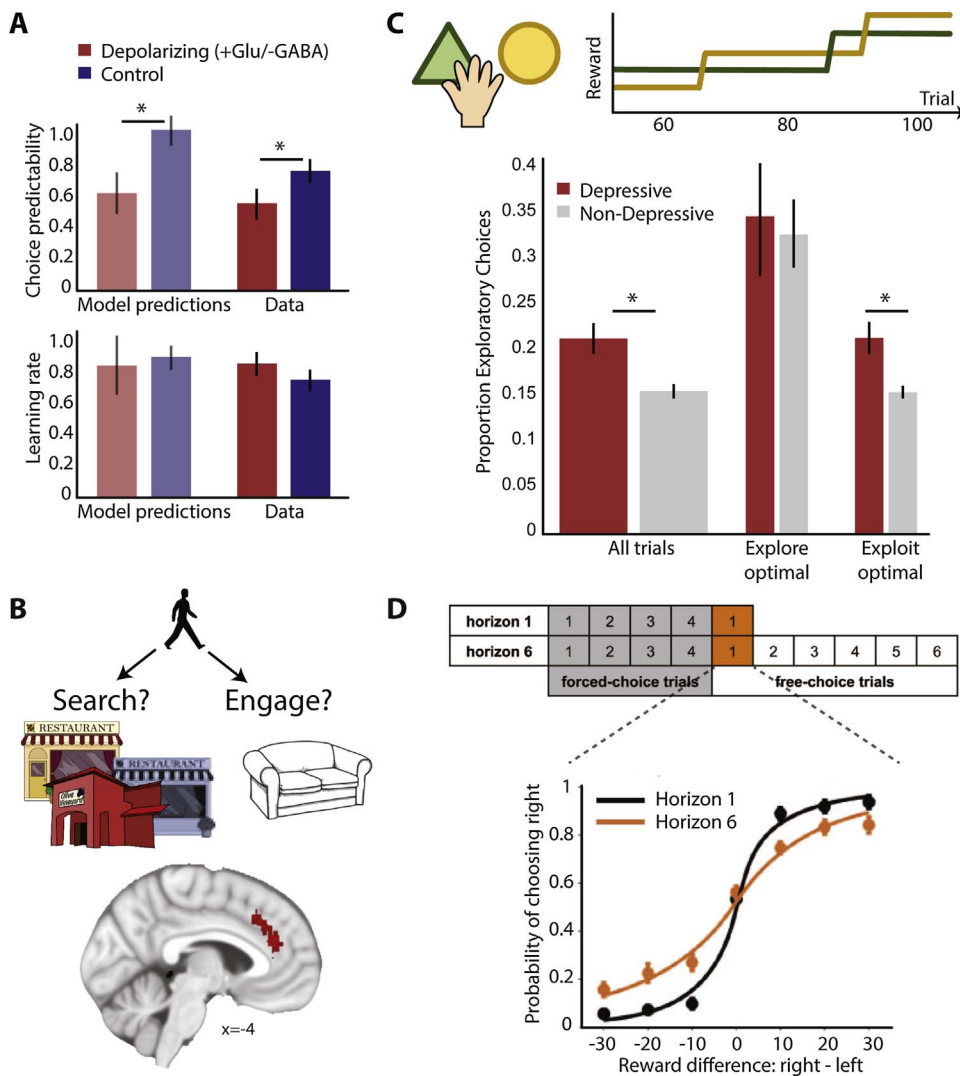


Fig. 9. Neural mechanisms underlying exploration. **A)** Participants performed a binary choice task in which the probability of reward had to be learnt while reward size was cued on the screen. Transcranial direct current stimulation (tDCS) targeted the vmPFC, a region critical for comparing the values of options when reward magnitude and probability need to be integrated. Anodal tDCS is thought to depolarize the underlying pyramidal neurons thus causing a shift in the E/I balance towards more excitation [250,251]. tDCS led to more exploration (smaller softmax inverse temperature) in agreement with predictions generated using a biophysical model (top). In contrast, learning (measured as learning rate in the model) was not changed (bottom). This dissociation would not have been possible without a learning model and suggests that the degree of exploration in this task is regulated via the balance of excitation and inhibition in vmPFC. Adapted from [218]. **B)** Example study in which choices were not framed as being between options A and B but between whether to engage with a currently present option (or in other words a 'default') or to explore for other opportunities). Studying the neural circuits underlying this ecological type of choice revealed the dACC as the key region representing the value of searching/ exploring. Adapted from [8]. **C)** Choices between two options are made and one of the options is consistently better until the alternative increase in value ('jumps up') at an unpredictable time (Leapfrog task). Exploratory choices (i.e., checking whether the alternative has changed) should become more frequent the more time has passed since the last 'jump'. Students with symptoms of depression were overall more exploratory (or in other words less consistent in picking the option with higher value, i.e. their behaviour was more random). However, most importantly, a closer look at when the exploration happened showed that depressed participants did not explore more on trials when exploring was advantageous (to check for 'jumps') but instead they explored more at times when exploiting would have been the optimal behaviour. Adapted from [29]. **D)** Wilson et al. [224] designed a task to measure whether humans adjust how much they explored, depending on how useful it was to do so. In each game of the task,

participants were presented with two slot machines. In the first four trials of each game they could not choose themselves which slot machine to play, but instead the computer selected an option for them ('forced choice trials'). Participants were only shown how much they won from the machine they played, not how much they could have won from the alternative. The key manipulation was that after these forced choices participants were given either a single trial ('horizon 1') or six trials ('horizon 6') on which they could choose freely between the two slot machines ('free choice trials'). In horizon 6 compared to horizon 1, it was thus more valuable to gather information about which slot machine was better because more choices were left where this information could be exploited. And indeed (bottom panel in D), participants were more likely, at the first free choice, to explore (and thus select the slot machine that had so far been less valuable) in horizon 6 compared to horizon 1 (see also to Fig. 2D).

whereby the competition between several choice options is resolved via mutual inhibition [219,220]. An increased concentration of GABA, and decreased concentration of glutamate, would imply increased levels of inhibition, a slower more precise comparison process and thus less random choices [217]. By contrast, increased levels of excitation will make the comparison process converge faster, thus generating more stochastic choices [218].

Interestingly, despite a relatively good neuro-computational understanding of stochasticity in decisions and decision-making problems being central to depression with obvious real-life problems associated with it, there are few studies showing changes in choice randomness in the lab, and most do not dissociate valuation from choice stochasticity in depression. Some reports find increased reaction times in simple binary choices in depression without changes in accuracy [221]. Changed reaction times could point at changed decision making as current decision models make predictions about both choice and reaction time patterns. This effect could, for example, be due to blunted valuation, in essence, making the choice harder because values seem more similar. However, if this is the case then decision-making should also be less accurate in depressed patients but this is not the case.

Alternatively, if valuation is unperturbed, increased reaction times can be achieved via increased inhibition/reduced excitation in the comparison process, but such changes would predict depression-related improvements in accuracy, which are not observed either. Interestingly, Huys et al.'s meta-analysis [46] reported reduced reward sensitivity in depression by fitting computational parameters to a reward-learning tasks. Because in the model, reward sensitivity could not be dissociated from stochasticity per se, this provides evidence for increased decision randomness in depression. In contrast, another recent study [157] that used a decision-making task, based only on explicitly shown, rather than learnt values, found no effect of depression on choice stochasticity. Overall it thus remains unclear whether the decision deficits reported in depression are truly deficits in the comparison process. However, maybe instead of more fine-grained measurements of choices and reaction times, we should start thinking differently about the reasons why a decision-maker might not always choose the best option in the first place. Such choices may not simply reflect stochasticity in the decision mechanism itself but instead they may reflect the operation of distinct goal-directed processes. In the next section we focus particularly on mechanisms for exploration.

3.7. Choice randomness in ecological environments enables exploration

It is important to consider why humans and animals would have developed a natural tendency to produce stochastic choices. What seems like suboptimal or noisy behaviour in simple laboratory settings turns out to actually provide an advantageous behavioural strategy in ecological settings, emphasizing once more the need for ecologically valid task designs. For example, the probability matching behaviour described earlier is the same as that found when an optimal Bayesian model is supplied with ecologically valid prior beliefs, such as a prior that events happening close in time are not unrelated [222]. Another aspect of natural environments is that they are not usually stable in terms of the outcomes predicted for a given choice. In an environment where friendships develop or the cook can change, it is optimal to occasionally explore the value of alternatives (e.g., how to spend your evening) to maximise overall gain over the longer term [223]. Few studies thus far have dissociated random choices from targeted exploration. However, Wilson and colleagues designed a task ('Horizon task'; Fig. 9D) that allowed dissociating random choice from purposeful exploration and showed that humans explored more when it was more useful to do so [224]. Furthermore, directed but not random exploration was affected by transcranial magnetic stimulation (TMS) over right frontopolar cortex [225] suggesting different neural systems support the two types of exploration. Other work in which choices were encoded in a frame of exploring versus exploiting showed that dACC holds a representation of the value of exploring the environment ([119,226], Fig. 9B). Not only does activity in dACC change when exploratory choices are made but when the outcomes of exploratory choices are evaluated [120,227,228]. Choice stochasticity, or in other words unpredictable choices, can also confer a critical advantage when faced with an opponent who is trying to predict our behaviour, and inactivation of rat ACC increases the degree of randomness ([122,229], Fig. 8A). Designing tasks that make it possible to determine the degree to which choice stochasticity is strategic, whether as a way of exploring and refining the model of the world or to confuse a competitor may be critical for understanding changes in decision-making in psychiatric disorders.

To our knowledge, only one study so far has distinguished targeted from random exploration in depression; this work does suggest that depressed patients may show altered exploratory behaviour. Blanco et al., [230] used the 'Leapfrog task' where the reward obtained from two options can only ever increase over time (Fig. 9C). The inferior option can jump up at unpredictable times, so that on a given trial the choice is between exploiting what is thought to be the best option, and exploring whether the previously inferior option might have exceeded it. In this task, exploration should be structured, with an increased likelihood of exploration as time since the last exploratory trial passes. Indeed, the majority of control participants' behaviour resembled such a pattern, while half of the patients suffering from depression (and particularly those with the highest levels of depression) lacked specifically this exploration strategy. Interestingly, overall, depressed patients produced more choices that could be seen as exploratory, but on trials when they should have exploited the current option. Thus, this study nicely illustrates the importance of ecological designs and of using behavioural models to better capture people's behaviour, and suggests that goal-directed exploration might be altered in depression.

3.8. Perils of the ecological approach and possible solutions

While we have tried to outline the various advantages of probing cognitive processes with tasks that incorporate ecological features, there are of course also some caveats that researchers need to be aware of. One concern is the increased complexity of ecological tasks, which may mean that a single behaviour (e.g. pressing a button) is now influenced by several different factors. This stands in contrast to more classical designs where all but one factor are kept constant. The

question that arises in ecological tasks then is how to dissociate different cognitive processes and define their unique contributions. Here we propose that this problem can be addressed by using computational models of the processes recruited by a given task (see for example [231,232] for an introduction). As one illustrative example, we can consider a study on reinforcement learning in depression by Huys et al. [46] (see also section on 'basic learning processes'). The authors showed that by using a computational model, a simple behaviour (i.e. choices on each trial that could be summarized as proportion correct) could be parsed into separate components, namely the speed of learning (learning rate) and the choice stochasticity (Fig. 2). Then, the next question that arises is about which model(s) to use. Choosing the right model is critical and the usual approach would be to construct several plausible models (e.g. in the example from Huys et al., models that consider an absence of reward as a punishment) and then select the model that best describes the data using model comparison [233]. Of course, model comparisons are limited to selecting the best model out of the models that are being considered and thus may rely on researchers' subjective choices. We can increase our confidence that we have indeed found a 'good enough' model by including model simulations, i.e. by letting the different models that have been fit to behaviour perform the same task as the participants. Then, we can check that the simulated behaviour captures key features of interest found in the behaviour of real participants [234] and that this is *only* the case for the model of interest, not for alternative models [235]. In general, the problems of complexity and model selection can be managed by increasing the sophistication of tasks gradually, by building on previous work, and by carefully defining the key cognition of interest in the first place.

Another source of concern might be the interpretation of findings obtained using more ecological tasks. As we have tried to illustrate, umbrella terms such as 'learning' often capture a range of different cognitive processes. Importantly, these processes can be described very precisely with separate parameters in a computational model, and thus computational modelling usually aids the interpretation of findings in more sophisticated tasks. For example finding that anxiety changes how well the learning rate is adjusted to match the environment should not be interpreted as a 'general learning deficit' (see section 'Environmental context and adaptive learning').

Given these caveats, it is clear that great care needs to be taken when designing and analysing experiments with ecological features and computational models. Fortunately, the field of computational neuroscience is moving towards open science, which we believe will accelerate the progress that can be made. It will enable direct access to resources, such as the code for a particular model and source data that can greatly aid replication and re-analysis.

4. Conclusions

In this review, we have argued that our understanding of learning and decision-making, in particular in the context of psychiatric disorders, could greatly benefit from considering tasks that probe more ecological features. Indeed, studying learning and decision-making in ways that capture aspects of naturalistic environments has started to reveal distinct cognitive processes that rely on different neural substrates, which would not be recruited in simple tasks. This diversity also highlights that generic umbrella terms such as 'learning' encompass a large variety of distinct and only partly overlapping neural processes. Thus, symptoms of psychiatric disorders may only emerge when probing the specific mechanisms that are also recruited in naturalistic scenarios relevant for the disorder under consideration. It is therefore essential to use tasks that incorporate ecological features and that are specifically designed to target the process of interest (Table 1). We have highlighted how computational modelling in combination with more ecological tasks can allow the dissociation of different behavioural processes and the characterization of different neural systems. We hope

this approach will enable better characterization of the diversity of an individual's behaviour (in other words creating computational 'fingerprints' of a person's cognitive abilities) and by extension enable mapping of subgroups of patients and symptoms more reliably. While psychiatric research has begun to apply some of those computational ideas, many fruitful avenues for future research remain. Ultimately, the results of these studies could help to build new unifying theories of psychiatric disorders that can be translated into patient treatment.

Funding

Jacqueline Scholl is funded by an MRC Skills Development fellowship (MR/N014448/1) and Miriam Klein-Flügge is funded by a Sir Henry Wellcome fellowship (103184/Z/13/Z).

Acknowledgments

The authors would like to thank Michael Browning, Laurence Hunt, Nils Kolling and Matthew Rushworth for helpful comments on an earlier version of this article.

References

- [1] C.M. Gillan, N.D. Daw, Taking psychiatry research online, *Neuron* 91 (2016) 19–23, <http://dx.doi.org/10.1016/j.neuron.2016.06.002>.
- [2] A.M. Chekroud, R. Gueorguieva, H.M. Krumholz, M.H. Trivedi, J.H. Krystal, G. McCarthy, Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach, *JAMA Psychiatry* (2017), <http://dx.doi.org/10.1001/jamapsychiatry.2017.0025>.
- [3] E. Fried, Moving forward: how depression heterogeneity hinders progress in treatment and research, *Expert Rev. Neurother.* 17 (2017) 423–425, <http://dx.doi.org/10.1080/14737175.2017.1307737>.
- [4] D. Borsboom, A.O.J. Cramer, Network analysis an integrative approach to the structure of psychopathology, *Annu. Rev. Clin. Psychol.* 9 (2013) 91–121, <http://dx.doi.org/10.1146/annurev-clinpsy-050212-185608>.
- [5] X.-J. Wang, J.H. Krystal, Computational psychiatry, *Neuron* 84 (2014) 638–654, <http://dx.doi.org/10.1016/j.neuron.2014.10.018>.
- [6] J.M. Pearson, K.K. Watson, M.L. Platt, Decision making the neuroethological turn, *Neuron* 82 (2014) 950–965, <http://dx.doi.org/10.1016/j.neuron.2014.04.037>.
- [7] A.J. Calhoun, B.Y. Hayden, The foraging brain, *Neuroeconomics* 5 (2015) 24–31, <http://dx.doi.org/10.1016/j.cobeha.2015.07.003>.
- [8] N. Kolling, T. Behrens, M. Wittmann, M. Rushworth, Multiple signals in anterior cingulate cortex, *Neurobiol. Cogn. Behav.* 37 (2016) 36–43, <http://dx.doi.org/10.1016/j.conb.2015.12.007>.
- [9] M.A. Addicott, J.M. Pearson, M.M. Sweitzer, D.L. Barack, M.L. Platt, A primer on foraging and the explore/exploit trade-off for psychiatry research, *Neuropsychopharmacology* (2017), <http://dx.doi.org/10.1038/npp.2017.108>.
- [10] D.R. Bach, P. Dayan, Algorithms for survival: a comparative perspective on emotions, *Nat. Rev. Neurosci.* 18 (2017) 311–319.
- [11] P.R. Corlett, P.C. Fletcher, Computational psychiatry: a Rosetta Stone linking the brain to mental illness, *Lancet Psychiatry* 1 (2014) 399–402.
- [12] R.A. Adams, Q.J.M. Huys, J.P. Roiser, Computational Psychiatry: towards a mathematically informed understanding of mental illness, *J. Neurol. Neurosurg. Am. Psychiatry* 87 (2015) 53, <http://dx.doi.org/10.1136/jnnp-2015-310737>.
- [13] P. Dayan, R.J. Dolan, K.J. Friston, P.R. Montague, Taming the shrewdness of neural function: methodological challenges in computational psychiatry, *Neuroeconomics* 5 (2015) 128–132, <http://dx.doi.org/10.1016/j.cobeha.2015.09.009>.
- [14] Q.J.M. Huys, T.V. Maia, M.J. Frank, Computational psychiatry as a bridge from neuroscience to clinical applications, *Nat. Neurosci.* 19 (2016) 404–413.
- [15] M. Moutoussis, E. Eldar, R.J. Dolan, Building a new field of computational psychiatry, *Biol. Psychiatry* (2016), <http://dx.doi.org/10.1016/j.biopsych.2016.10.007>.
- [16] H.R. Brown, P. Zeidman, P. Smittenaar, R.A. Adams, F. McNab, R.B. Rutledge, R.J. Dolan, Crowdsourcing for cognitive science? The utility of smartphones, *PLoS One* 9 (2014) e100662, <http://dx.doi.org/10.1371/journal.pone.0100662>.
- [17] R.B. Rutledge, M. Moutoussis, P. Smittenaar, P. Zeidman, T. Taylor, L. Hryniewicz, J. Lam, N. Skandali, J.Z. Siegel, O.T. Ousdal, G. Prabhu, P. Dayan, P. Fonagy, R.J. Dolan, Association of neural and emotional impacts of reward prediction errors with major depression, *JAMA Psychiatry* (2017), <http://dx.doi.org/10.1001/jamapsychiatry.2017.1713>.
- [18] A.T. Beck, Cognitive therapy: nature and relation to behavior therapy, *Behav. Ther.* 1 (1970) 184–200.
- [19] J. Wolpe, Reciprocal inhibition as the main basis of psychotherapeutic effects, *AMA Arch. Neurol. Psychiatry* 72 (1954) 205–226.
- [20] A.A. Lazarus, New methods in psychotherapy: a case study, *South Afr. Med. J.* 32 (1958) 660–663.
- [21] A.A. Lazarus, Has behavior therapy outlived its usefulness? *Am. Psychol.* 32 (1977) 550.
- [22] W. Schultz, P. Dayan, P.R. Montague, A neural substrate of prediction and reward, *Science* 275 (1997) 1593–1599.
- [23] J.P. O'Doherty, P. Dayan, K. Friston, H. Critchley, R.J. Dolan, Temporal difference models and reward-related learning in the human brain, *Neuron* 38 (2003) 329–337.
- [24] H.M. Bayer, P.W. Glimcher, Midbrain dopamine neurons encode a quantitative reward prediction error signal, *Neuron* 47 (2005) 129–141, <http://dx.doi.org/10.1016/j.neuron.2005.05.020>.
- [25] R.B. Rutledge, M. Dean, A. Caplin, P.W. Glimcher, Testing the reward prediction error hypothesis with an axiomatic model, *J. Neurosci.* 30 (2010) 13525–13536, <http://dx.doi.org/10.1523/JNEUROSCI.1747-10.2010>.
- [26] J. Garrison, B. Erdeniz, J. Done, Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies, *Neurosci. Biobehav. Rev.* 37 (2013) 1297–1310.
- [27] P. Dayan, K.C. Berridge, Model-based and model-free pavlovian reward learning: reevaluation, revision, and revelation, *Cogn. Behav. Neurosci.* 14 (2014) 473–492, <http://dx.doi.org/10.3758/s13415-014-0277-8>.
- [28] A.G.E. Collins, M.J. Frank, How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neuro-genetic analysis: working memory in reinforcement learning, *Eur. J. Neurosci.* 35 (2012) 1024–1035, <http://dx.doi.org/10.1111/j.1460-9568.2011.07980>.
- [29] N.J. Blanco, A.R. Otto, W.T. Maddox, C.G. Beavers, B.C. Love, The influence of depression symptoms on exploratory decision-making, *Cognition* 129 (2013) 563–568, <http://dx.doi.org/10.1016/j.cognition.2013.08.018>.
- [30] H.W. Chase, M.J. Frank, A. Michael, E.T. Bullmore, B.J. Sahakian, T.W. Robbins, Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia, *Psychol. Med.* 40 (2010) 433, <http://dx.doi.org/10.1017/S0033291709990468>.
- [31] M. Cella, S. Dymond, A. Cooper, Impaired flexible decision-making in major depressive disorder, *J. Affect. Disord.* 124 (2010) 207–210, <http://dx.doi.org/10.1016/j.jad.2009.11.013>.
- [32] M.M. Herzallah, A.A. Moustafa, J.Y. Natsheh, S.M. Abdellatif, M.B. Taha, Y.I. Tayem, M.A. Sehwal, I. Amleh, G. Petrides, C.E. Myers, M.A. Gluck, Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants, *Front. Integr. Neurosci.* 7 (2013), <http://dx.doi.org/10.3389/fnint.2013.00067>.
- [33] Y. Kunisato, Y. Okamoto, K. Ueda, K. Onoda, G. Okada, S. Yoshimura, S. Suzuki, K. Samejima, S. Yamawaki, Effects of depression on reward-based decision making and variability of action in probabilistic learning, *J. Behav. Ther. Exp. Psychiatry* 43 (2012) 1088–1094, <http://dx.doi.org/10.1016/j.jbtep.2012.05.007>.
- [34] E.M. Mueller, P. Pechtel, A.L. Cohen, S.R. Douglas, D.A. Pizzagalli, Potentiated processing of negative feedback in depression is attenuated by anhedonia: research Article: anhedonia and Feedback Processing in MDD, *Depress. Anxiety* 32 (2015) 296–305, <http://dx.doi.org/10.1002/da.22338>.
- [35] A. Must, Z. Szabó, N. Bódi, A. Szász, Z. Janka, S. Kéri, Sensitivity to reward and punishment and the prefrontal cortex in major depression, *J. Affect. Disord.* 90 (2006) 209–215, <http://dx.doi.org/10.1016/j.jad.2005.12.005>.
- [36] P. Pechtel, S.J. Dutra, E.L. Goetz, D.A. Pizzagalli, Blunted reward responsiveness in remitted depression, *J. Psychiatr. Res.* 47 (2013) 1864–1869, <http://dx.doi.org/10.1016/j.jpsychires.2013.08.011>.
- [37] P.L. Remijnse, M.M.A. Nielen, A.J.L.M. van Balkom, G.-J. Hendriks, W.J. Hoogendijk, H.B.M. Uylings, D.J. Veltman, Differential frontal-striatal and paralimbic activity during reversal learning in major depressive disorder and obsessive-compulsive disorder, *Psychol. Med.* 39 (2009) 1503, <http://dx.doi.org/10.1017/S0033291708005072>.
- [38] J.V. Taylor Tavares, L. Clark, M.L. Furey, G.B. Williams, B.J. Sahakian, W.C. Drevets, Neural basis of abnormal response to negative feedback in unmedicated mood disorders, *Neuroimage* 42 (2008) 1118–1126, <http://dx.doi.org/10.1016/j.neuroimage.2008.05.049>.
- [39] P. Thoma, C. Norra, G. Juckel, B. Suchan, C. Bellebaum, Performance monitoring and empathy during active and observational learning in patients with major depression, *Biol. Psychol.* 109 (2015) 222–231, <http://dx.doi.org/10.1016/j.biopsycho.2015.06.002>.
- [40] E. Vrieze, D.A. Pizzagalli, K. Demyttenaere, T. Hompe, P. Sienaert, P. de Boer, M. Schmidt, S. Claes, Reduced reward learning predicts outcome in major depressive disorder, *Biol. Psychiatry* 73 (2013) 639–645, <http://dx.doi.org/10.1016/j.biopsych.2012.10.014>.
- [41] R.M. Msetfi, R.A. Murphy, D.E. Kornbrot, Dysphoric mood states are related to sensitivity to temporal changes in contingency, *Front. Psychol.* 3 (2012), <http://dx.doi.org/10.3389/fpsyg.2012.00368>.
- [42] P. Kumar, G. Waiter, T. Ahearn, M. Milders, I. Reid, J.D. Steele, Abnormal temporal difference reward-learning signals in major depression, *Brain* 131 (2008) 2084–2093, <http://dx.doi.org/10.1093/brain/awn136>.
- [43] V.B. Gradin, P. Kumar, G. Waiter, T. Ahearn, C. Stickle, M. Milders, I. Reid, J. Hall, J.D. Steele, Expected value and prediction error abnormalities in depression and schizophrenia, *Brain* 134 (2011) 1751–1764, <http://dx.doi.org/10.1093/brain/awr059>.
- [44] O.J. Robinson, R. Cools, C.O. Carlisi, B.J. Sahakian, W.C. Drevets, Ventral striatum response during reward and punishment reversal learning in unmedicated major depressive disorder, *Am. J. Psychiatry* 169 (2012) 152–159.
- [45] C. Chen, T. Takahashi, S. Nakagawa, T. Inoue, I. Kusumi, Reinforcement learning in depression: a review of computational research, *Neurosci. Biobehav. Rev.* 55 (2015) 247–267, <http://dx.doi.org/10.1016/j.neubiorev.2015.05.005>.
- [46] Q.J. Huys, D.A. Pizzagalli, R. Bogdan, P. Dayan, Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis, *Biol. Mood Anxiety Disord.* 3 (2013) 12.

- [47] C.M. Gillan, M. Kosinski, R. Whelan, E.A. Phelps, N.D. Daw, Characterizing a psychiatric symptom dimension related to deficits in goal-directed control, *Elife* 5 (2016) e11305.
- [48] N. Garrett, T. Sharot, P. Faulkner, C.W. Korn, J.P. Roiser, R.J. Dolan, Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression, *Front. Hum. Neurosci.* 8 (2014), <http://dx.doi.org/10.3389/fnhum.2014.00639>.
- [49] C.W. Korn, T. Sharot, H. Walter, H.R. Heekeren, R.J. Dolan, Depression is related to an absence of optimistically biased belief updating about future life events, *Psychol. Med.* 44 (2014) 579–592, <http://dx.doi.org/10.1017/S0033291713001074>.
- [50] J.X. O'Reilly, *Making Predictions in a Changing World—Inference, Uncertainty, and Learning*, (2013).
- [51] T.E.J. Behrens, M.W. Woolrich, M.E. Walton, M.F.S. Rushworth, Learning the value of information in an uncertain world, *Nat. Neurosci.* 10 (2007) 1214–1221, <http://dx.doi.org/10.1038/nn1954>.
- [52] C. Mathys, J. Daunizeau, K.J. Friston, K.E. Stephan, A Bayesian foundation for individual learning under uncertainty, *Front. Hum. Neurosci.* 5 (2011) 39.
- [53] M. Browning, T.E. Behrens, G. Jocham, J.X. O'Reilly, S.J. Bishop, Anxious individuals have difficulty learning the causal statistics of aversive environments, *Nat. Neurosci.* 18 (2015) 590–596.
- [54] A.O. De Berker, R.B. Rutledge, C. Mathys, L. Marshall, G.F. Cross, R.J. Dolan, S. Bestmann, Computations of uncertainty mediate acute stress responses in humans, *Nat. Commun.* 7 (2016).
- [55] L. Marshall, C. Mathys, D. Ruge, A.O. de Berker, P. Dayan, K.E. Stephan, S. Bestmann, Pharmacological fingerprints of contextual uncertainty, *PLoS Biol.* 14 (2016) e1002575.
- [56] J. Scholl, N. Kolling, N. Nelissen, M. Browning, M.F. Rushworth, C.J. Harmer, Beyond negative valence: 2-week administration of a serotonergic antidepressant enhances both reward and effort learning signals, *PLoS Biol.* 15 (2017) e2000756.
- [57] C.J. Harmer, R.S. Duman, P.J. Cowen, How do antidepressants work? New perspectives for refining future treatment approaches, *Lancet Psychiatry* (2017) <http://www.sciencedirect.com/science/article/pii/S2215036617300159> (Accessed 21 June, 2017).
- [58] R.C. Wilson, Y. Niv, Inferring relevance in a changing world, *Front. Hum. Neurosci.* 5 (2012) 189.
- [59] Y. Niv, R. Daniel, A. Geana, S.J. Gershman, Y.C. Leong, A. Radulescu, R.C. Wilson, Reinforcement learning in multidimensional environments relies on attention mechanisms, *J. Neurosci.* 35 (2015) 8145–8157, <http://dx.doi.org/10.1523/JNEUROSCI.2978-14.2015>.
- [60] R. Akaiishi, N. Kolling, J.W. Brown, M. Rushworth, Neural mechanisms of credit assignment in a multicue environment, *J. Neurosci.* 36 (2016) 1096–1112, <http://dx.doi.org/10.1523/JNEUROSCI.3159-15.2016>.
- [61] Y.C. Leong, A. Radulescu, R. Daniel, V. DeWoskin, Y. Niv, Dynamic interaction between reinforcement learning and attention in multidimensional environments, *Neuron* 93 (2017) 451–463, <http://dx.doi.org/10.1016/j.neuron.2016.12.040>.
- [62] L.S. Hallion, A.M. Ruscio, A meta-analysis of the effect of cognitive bias modification on anxiety and depression, *Psychol. Bull.* 137 (2011) 940–958, <http://dx.doi.org/10.1037/a0024355>.
- [63] A. Stolcyn, J.D. Steele, J.W. Serietà, Conditioned task-set competition: neural mechanisms of emotional interference in depression, *Cogn. Affect. Behav. Neurosci.* (2016), <http://dx.doi.org/10.3758/s13415-016-0478-4>.
- [64] G. Rowe, J.B. Hirsh, A.K. Anderson, Positive affect increases the breadth of attentional selection, *Proc. Natl. Acad. Sci.* 104 (2007) 383–388.
- [65] P. Gable, E. Harmon-Jones, The motivational dimensional model of affect: implications for breadth of attention, memory, and cognitive categorisation, *Cogn. Emot.* 24 (2010) 322–337, <http://dx.doi.org/10.1080/02699930903378305>.
- [66] M. Grol, E.H.W. Koster, L. Bruyneel, R. De Raedt, Effects of positive mood on attention broadening for self-related information, *Psychol. Res.* (2013), <http://dx.doi.org/10.1007/s00426-013-0508-6>.
- [67] E. Eldar, J.D. Cohen, Y. Niv, The effects of neural gain on attention and learning, *Nat. Neurosci.* 16 (2013) 1146–1153, <http://dx.doi.org/10.1038/nn.3428>.
- [68] C. Padoa-Schioppa, J.A. Assad, Neurons in the orbitofrontal cortex encode economic value, *Nature* 441 (2006) 223–226, <http://dx.doi.org/10.1038/nature04676>.
- [69] M.C. Klein-Flügge, H.C. Barron, K.H. Brodersen, R.J. Dolan, T.E.J. Behrens, Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex, *J. Neurosci.* 33 (2013) 3202–3211, <http://dx.doi.org/10.1523/JNEUROSCI.2532-12.2013>.
- [70] A.-M. Schiffer, T. Muller, N. Yeung, F. Waszak, Reward activates stimulus-specific and task-dependent representations in visual association cortices, *J. Neurosci.* 34 (2014) 15610–15620, <http://dx.doi.org/10.1523/JNEUROSCI.1640-14.2014>.
- [71] E.D. Boorman, V.G. Rajendran, J.X. O'Reilly, T.E. Behrens, Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus, *Neuron* 89 (2016) 1343–1354, <http://dx.doi.org/10.1016/j.neuron.2016.02.014>.
- [72] M.P. Noonan, M.E. Walton, T.E.J. Behrens, J. Sallet, M.J. Buckley, M.F.S. Rushworth, Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex, *Proc. Natl. Acad. Sci.* 107 (2010) 20547–20552.
- [73] M.E. Walton, T.E.J. Behrens, M.J. Buckley, P.H. Rudebeck, M.F.S. Rushworth, Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning, *Neuron* 65 (2010) 927–939, <http://dx.doi.org/10.1016/j.neuron.2010.02.027>.
- [74] M.E. Walton, T.E.J. Behrens, M.P. Noonan, M.F.S. Rushworth, Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world: orbitofrontal cortex: value assignment and comparison, *Ann. N. Y. Acad. Sci.* 1239 (2011) 14–24, <http://dx.doi.org/10.1111/j.1749-6632.2011.06257.x>.
- [75] M.P. Noonan, B. Chau, M.F. Rushworth, L.K. Fellows, Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision making in humans, *J. Neurosci.* (2017), <http://dx.doi.org/10.1523/JNEUROSCI.0692-17.2017>.
- [76] M. Noonan, R. Mars, M. Rushworth, Distinct roles of three frontal cortical areas in reward-guided behavior, *J. Neurosci.* 31 (2011) 14399–14412.
- [77] E.D. Boorman, J.P. O'Doherty, R. Adolphs, A. Rangel, The behavioral and neural mechanisms underlying the tracking of expertise, *Neuron* 80 (2013) 1558–1571, <http://dx.doi.org/10.1016/j.neuron.2013.10.024>.
- [78] B.K.H. Chau, J. Sallet, G.K. Papageorgiou, M.P. Noonan, A.H. Bell, M.E. Walton, M.F.S. Rushworth, Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques, *Neuron* 87 (2015) 1106–1118, <http://dx.doi.org/10.1016/j.neuron.2015.08.018>.
- [79] G. Jocham, K.H. Brodersen, A.O. Constantinescu, M.C. Kahn, A.M. Ianni, M.E. Walton, M.F.S. Rushworth, T.E.J. Behrens, Reward-guided learning with and without causal attribution, *Neuron* 90 (2016) 177–190, <http://dx.doi.org/10.1016/j.neuron.2016.02.018>.
- [80] S.W. Lee, J.P. O'Doherty, S. Shimojo, Neural computations mediating one-shot learning in the human brain, *PLoS Biol.* 13 (2015) e1002137, <http://dx.doi.org/10.1371/journal.pbio.1002137>.
- [81] L.Y. Abramson, G.I. Metalsky, L.B. Alloy, Hopelessness depression: a theory-based subtype of depression, *Psychol. Rev.* 96 (1989) 358.
- [82] L.B. Alloy, L.Y. Abramson, B.E. Gibb, A.G. Crossfield, A.M. Pieracci, J. Spasojevic, J.A. Steinberg, Developmental antecedents of cognitive vulnerability to depression: review of findings from the cognitive vulnerability to depression project, *J. Cogn. Psychother.* 18 (2004) 115–133, <http://dx.doi.org/10.1891/jcop.18.2.115.65957>.
- [83] L.Y. Abramson, M.E. Seligman, J.D. Teasdale, Learned helplessness in humans: critique and reformulation, *J. Abnorm. Psychol.* 87 (1978) 49.
- [84] C. Peterson, A. Semmel, C. Von Baeyer, L.Y. Abramson, G.I. Metalsky, M.E. Seligman, The attributional style questionnaire, *Cogn. Ther. Res.* 6 (1982) 287–299.
- [85] Gerald J. Haefel, Brandon E. Gibb, Gerald I. Metalsky, Lauren B. Alloy, Lyn Y. Abramson, Benjamin L. Hankin, Thomas E. Joiner, Joel D. Swendsen, Measuring cognitive vulnerability to depression: Development and validation of the cognitive style questionnaire, *Clin. Psychol. Rev.* 28 (5) (2008) 824–836, <http://dx.doi.org/10.1016/j.cpr.2007.12.001> ISSN 0272-7358 <http://www.sciencedirect.com/science/article/pii/S0272735807002036>.
- [86] B.L. Hankin, L.Y. Abramson, N. Miller, G.J. Haefel, Cognitive vulnerability-stress theories of depression: examining affective specificity in the prediction of depression versus anxiety in three prospective studies, *Cogn. Ther. Res.* 28 (2004) 309–345.
- [87] G.J. Haefel, B.E. Gibb, G.I. Metalsky, L.B. Alloy, L.Y. Abramson, B.L. Hankin, T.E. Joiner, J.D. Swendsen, Measuring cognitive vulnerability to depression: development and validation of the cognitive style questionnaire, *Clin. Psychol. Rev.* 28 (2008) 824–836, <http://dx.doi.org/10.1016/j.cpr.2007.12.001>.
- [88] L.B. Alloy, L.Y. Abramson, Judgment of contingency in depressed and non-depressed students: sadder but wiser? *J. Exp. Psychol. Gen.* 108 (1979) 441.
- [89] R.M. Msetfi, R.A. Murphy, J. Simpson, Depressive realism and the effect of inter-trial interval on judgements of zero, positive, and negative contingencies, *Q. J. Exp. Psychol.* 60 (2007) 461–481, <http://dx.doi.org/10.1080/17470210601002595>.
- [90] R.M. Msetfi, C. Wade, R.A. Murphy, Context and time in causal learning: contingency and mood dependent effects, *PLoS One* 8 (2013) e64063, <http://dx.doi.org/10.1371/journal.pone.0064063>.
- [91] F. Blanco, H. Matute, M.A. Vadillo, Mediating role of activity level in the depressive realism effect, *PLoS One* 7 (2012) e46203.
- [92] N.C. Byrom, R.M. Msetfi, R.A. Murphy, Two pathways to causal control: use and availability of information in the environment in people with and without signs of depression, *Acta Psychol (Amst.)* 157 (2015) 1–12, <http://dx.doi.org/10.1016/j.actpsy.2015.02.004>.
- [93] L.G. Allan, S. Siegel, S. Hannah, The sad truth about depressive realism, *Q. J. Exp. Psychol.* 60 (2007) 482–495, <http://dx.doi.org/10.1080/17470210601002686>.
- [94] B.B. Doll, D.A. Simon, N.D. Daw, The ubiquity of model-based reinforcement learning, *Curr. Opin. Neurobiol.* 22 (2012) 1075–1081, <http://dx.doi.org/10.1016/j.conb.2012.08.003>.
- [95] Q.J.M. Huys, N.D. Daw, P. Dayan, Depression a decision-theoretic analysis, *Annu. Rev. Neurosci.* 38 (2015) 1–23, <http://dx.doi.org/10.1146/annurev-neuro-071714-033928>.
- [96] J.P. O'Doherty, S.W. Lee, D. McNamee, The structure of reinforcement-learning mechanisms in the human brain, *Curr. Opin. Behav. Sci.* 1 (2015) 94–100.
- [97] S.W. Lee, S. Shimojo, J.P. O'Doherty, Neural computations underlying arbitration between model-based and model-free learning, *Neuron* 81 (2014) 687–699, <http://dx.doi.org/10.1016/j.neuron.2013.11.028>.
- [98] J. Gläscher, N. Daw, P. Dayan, J.P. O'Doherty, States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning, *Neuron* 66 (2010) 585–595, <http://dx.doi.org/10.1016/j.neuron.2010.04.016>.
- [99] N.D. Daw, S.J. Gershman, B. Seymour, P. Dayan, R.J. Dolan, Model-based influences on humans' choices and striatal prediction errors, *Neuron* 69 (2011) 1204–1215, <http://dx.doi.org/10.1016/j.neuron.2011.02.027>.
- [100] T. Lohrenz, K. McCabe, C.F. Camerer, P.R. Montague, Neural signature of fictive learning signals in a sequential investment task, *Proc. Natl. Acad. Sci.* 104 (2007) 9493–9498.
- [101] E.D. Boorman, T.E.J. Behrens, M.W. Woolrich, M.F.S. Rushworth, How green is the

- grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action, *Neuron* 62 (2009) 733–743, <http://dx.doi.org/10.1016/j.neuron.2009.05.014>.
- [102] J. Scholl, N. Kolling, N. Nelissen, M.K. Wittmann, C.J. Harmer, M.F. Rushworth, The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort, *J. Neurosci.* 35 (2015) 11233–11251.
- [103] T.J. Vickery, M.M. Chun, D. Lee, Ubiquity and specificity of reinforcement signals throughout the human brain, *Neuron* 72 (2011) 166–177.
- [104] I. Volman, K. Roelofs, S. Koch, L. Verhagen, I. Toni, Anterior prefrontal cortex inhibition impairs control over social emotional actions, *Curr. Biol.* 21 (2011) 1766–1770, <http://dx.doi.org/10.1016/j.cub.2011.08.050>.
- [105] I. Volman, L. Verhagen, H.E. den Ouden, G. Fernández, M. Rijpkema, B. Franke, I. Toni, K. Roelofs, Reduced serotonin transporter availability decreases prefrontal control of the amygdala, *J. Neurosci.* 33 (2013) 8974–8979.
- [106] I. Volman, A. Katinka Louise von Borries, B. Hendrik Bulten, R. Jan Verkes, I. Toni, K. Roelofs, Testosterone modulates altered prefrontal control of emotional actions in psychopathic offenders, *eNeuro* 3 (2016), <http://dx.doi.org/10.1523/ENEURO.0107-15.2016>.
- [107] A.R. Vaidya, L.K. Fellows, Necessary contributions of human frontal lobe subregions to reward learning in a dynamic, multidimensional environment, *J. Neurosci.* 36 (2016) 9843–9858, <http://dx.doi.org/10.1523/JNEUROSCI.1337-16.2016>.
- [108] A.T. Beck, The evolution of the cognitive model of depression and its neurobiological correlates, *Am. J. Psychiatry* 165 (2008) 969–977.
- [109] D.A. Clark, A.T. Beck, Cognitive theory and therapy of anxiety and depression: convergence with neurobiological findings, *Trends Cogn. Sci.* 14 (2010) 418–424, <http://dx.doi.org/10.1016/j.tics.2010.06.007>.
- [110] S.G. Disner, C.G. Beevers, E.A.P. Haigh, A.T. Beck, Neural mechanisms of the cognitive model of depression, *Nat. Rev. Neurosci.* 12 (2011) 467–477, <http://dx.doi.org/10.1038/nrn3027>.
- [111] S.J. Gershman, Y. Niv, Learning latent structure: carving nature at its joints, *Curr. Opin. Neurobiol.* 20 (2010) 251–256.
- [112] R.C. Wilson, Y.K. Takahashi, G. Schoenbaum, Y. Niv, Orbitofrontal cortex as a cognitive map of task space, *Neuron* 81 (2014) 267–279.
- [113] S.C.Y. Chan, Y. Niv, K.A. Norman, A probability distribution over latent causes, in the orbitofrontal cortex, *J. Neurosci.* 36 (2016) 7817–7828, <http://dx.doi.org/10.1523/JNEUROSCI.0659-16.2016>.
- [114] A.O. Constantinescu, J.X. O'Reilly, T.E. Behrens, Organizing conceptual knowledge in humans with a gridlike code, *Science* 352 (2016) 1464–1468.
- [115] N.W. Schuck, M.B. Cai, R.C. Wilson, Y. Niv, Human orbitofrontal cortex represents a cognitive map of state space, *Neuron* 91 (2016) 1402–1412.
- [116] C. Amiez, Reward encoding in the monkey anterior cingulate cortex, *Cereb. Cortex* 16 (2005) 1040–1055, <http://dx.doi.org/10.1093/cercor/bhj046>.
- [117] S.W. Kennerley, M.E. Walton, T.E.J. Behrens, M.J. Buckley, M.F.S. Rushworth, Optimal decision making and the anterior cingulate cortex, *Nat. Neurosci.* 9 (2006) 940–947, <http://dx.doi.org/10.1038/nrn1724>.
- [118] M.P. Karlsson, D.G. Tervo, A.Y. Karpova, Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty, *Science* 338 (2012) 135–139.
- [119] N. Kolling, T.E. Behrens, R.B. Mars, M.F. Rushworth, Neural mechanisms of foraging, *Science* 336 (2012) 95–98.
- [120] N. Kolling, M. Wittmann, M.F.S. Rushworth, Multiple neural mechanisms of decision making and their competition under changing risk pressure, *Neuron* 81 (2014) 1190–1202, <http://dx.doi.org/10.1016/j.neuron.2014.01.033>.
- [121] J.X. O'Reilly, U. Schüffelgen, S.F. Cuell, T.E. Behrens, R.B. Mars, M.F. Rushworth, Dissociable effects of surprise and model update in parietal and anterior cingulate cortex, *Proc. Natl. Acad. Sci.* 110 (2013) E3660–E3669.
- [122] D.G.R. Tervo, M. Proskurin, M. Manakov, M. Kabra, A. Vollmer, K. Branson, A.Y. Karpova, Behavioral variability through stochastic choice and its gating by anterior cingulate cortex, *Cell* 159 (2014) 21–32, <http://dx.doi.org/10.1016/j.cell.2014.08.037>.
- [123] B.B. Doll, K.D. Duncan, D.A. Simon, D. Shohamy, N.D. Daw, Model-based choices involve prospective neural activity, *Nat. Neurosci.* 18 (2015) 767–772, <http://dx.doi.org/10.1038/nrn.3981>.
- [124] N.W. Schuck, R. Gaschler, D. Wenke, J. Heinze, P.A. Frensch, J.-D. Haynes, C. Reverberi, Medial prefrontal cortex predicts internally driven strategy shifts, *Neuron* 86 (2015) 331–340.
- [125] M.K. Wittmann, N. Kolling, R. Akaishi, B.K. Chau, J.W. Brown, N. Nelissen, M.F. Rushworth, Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex, *Nat. Commun.* 7 (2016).
- [126] J. Scholl, N. Kolling, N. Nelissen, C.J. Stagg, C.J. Harmer, M.F. Rushworth, Excitation and inhibition in anterior cingulate predict use of past experiences, *eLife* 6 (2017), <http://dx.doi.org/10.7554/eLife.20365>.
- [127] K.E. Stephan, Z.M. Manjaly, C.D. Mathys, L.A.E. Weber, S. Paliwal, T. Gard, M. Tittgemeyer, S.M. Fleming, H. Haker, A.K. Seth, F.H. Petzschner, Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression, *Front. Hum. Neurosci.* 10 (2016), <http://dx.doi.org/10.3389/fnhum.2016.00550>.
- [128] A.T. Drysdale, L. Grosenick, J. Downar, K. Dunlop, F. Mansouri, Y. Meng, R.N. Fretschel, B. Zebley, D.J. Oathes, A. Etkin, A.F. Schatzberg, K. Sudheimer, J. Keller, H.S. Mayberg, F.M. Gunning, G.S. Alexopoulos, M.D. Fox, A. Pascual-Leone, H.U. Voss, B.J. Casey, M.J. Dubin, C. Liston, Resting-state connectivity biomarkers define neurophysiological subtypes of depression, *Nat. Med.* 23 (2017) 28–38, <http://dx.doi.org/10.1038/nm.4246>.
- [129] M.F.S. Rushworth, M.P. Noonan, E.D. Boorman, M.E. Walton, T.E. Behrens, Frontal cortex and reward-guided learning and decision-making, *Neuron* 70 (2011) 1054–1069, <http://dx.doi.org/10.1016/j.neuron.2011.05.014>.
- [130] M.F. Rushworth, N. Kolling, J. Sallet, R.B. Mars, Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr. Opin. Neurobiol.* 22 (2012) 946–955, <http://dx.doi.org/10.1016/j.conb.2012.04.011>.
- [131] D. Öngür, J.L. Price, The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans, *Cereb. Cortex* 10 (2000) 206–219, <http://dx.doi.org/10.1093/cercor/10.3.206>.
- [132] G. Schoenbaum, A.A. Chiba, M. Gallagher, Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning, *Nat. Neurosci.* 1 (1998) 155–159, <http://dx.doi.org/10.1038/407>.
- [133] G. Schoenbaum, A.A. Chiba, M. Gallagher, Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning, *J. Neurosci.* 19 (1999) 1876–1884.
- [134] L. Tremblay, W. Schultz, Relative reward preference in primate orbitofrontal cortex, *Nature* 398 (1999) 704–708, <http://dx.doi.org/10.1038/19525>.
- [135] K. Hikosaka, M. Watanabe, Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards, *Cereb. Cortex N. Y. N 10* (2000) 263–271.
- [136] J.D. Wallis, E.K. Miller, Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task, *Eur. J. Neurosci.* 18 (2003) 2069–2081.
- [137] M.R. Roesch, C.R. Olson, Neuronal activity related to reward value and motivation in primate frontal cortex, *Science* 304 (2004) 307–310, <http://dx.doi.org/10.1126/science.1093223>.
- [138] P.H. Rudebeck, T.E. Behrens, S.W. Kennerley, M.G. Baxter, M.J. Buckley, M.E. Walton, M.F.S. Rushworth, Frontal cortex subregions play distinct roles in choices between actions and stimuli, *J. Neurosci.* 28 (2008) 13775–13785, <http://dx.doi.org/10.1523/jneurosci.3541-08.2008>.
- [139] S.W. Kennerley, A.F. Dahmubed, A.H. Lara, J.D. Wallis, Neurons in the frontal lobe encode the value of multiple decision variables, *J. Cogn. Neurosci.* 21 (2009) 1162–1178, <http://dx.doi.org/10.1162/jocn.2009.21100>.
- [140] M.Z. Wang, B.Y. Hayden, Reactivation of associative structure specific outcome responses during prospective evaluation in reward-based choices, *Nat. Commun.* 8 (2017), <http://dx.doi.org/10.1038/ncomms15821>.
- [141] B.Y. Hayden, M.L. Platt, Neurons in anterior cingulate cortex multiplex information about reward and action, *J. Neurosci.* 30 (2010) 3339–3346, <http://dx.doi.org/10.1523/jneurosci.4874-09.2010>.
- [142] B. Lau, P.W. Glimcher, Value representations in the primate striatum during matching behavior, *Neuron* 58 (2008) 451–463, <http://dx.doi.org/10.1016/j.neuron.2008.02.021>.
- [143] K. Matsumoto, W. Suzuki, K. Tanaka, Neuronal correlates of goal-based motor selection in the prefrontal cortex, *Science* 301 (2003) 229–232, <http://dx.doi.org/10.1126/science.1084204>.
- [144] K. Samejima, Y. Ueda, K. Doya, M. Kimura, Representation of action-specific reward values in the striatum, *Science* 310 (2005) 1337–1340, <http://dx.doi.org/10.1126/science.1115270>.
- [145] N. Camille, A. Tsuchida, L.K. Fellows, Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage, *J. Neurosci.* 31 (2011) 15048–15052, <http://dx.doi.org/10.1523/jneurosci.3164-11.2011>.
- [146] R.P. Dum, P.L. Strick, Spinal cord terminations of the medial wall motor areas in macaque monkeys, *J. Neurosci.* 16 (1996) 6513–6525.
- [147] R.J. Morecraft, G.W. Van Hoesen, Frontal granular cortex input to the cingulate (M3), supplementary (M2) and primary (M1) motor cortices in the rhesus monkey, *J. Comp. Neurol.* 337 (1993) 669–689, <http://dx.doi.org/10.1002/cne.903370411>.
- [148] L.T. Hunt, R.J. Dolan, T.E.J. Behrens, Hierarchical competitions subserving multi-attribute choice, *Nat. Neurosci.* 17 (2014) 1613–1622, <http://dx.doi.org/10.1038/nrn.3836>.
- [149] A. Meyer-Lindenberg, H. Tost, Neural mechanisms of social risk for psychiatric disorders, *Nat. Neurosci.* 15 (2012) 663–668, <http://dx.doi.org/10.1038/nn.3083>.
- [150] M.F. Rushworth, R.B. Mars, J. Sallet, Are there specialized circuits for social cognition and are they unique to humans? *Curr. Opin. Neurobiol.* 23 (2013) 436–442, <http://dx.doi.org/10.1016/j.conb.2012.11.013>.
- [151] C.C. Ruff, E. Fehr, The neurobiology of rewards and values in social decision making, *Nat. Rev. Neurosci.* 15 (2014) 549–562, <http://dx.doi.org/10.1038/nrn3776>.
- [152] N. Eshel, J.P. Roiser, Reward and punishment processing in depression, *Biol. Psychiatry* 68 (2010) 118–124, <http://dx.doi.org/10.1016/j.biopsych.2010.01.027>.
- [153] J.P. Roiser, B.J. Sahakian, Hot and cold cognition in depression, *CNS Spectr.* 18 (2013) 139–149, <http://dx.doi.org/10.1017/S1092852913000072>.
- [154] J.D. Amsterdam, R.G. Settle, R.L. Doty, E. Aelman, A. Winokur, Taste and smell perception in depression, *Biol. Psychiatry* 22 (1987) 1481–1485.
- [155] I. Berlin, L. Givry-Steiner, Y. Lecrubier, A.J. Puech, Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects, *Eur. Psychiatry J. Assoc. Eur. Psychiatr.* 13 (1998) 303–309, [http://dx.doi.org/10.1016/s0924-9338\(98\)80048-8](http://dx.doi.org/10.1016/s0924-9338(98)80048-8).
- [156] L. Sherdell, C.E. Waugh, I.H. Gotlib, Anticipatory pleasure predicts motivation for reward in major depression, *J. Abnorm. Psychol.* 121 (2012) 51–60, <http://dx.doi.org/10.1037/a0024945>.
- [157] D. Chung, K. Kadlec, J.A. Aimone, K. McCarry, B. King-Casas, P.H. Chiu, Valuation in major depression is intact and stable in a non-learning environment, *Sci. Rep.* 7 (2017), <http://dx.doi.org/10.1038/srep44374>.
- [158] J. Rottenberg, K.L. Kasch, J.J. Gross, I.H. Gotlib, Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder, *Emot. Wash. DC* 2 (2002) 135–146.

- [159] B.C. Beats, B.J. Sahakian, R. Levy, Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed, *Psychol. Med.* 26 (1996) 591–603.
- [160] R. Elliott, B.J. Sahakian, A.P. McKay, J.J. Herrod, T.W. Robbins, E.S. Paykel, Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance, *Psychol. Med.* 26 (1996) 975–989, <http://dx.doi.org/10.1017/S0033291700035303>.
- [161] J.D. Steele, P. Kumar, K.P. Ebmeier, Blunted response to feedback information in depressive illness, *Brain* 130 (2007) 2367–2374, <http://dx.doi.org/10.1093/brain/awm150>.
- [162] L.M. Bylsma, B.H. Morris, J. Rottenberg, A meta-analysis of emotional reactivity in major depressive disorder, *Clin. Psychol. Rev.* 28 (2008) 676–691, <http://dx.doi.org/10.1016/j.cpr.2007.10.001>.
- [163] R.J. Dolan, P. Dayan, Goals and habits in the brain, *Neuron* 80 (2013) 312–325, <http://dx.doi.org/10.1016/j.neuron.2013.09.007>.
- [164] M. Guitart-Masip, E. Duzel, R. Dolan, P. Dayan, Action versus valence in decision making, *Trends Cogn. Sci.* 18 (2014) 194–202, <http://dx.doi.org/10.1016/j.tics.2014.01.003>.
- [165] M. Guitart-Masip, Q.J.M. Huys, L. Fuentemilla, P. Dayan, E. Duzel, R.J. Dolan, Go and no-go learning in reward and punishment: interactions between affect and effect, *NeuroImage* 62 (2012) 154–166, <http://dx.doi.org/10.1016/j.neuroimage.2012.04.024>.
- [166] K. Amemori, A.M. Graybiel, Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making, *Nat. Neurosci.* 15 (2012) 776–785, <http://dx.doi.org/10.1038/nn.3088>.
- [167] J.P. Aggleton, N.F. Wright, D.L. Rosene, R.C. Saunders, Complementary patterns of direct amygdala and hippocampal projections to the macaque prefrontal cortex, *Cereb. Cortex N. Y. N* 25 (2015) 1991 4351–4373, <http://dx.doi.org/10.1093/cercor/bhv019>.
- [168] J.L. Price, W.C. Drevets, Neurocircuitry of mood disorders, *Neuropsychopharmacology* 35 (2010) 192–216, <http://dx.doi.org/10.1038/npp.2009.104>.
- [169] Q.J.M. Huys, R. Cools, M. Gölzer, E. Friedel, A. Heinz, R.J. Dolan, P. Dayan, Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding, *PLoS Comput. Biol.* 7 (2011) e1002028, <http://dx.doi.org/10.1371/journal.pcbi.1002028>.
- [170] P.C. Holland, M. Gallagher, Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer, *Eur. J. Neurosci.* 17 (2003) 1680–1694.
- [171] L.H. Corbit, B.W. Balleine, Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer, *J. Neurosci.* 25 (2005) 962–970, <http://dx.doi.org/10.1523/jneurosci.4507-04.2005>.
- [172] D. Talmi, B. Seymour, P. Dayan, R.J. Dolan, Human pavlovian-instrumental transfer, *J. Neurosci.* 28 (2008) 360–368, <http://dx.doi.org/10.1523/JNEUROSCI.4028-07.2008>.
- [173] C. Prévost, M. Liljeholm, J.M. Tyszka, J.P. O'Doherty, Neural correlates of specific and general pavlovian-to-instrumental transfer within human amygdalar subregions: a high-resolution fMRI study, *J. Neurosci.* 32 (2012) 8383–8390, <http://dx.doi.org/10.1523/JNEUROSCI.6237-11.2012>.
- [174] M. Lebreton, S. Jorge, V. Michel, B. Thirion, M. Pessiglione, An automatic valuation system in the human brain: evidence from functional neuroimaging, *Neuron* 64 (2009) 431–439, <http://dx.doi.org/10.1016/j.neuron.2009.09.040>.
- [175] Q.J.M. Huys, M. Gölzer, E. Friedel, A. Heinz, R. Cools, P. Dayan, R.J. Dolan, The specificity of Pavlovian regulation is associated with recovery from depression, *Psychol. Med.* 46 (2016) 1027–1035, <http://dx.doi.org/10.1017/S0033291715002597>.
- [176] D.R. Bach, M. Guitart-Masip, P.A. Packard, J. Miró, M. Falip, L. Fuentemilla, R.J. Dolan, Human hippocampus arbitrates approach-avoidance conflict, *Curr. Biol.* CB 24 (2014) 541–547, <http://dx.doi.org/10.1016/j.cub.2014.01.046>.
- [177] D.E.M. Geurts, Q.J.M. Huys, H.E.M. den Ouden, R. Cools, Serotonin and aversive Pavlovian control of instrumental behavior in humans, *J. Neurosci.* 33 (2013) 18932–18939, <http://dx.doi.org/10.1523/jneurosci.2749-13.2013>.
- [178] W.C. Drevets, J.L. Price, J.R. Simpson Jr, R.D. Todd, T. Reich, M. Vannier, M.E. Raichle, Subgenual prefrontal cortex abnormalities in mood disorders, *Nature* 386 (1997) 824–827, <http://dx.doi.org/10.1038/386824a0>.
- [179] K.N. Botteron, M.E. Raichle, W.C. Drevets, A.C. Heath, R.D. Todd, Volumetric reduction in left subgenual prefrontal cortex in early onset depression, *Biol. Psychiatry* 51 (2002) 342–344.
- [180] W.C. Drevets, Neuroimaging abnormalities in the amygdala in mood disorders, *Ann. N. Y. Acad. Sci.* 985 (2003) 420–444.
- [181] R. Admon, R.H. Kaiser, D.G. Dillon, M. Beltzer, F. Goer, D.P. Olson, G. Vitaliano, D.A. Pizzagalli, Dopaminergic enhancement of striatal response to reward in major depression, *Am. J. Psychiatry.* (2016), <http://dx.doi.org/10.1176/appi.ajp.2016.16010111>
- [182] E. Eldar, Y. Niv, Interaction between emotional state and learning underlies mood instability, *Nat. Commun.* 6 (2015) 6149, <http://dx.doi.org/10.1038/ncomms7149>.
- [183] E. Eldar, R.B. Rutledge, R.J. Dolan, Y. Niv, Mood as representation of momentum, *Trends Cogn. Sci.* 20 (2016) 15–24.
- [184] I.T. Kurniawan, B. Seymour, D. Talmi, W. Yoshida, N. Chater, R.J. Dolan, Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action, *J. Neurophysiol.* 104 (2010) 313–321, <http://dx.doi.org/10.1152/jn.00027.2010>.
- [185] M.C. Klein-Flügge, S.W. Kennerley, K. Friston, S. Bestmann, Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off, *J. Neurosci.* 36 (2016) 10002–10015, <http://dx.doi.org/10.1523/JNEUROSCI.0292-16.2016>.
- [186] P.L. Lockwood, M. Hamonet, S.H. Zhang, A. Ratnavel, F.U. Salmony, M. Husain, M.A.J. Apps, Prosocial apathy for helping others when effort is required, *Nat. Hum. Behav.* 1 (2017), <http://dx.doi.org/10.1038/s41562-017-0131-s41562-17-131-17>.
- [187] L.T. Hunt, N. Kolling, A. Soltani, M.W. Woolrich, M.F.S. Rushworth, T.E.J. Behrens, Mechanisms underlying cortical activity during value-guided choice, *Nat. Neurosci.* 15 (2012) 470–476, <http://dx.doi.org/10.1038/nn.3017>.
- [188] J.W. Kable, P.W. Glimcher, The neural correlates of subjective value during intertemporal choice, *Nat. Neurosci.* 10 (2007) 1625–1633, <http://dx.doi.org/10.1038/nn2007>.
- [189] A. Nicole, M.C. Klein-Flügge, L.T. Hunt, I. Vlaev, R.J. Dolan, T.E.J. Behrens, An agent independent axis for executed and modeled choice in medial prefrontal cortex, *Neuron* 75 (2012) 1114–1121, <http://dx.doi.org/10.1016/j.neuron.2012.07.023>.
- [190] C. Prévost, M. Pessiglione, E. Méteureau, M.-L. Cléry-Melin, J.-C. Dreher, Separate valuation subsystems for delay and effort decision costs, *J. Neurosci.* 30 (2010) 14080–14090, <http://dx.doi.org/10.1523/jneurosci.2752-10.2010>.
- [191] M.E. Walton, D.M. Bannerman, K. Alterescu, M.F.S. Rushworth, Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions, *J. Neurosci.* 23 (2003) 6475–6479.
- [192] M.E. Walton, S.W. Kennerley, D.M. Bannerman, P.E.M. Phillips, M.F.S. Rushworth, Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making, *Neural Netw.* 19 (2006) 1302–1314, <http://dx.doi.org/10.1016/j.neunet.2006.03.005>.
- [193] P.H. Rudebeck, M.E. Walton, A.N. Smyth, D.M. Bannerman, M.F.S. Rushworth, Separate neural pathways process different decision costs, *Nat. Neurosci.* 9 (2006) 1161–1168, <http://dx.doi.org/10.1038/nn1756>.
- [194] P.L. Croxson, M.E. Walton, J.X. O'Reilly, T.E.J. Behrens, M.F.S. Rushworth, Effort-based cost-benefit valuation and the human brain, *J. Neurosci.* 29 (2009) 4531–4541, <http://dx.doi.org/10.1523/jneurosci.4515-08.2009>.
- [195] K.L. Hillman, D.K. Bilkey, Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task, *J. Neurosci.* 30 (2010) 7705–7713, <http://dx.doi.org/10.1523/jneurosci.1273-10.2010>.
- [196] T. Hosokawa, S.W. Kennerley, J. Sloan, J.D. Wallis, Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex, *J. Neurosci.* 33 (2013) 17385–17397, <http://dx.doi.org/10.1523/jneurosci.2221-13.2013>.
- [197] V. Bonnelle, S. Manohar, T. Behrens, M. Husain, Individual differences in pre-motor brain systems underlie behavioral apathy, *Cereb. Cortex N. Y. NY.* 26 (2016) 807–819, <http://dx.doi.org/10.1093/cercor/bhv247>.
- [198] M.T. Treadway, N.A. Bossaller, R.C. Shelton, D.H. Zald, Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia, *J. Abnorm. Psychol.* 121 (2012) 553–558, <http://dx.doi.org/10.1037/a0028813>.
- [199] M.-L. Cléry-Melin, L. Schmidt, G. Lafargue, N. Baup, P. Fossati, M. Pessiglione, Why don't you try harder? An investigation of effort production in major depression, *PLoS One* 6 (2011) e23178, <http://dx.doi.org/10.1371/journal.pone.0023178>.
- [200] M.N. Hartmann, O.M. Hager, A.V. Reimann, J.R. Chumbley, M. Kirschner, E. Seifritz, P.N. Tobler, S. Kaiser, Apathy but not diminished expression in schizophrenia is associated with discounting of monetary rewards by physical effort, *Schizophr. Bull.* (2014), <http://dx.doi.org/10.1093/schbul/sbu102>.
- [201] V. Bonnelle, K.-R. Veromann, S. Burnett Heyes, E. Lo Sterzo, S. Manohar, M. Husain, Characterization of reward and effort mechanisms in apathy, *J. Physiol. Paris* 109 (2015) 16–26, <http://dx.doi.org/10.1016/j.jphysparis.2014.04.002>.
- [202] E.I. Fried, S. Epskamp, R.M. Nesse, F. Tuerlinckx, D. Borsboom, What are good depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis, *J. Affect. Disord.* 189 (2016) 314–320, <http://dx.doi.org/10.1016/j.jad.2015.09.005>.
- [203] D.-J. Saaltink, E. Vreugdenhil, Stress, glucocorticoid receptors, and adult neurogenesis: a balance between excitation and inhibition? *Cell Mol. Life Sci. CMLS* 71 (2014) 2499–2515, <http://dx.doi.org/10.1007/s0018-014-1568-5>.
- [204] J.F. Maya Vetencourt, A. Sale, A. Viegi, L. Baroncelli, R. De Pasquale, O.F. O'Leary, E. Castrén, L. Maffei, The antidepressant fluoxetine restores plasticity in the adult visual cortex, *Science* 320 (2008) 385–388, <http://dx.doi.org/10.1126/science.1150516>.
- [205] G. Hasler, J.W. van der Veen, T. Tuminis, N. Meyers, J. Shen, W.C. Drevets, Reduced prefrontal glutamate/glutamine and gamma-aminobutyric acid levels in major depression determined using proton magnetic resonance spectroscopy, *Arch. Gen. Psychiatry* 64 (2007) 193–200, <http://dx.doi.org/10.1001/archpsyc.64.2.193>.
- [206] V. Gabbay, X. Mao, R.G. Klein, B.A. Ely, J.S. Babb, A.M. Panzer, C.M. Alonso, D.C. Shungu, Anterior cingulate cortex γ -aminobutyric acid in depressed adolescents: relationship to anhedonia, *Arch. Gen. Psychiatry* 69 (2012) 139–149, <http://dx.doi.org/10.1001/archgenpsychiatry.2011.131>.
- [207] C. Yüksel, D. Öngür, Magnetic resonance spectroscopy studies of glutamate-related abnormalities in mood disorders, *Biol. Psychiatry* 68 (2010) 785–794, <http://dx.doi.org/10.1016/j.biopsych.2010.06.016>.
- [208] J.J. Luyckx, K.G. Laban, M.P. van den Heuvel, M.P.M. Boks, R.C.W. Mandl, R.S. Kahn, S.C. Bakker, Region and state specific glutamate downregulation in major depressive disorder: a meta-analysis of (1)H-MRS findings, *Neurosci. Biobehav. Rev.* 36 (2012) 198–205, <http://dx.doi.org/10.1016/j.neubiorev.2011.05.014>.
- [209] J. Scholl, J. Günthner, N. Kolling, E. Favaron, M.F. Rushworth, C.J. Harmer,

- A. Reinecke, A role beyond learning for NMDA receptors in reward-based decision-making—a pharmacological study using d-cycloserine, *Neuropsychopharmacology* 39 (2014) 2900–2909, <http://dx.doi.org/10.1038/npp.2014.144>.
- [210] K.E. Binns, T.E. Salt, Importance of NMDA receptors for multimodal integration in the deep layers of the cat superior colliculus, *J. Neurophysiol.* 75 (1996) 920–930.
- [211] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, fourth edition, DSM-IV-TR[®]*, Washington, DC, 2000.
- [212] M.H. Radford, L. Mann, R.S. Kalucy, Psychiatric disturbance and decision-making, *Aust. N. Z. J. Psychiatry* 20 (1986) 210–217, <http://dx.doi.org/10.3109/00048678609161333>.
- [213] M.H. Radford, Y. Nakane, Y. Ohta, L. Mann, R.S. Kalucy, Decision making in clinically depressed patients. A transcultural social psychological study, *J. Nerv. Ment. Dis.* 179 (1991) 711–719.
- [214] D.E. Saunders, G.W. Peterson, J.P. Sampson, R.C. Reardon, Relation of depression and dysfunctional career thinking to career indecision, *J. Vocat. Behav.* 2 (2000) 288–298, <http://dx.doi.org/10.1006/jvbe.1999.1715>.
- [215] J.O. Okwumabua, S.P. Wong, E.J. Duryea, Depressive symptoms and decision making among African American youth, *J. Adolesc. Res.* 18 (2003) 436–453, <http://dx.doi.org/10.1177/0743558403255062>.
- [216] Y. Leykin, C.S. Roberts, R.J. Derubeis, Decision-making and depressive symptomatology, *Cogn. Ther. Res.* 35 (2011) 333–341, <http://dx.doi.org/10.1007/s10608-010-9308-0>.
- [217] G. Jochem, L.T. Hunt, J. Near, T.E.J. Behrens, A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex, *Nat. Neurosci.* 15 (2012) 960–961, <http://dx.doi.org/10.1038/nn.3140>.
- [218] D. Hämmerer, J. Bonaiuto, M. Klein-Flügge, M. Bikson, S. Bestmann, Selective alteration of human value decisions with medial frontal tDCS is predicted by changes in attractor dynamics, *Sci. Rep.* 6 (2016), <http://dx.doi.org/10.1038/srep25160> Accessed 31 March, 2017.
- [219] X.-J. Wang, Probabilistic decision making by slow reverberation in cortical circuits, *Neuron* 36 (2002) 955–968.
- [220] X.-J. Wang, Decision making in recurrent neuronal circuits, *Neuron* 60 (2008) 215–234, <http://dx.doi.org/10.1016/j.neuron.2008.09.034>.
- [221] F.C. Murphy, J.S. Rubinsztein, A. Michael, R.D. Rogers, T.W. Robbins, E.S. Paykel, B.J. Sahakian, Decision-making cognition in mania and depression, *Psychol. Med.* 31 (2001) 679–693.
- [222] C.S. Green, C. Benson, D. Kersten, P. Schrater, Alterations in choice behavior by manipulations of world model, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 16401–16406, <http://dx.doi.org/10.1073/pnas.1001709107>.
- [223] N.D. Daw, J.P. O'Doherty, P. Dayan, B. Seymour, R.J. Dolan, Cortical substrates for exploratory decisions in humans, *Nature* 441 (2006) 876–879, <http://dx.doi.org/10.1038/nature04766>.
- [224] R.C. Wilson, A. Geana, J.M. White, E.A. Ludvig, J.D. Cohen, Humans use directed and random exploration to solve the explore-exploit dilemma, *J. Exp. Psychol. Gen.* 143 (2014) 2074–2081, <http://dx.doi.org/10.1037/a0038199>.
- [225] W.K. Zajkowski, M. Kossut, R.C. Wilson, A causal role for right frontopolar cortex in directed, but not random, exploration, *bioRxiv* (2017) 127704, <http://dx.doi.org/10.1101/127704>.
- [226] E.D. Boorman, M.F. Rushworth, T.E. Behrens, Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice, *J. Neurosci.* 33 (2013) 2242–2253, <http://dx.doi.org/10.1523/JNEUROSCI.3022-12.2013>.
- [227] R. Quilodran, M. Rothé, E. Procyk, Behavioral shifts and action valuation in the anterior cingulate cortex, *Neuron* 57 (2008) 314–325, <http://dx.doi.org/10.1016/j.neuron.2007.11.031>.
- [228] B.Y. Hayden, J.M. Pearson, M.L. Platt, Neuronal basis of sequential foraging decisions in a patchy environment, *Nat. Neurosci.* 14 (2011) 933–939, <http://dx.doi.org/10.1038/nn.2856>.
- [229] H. Seo, X. Cai, C.H. Donahue, D. Lee, Neural correlates of strategic reasoning during competitive games, *Science* 346 (2014) 340–343, <http://dx.doi.org/10.1126/science.1256254>.
- [230] N.J. Blanco, A.R. Otto, W.T. Maddox, C.G. Beevers, B.C. Love, The influence of depression symptoms on exploratory decision-making, *Cognition* 129 (2013) 563–568, <http://dx.doi.org/10.1016/j.cognition.2013.08.018>.
- [231] J.P. Gläscher, J.P. O'Doherty, Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data, *Wiley Interdiscip. Rev. Cogn. Sci.* (2010), <http://dx.doi.org/10.1002/wcs.57> n/a–n/a.
- [232] P. Schwartenbeck, K. Friston, Computational phenotyping in psychiatry: a worked example, *Eneuro* 3 (2016), <http://dx.doi.org/10.1523/ENEURO.0049-16.2016>.
- [233] N.D. Daw, Trial-by-trial data analysis using computational models, *Decis. Mak. Affect. Learn. Atten. Perform.* XXIII (23) (2011) 3–38.
- [234] M.R. Nassar, J.I. Gold, A healthy fear of the unknown: perspectives on the interpretation of parameter fits from computational models in neuroscience, *PLoS Comput. Biol.* 9 (2013) e1003015, <http://dx.doi.org/10.1371/journal.pcbi.1003015>.
- [235] S. Palminteri, V. Wyart, E. Koehlin, The importance of falsification in computational cognitive modeling, *Trends Cogn. Sci.* 21 (2017) 425–433, <http://dx.doi.org/10.1016/j.tics.2017.03.011>.
- [236] B. Vogt, *Cingulate Neurobiology and Disease*, OUP Oxford, 2009.
- [237] E. Procyk, C.R.E. Wilson, F.M. Stoll, M.C.M. Faraut, M. Petrides, C. Amiez, Midcingulate motor map and feedback detection: converging data from humans and monkeys, *Cereb. Cortex* N. Y. N 26 (2016) (1991) 467–476, <http://dx.doi.org/10.1093/cercor/bhu213>.
- [238] B.A. Vogt, E.A. Nimchinsky, L.J. Vogt, P.R. Hof, Human cingulate cortex: surface features, flat maps, and cytoarchitecture, *J. Comp. Neurol.* 359 (1995) 490–506, <http://dx.doi.org/10.1002/cne.903590310>.
- [239] J. Sallet, R.B. Mars, M.P. Noonan, F.-X. Neubert, S. Jbabdi, J.X. O'Reilly, N. Filippini, A.G. Thomas, M.F. Rushworth, The organization of dorsal frontal cortex in humans and macaques, *J. Neurosci.* 33 (2013) 12255–12274, <http://dx.doi.org/10.1523/jneurosci.5108-12.2013>.
- [240] F.-X. Neubert, R.B. Mars, A.G. Thomas, J. Sallet, M.F.S. Rushworth, Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex, *Neuron* 81 (2014) 700–713, <http://dx.doi.org/10.1016/j.neuron.2013.11.012>.
- [241] F.-X. Neubert, R.B. Mars, J. Sallet, M.F.S. Rushworth, Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) E2695–2704, <http://dx.doi.org/10.1073/pnas.1410767112>.
- [242] H.C. Barron, M.M. Garvert, T.E.J. Behrens, Repetition suppression: a means to index neural representations using BOLD? *Philos. Trans. R. Soc. B* 371 (2016) 20150355, <http://dx.doi.org/10.1098/rstb.2015.0355>.
- [243] A. Tversky, D. Kahneman, Advances in prospect theory: cumulative representation of uncertainty, *J. Risk Uncertain* 5 (1992) 297–323.
- [244] D. Kahneman, A. Tversky, Prospect theory – analysis of decision under risk, *Econometrica* 47 (1979) 263–291.
- [245] D. Prelec, The probability weighting function, *Econometrica* 66 (1998) 497–527, <http://dx.doi.org/10.2307/2998573>.
- [246] M.N. Hartmann, O.M. Hager, P.N. Tobler, S. Kaiser, Parabolic discounting of monetary rewards by physical effort, *Behav. Processes* 100 (2013) 192–196, <http://dx.doi.org/10.1016/j.beproc.2013.09.014>.
- [247] M.C. Klein-Flügge, S.W. Kennerley, A.C. Saraiva, W.D. Penny, S. Bestmann, Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation, *PLoS Comput. Biol.* 11 (2015) e1004116, <http://dx.doi.org/10.1371/journal.pcbi.1004116>.
- [248] J. Myerson, L. Green, Discounting of delayed rewards: models of individual choice, *J. Exp. Anal. Behav.* 64 (1995) 263–276.
- [249] G.J. Madden, W.K. Bickel, E.A. Jacobs, Discounting of delayed rewards in opioid-dependent outpatients: exponential or hyperbolic discounting functions? *Exp. Clin. Psychopharmacol.* 7 (1999) 284–293, <http://dx.doi.org/10.1037/1064-1297.7.3.284>.
- [250] C.J. Stagg, M.A. Nitsche, Physiological basis of transcranial direct current stimulation, *Neuroscientist* 17 (2011) 37–53, <http://dx.doi.org/10.1177/1073858410386614>.
- [251] A. Rahman, D. Reato, M. Arlotti, F. Gasca, A. Datta, L.C. Parra, M. Bikson, Cellular effects of acute direct current stimulation: somatic and synaptic terminal effects, *J. Physiol.* 591 (2013) 2563–2578, <http://dx.doi.org/10.1113/jphysiol.2012.247171>.
- [252] Tali Sharot, W. Korn Christoph, J. Dolan Raymond, How unrealistic optimism is maintained in the face of reality, *Nature neuroscience* 14 (11) (2011) 1475–1479.