

REVIEW ARTICLE

Interactions between ventrolateral prefrontal and anterior cingulate cortex during learning and behavioural change

Ilya E. Monosov^{1,2,3,4,5}✉ and Matthew F. S. Rushworth⁶✉

© The Author(s), under exclusive licence to American College of Neuropsychopharmacology 2021

Hypotheses and beliefs guide credit assignment – the process of determining which previous events or actions caused an outcome. Adaptive hypothesis formation and testing are crucial in uncertain and changing environments in which associations and meanings are volatile. Despite primates' abilities to form and test hypotheses, establishing what is causally responsible for the occurrence of particular outcomes remains a fundamental challenge for credit assignment and learning. Hypotheses about what surprises are due to stochasticity inherent in an environment as opposed to real, systematic changes are necessary for identifying the environment's predictive features, but are often hard to test. We review evidence that two highly interconnected frontal cortical regions, anterior cingulate cortex and ventrolateral prefrontal area 47/12o, provide a biological substrate for linking two crucial components of hypothesis-formation and testing: the control of information seeking and credit assignment. Neuroimaging, targeted disruptions, and neurophysiological studies link an anterior cingulate – 47/12o circuit to generation of exploratory behaviour, non-instrumental information seeking, and interpretation of subsequent feedback in the service of credit assignment. Our observations support the idea that information seeking and credit assignment are linked at the level of neural circuits and explain why this circuit is important for ensuring behaviour is flexible and adaptive.

Neuropsychopharmacology (2022) 47:196–210; <https://doi.org/10.1038/s41386-021-01079-2>

INTRODUCTION

Although it comes in many guises, an enduring idea about the prefrontal cortex is that it is essential for flexible behaviour [1, 2]. A behaviour that is successful in one situation may not lead to the desired outcome in another situation. Therefore, to ensure that behaviour has the best chance of being successful, it is necessary for it to change from one context to the next. Successful flexible behaviour involves repeated cycles of exploratory behaviour, in which the best course of action or the most useful source of guiding information is sought, interleaved with periods in which this knowledge is exploited. Each time the environment changes (or the internal motivation state of the agent changes), it is necessary to re-enter the exploratory phase and identify which sources of information or courses of action should be followed.

We argue that such exploratory behaviour is guided by hypotheses that animals form which guide their interpretation of the feedback they receive after making choices. These hypotheses allow them to assign credit for the feedback they receive to particular previous events or actions. Not only do primates form hypotheses about which antecedent events might be causally responsible for subsequent outcomes, but they also form hypotheses about what surprises are due to inherent stochasticity in the environment and which are due to systematic changes. Here we review how two regions of frontal cortex, anterior cingulate cortex (ACC) and a region centred on area 47/

12o in ventrolateral prefrontal cortex (Fig. 1), are central to this process. First, they allow primates to explore and identify important uncertainty-reducing and information-providing features in the environment. Second, these brain regions establish the predictive significance of these items so that they can be used to guide future behaviour. We argue that the ACC-47/12o network drives information seeking and that a region roughly centred on 47/12o also prominently participates in a second process of credit assignment – the linking of choices with outcomes as a result of information seeking. Once predictive significance is established then behaviour can re-enter a stable period in which the information gained can be exploited.

In the next sections, first we review the evidence suggesting that these brain regions play preeminent roles in flexible behaviour as assessed with a widely used task – reversal learning. We then go on to argue that their importance for flexible behaviour is the consequence of the roles that the areas have in information seeking and credit assignment.

LINKING 47/12O AND ACC TO BEHAVIOURAL FLEXIBILITY

Reversal tasks are the classic way to assess whether behaviour is adaptive and flexible. In reversal tasks an animal is taught to discriminate between two objects; it learns that choosing one object leads to reward while choosing the other does not. Once

¹Department of Neuroscience, Washington University School of Medicine, St. Louis, MO, USA. ²Department of Biomedical Engineering, Washington University, St. Louis, MO, USA.

³Department of Electrical Engineering, Washington University, St. Louis, MO, USA. ⁴Department of Neurosurgery, Washington University, St. Louis, MO, USA. ⁵Pain Center, Washington University, St. Louis, MO, USA. ⁶Wellcome Centre for Integrative Neuroimaging (WIN), Department of Experimental Psychology, University of Oxford, Oxford, UK.

✉email: ilya.monosov@gmail.com; matthew.rushworth@psy.ox.ac.uk

Received: 7 April 2021 Revised: 27 May 2021 Accepted: 15 June 2021

Published online: 7 July 2021

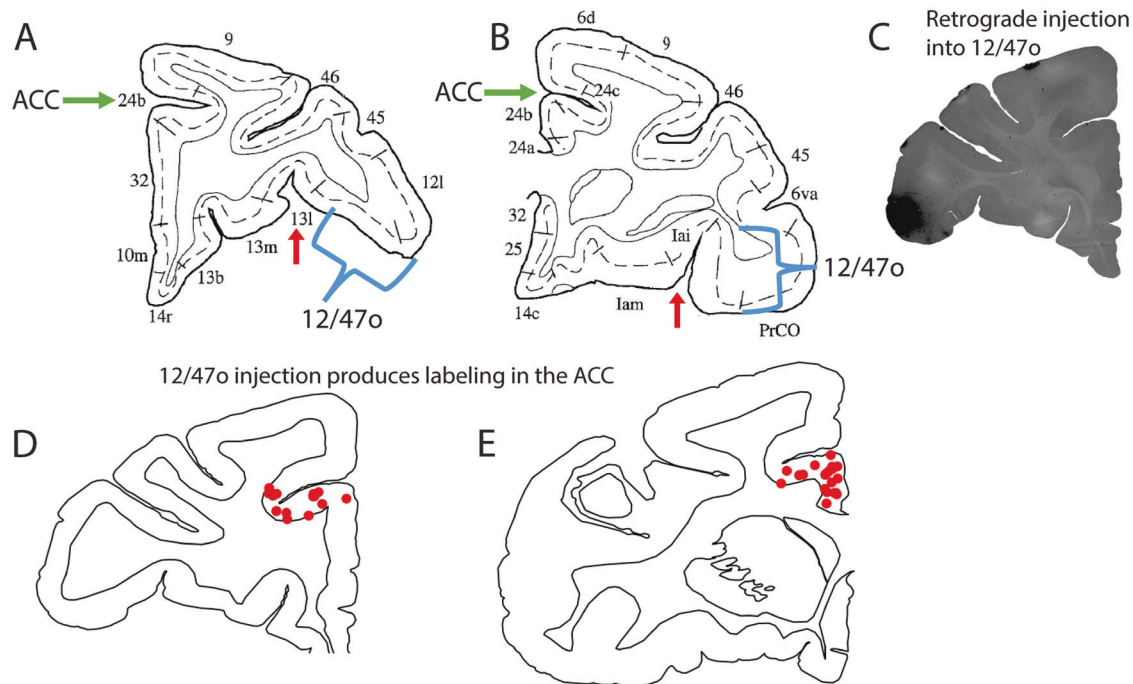


Fig. 1 Coronal cross sections through the macaque frontal cortex. The cytoarchitectonic areas identified by Carmichael and Price (1994) are illustrated on a more anterior (**A**) and posterior (**B**) coronal section. Area 47/12o lies on the boundary between the regions that are typically taken as representative of orbitofrontal cortex (orbitofrontal cortex is typically regarded as lying medial to the lateral orbitofrontal sulcus indicated with red arrows in panels **A** and **B**) and the lateral prefrontal cortex (lateral prefrontal cortex is typically regarded as lying on the lateral surface of the prefrontal cortex). The area that we refer to as 47/12o was identified by Carmichael and Price as area 12o, a subdivision of area 12, that lies in and lateral to the lateral orbitofrontal sulcus that they referred to as orbital area 12 or 12o. Because panels **A** and **B** are taken from Carmichael and Price it is shown as 12/47o here. It is lateral to the lateral orbital sulcus in panel **A** and in the lateral orbital sulcus in panel **B**. Mackey and Petrides [12] have pointed out the similarities between macaque area 12 and human area 47 and so we have followed their suggestion and referred to this region as 47/12o. In addition to 47/12o, we also focus on a dorsal part of ACC in and just ventral to the cingulate sulcus (marked by 24b and 24c labels and indicated by green arrows). Although there is a lack of connections between many parts of ACC and many parts of orbital and ventrolateral prefrontal cortex, ACC and 47/12o are interconnected. When Lucifer Yellow tracer was injected into 47/12o (**C**) retrograde labelling was seen in anterior (**D**) and more posterior (**E**) parts of ACC [23].

the discrimination has been mastered, however, the rules change. The object that had been rewarded is no longer rewarded and now it is the other object that is rewarded. A number of studies have reported that orbitofrontal lesions impair object discrimination reversal learning and traditionally emphasis has been placed on the idea that it is the need for inhibiting the previously rewarded behaviour that makes the reversal tasks difficult [1, 3–7].

Our understanding of the cognitive processes that are needed to perform reversal tasks and the critical brain regions on which they depend has, however, begun to evolve. In one important study, Rudebeck et al. [8] carefully made excitotoxic lesions of orbitofrontal cortex that spared adjacent white matter. They showed that when lesions were made in this way then object discrimination reversal learning was not impaired even though small aspiration lesions, which compromise white matter connection pathways, do impair object discrimination reversal.

The lesions made by Rudebeck et al. encompassed the orbitofrontal region between the lateral orbital sulcus and the rostral sulcus (Fig. 1) but they left the tissue lateral to the lateral orbitofrontal sulcus intact. While most of the tissue between these sulci at the level of the coronal section shown in Fig. 1A was removed, the lesions did not extend as far posteriorly as the coronal section shown in Fig. 1B. The region that lies lateral to the one in which Rudebeck et al. made lesions has sometimes been called lateral OFC (lOFC) and sometimes included within the area referred to as ventrolateral prefrontal cortex (vlPFC) [9–11]. Rather than trying to decide whether it is better to call it vlPFC or lOFC, it may be better simply to refer to this region as 47/12o [12]. The inclusion of “12” in the name reflects the fact that the area is

cytoarchitecturally similar to that in adjacent parts of vlPFC in the inferior convexity [13] which, in the macaque, had been given the designation 12 [14]. The inclusion of “47” in the name acknowledges, however, that the cytoarchitecture of this region of the macaque brain also resembles part of a region that had been labelled 47 earlier in a different primate species – the human – by Brodmann [15]; it has been argued that human area 47 and macaque area 12 are homologous [16, 17]. The “o” in 47/12o reflects the fact that it is the part of area 47/12 that is closest to OFC and the orbit of the eye. We link the activity in this area to reward-guided learning and credit assignment. It is clear that there is important activity present within 47/12o but whether it is confined to 47/12o or whether it extends into immediately adjacent areas such as 13l, 12l, and 12m is not certain. 12/47o lies at the centre of this set of areas and they appear to constitute a network; we know that they are anatomically interconnected [18].

Area 47/12o is not just interesting because it has often been overlooked by investigators focused on more lateral vlPFC or more medial orbitofrontal cortex, it also interesting because it has a very special pattern of anatomical connections that suggests it might play an equally special role in cognition. Unlike other medial and orbital regions, it receives highly processed polymodal and visual information from temporal lobe areas including the superior temporal sulcus, TE, and perirhinal cortex [19–22]. Moreover, again unlike many adjacent areas, 47/12o is interconnected with much of both the orbital and medial frontal cortex [18] and particularly strongly with the ACC [23; Fig. 1c–e]. Also, within the basal ganglia, the same regions of the striatum that receive dense ACC inputs also receive dense 47/12o

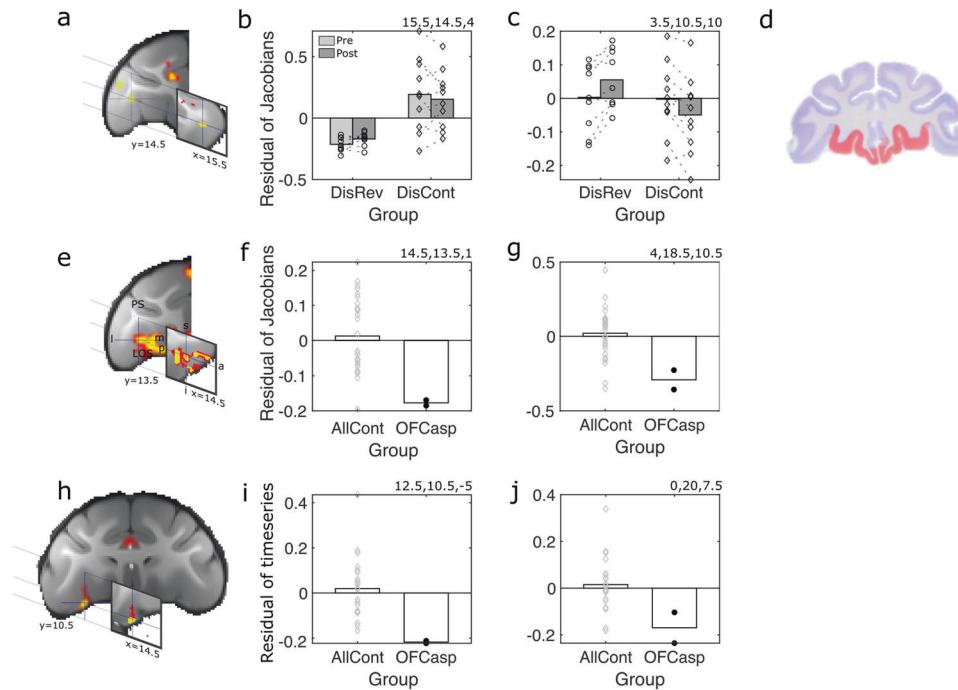


Fig. 2 Structural and functional changes occur in 47/12o and ACC as performance on reversal tasks improves and the same brain areas are compromised by lesion that have traditionally been thought to impair reversal task performance. **a** Coronal section illustrating areas of relative grey matter increase in 47/12o (shown also in inset) and ACC associated with the learning of discrimination reversals (as opposed to reversal learning tasks). Grey matter increases in 47/12o (**b**) and ACC (**c**) during reversal learning (left-hand side of each panel) but decreases in animals that perform a single discrimination learning task without reversals over the same time period (right-hand side of each panel). The dependent variable plotted on the ordinate is a scalar variable representing how much each voxel in an individual's brain needs to be expanded or compressed to match the group average brain. It is a residual effect after controlling for age and sex. Because in the initial stages of analysis, all scans from both before and after training are registered to a template derived from their group average, the baseline residual Jacobian values in each figure lie close to the mean. **d** A lesion in orbital and ventromedial prefrontal cortex placed between the lateral orbital and rostral sulci leads to grey matter decreases in two animals in 47/12o (**e, f**) and ACC (**g**) compared to grey matter levels in a group of 28 controls. The orbitofrontal and ventromedial prefrontal lesions also changed the pattern of activity coupling between anterior insula and 47/12 (**h, i**) and ACC (**j**). Control and lesion data are shown on the left and the right respectively in panels **f, g, i, and j**.

inputs [23], though note this remains to be tested at a granular level across the entire striatum.

There is increasing evidence that 47/12o is important for object-value discrimination reversal learning. Even though it might be intact when aspiration lesions are made in OFC, its interactions with other brain areas are disrupted by the aspiration lesions that impair object-value discrimination reversal learning; such aspiration lesions damage the white matter pathways that link 47/12o to other parts of prefrontal cortex, anterior insula, and ACC. It is this extended neural circuit, comprising in part ACC and 47/12o, that is important for performing reversal tasks. This was demonstrated by Sallet et al. [24] in two ways. In one experiment they used magnetic resonance imaging (MRI) scans to look for regions of structural change across the frontal lobes of macaques that had been trained to perform reversal tasks. They found that, with training on discrimination reversal task, grey matter in 47/12o expands (Fig. 2a, b). Rather than neurogenesis, grey matter expansion is likely to reflect changes in dendritic organisation, myelination, and angiogenesis [25–28]. Interestingly, grey matter changes also occur in a number of other locations in the frontal lobes including ACC (Fig. 2c). Moreover, activity coupling between these areas increases when animals learn to perform discrimination reversal tasks efficiently. In the second experiment, Sallet et al. took a different approach and examined the impact of aspiration lesions placed in the OFC region between the lateral orbital sulcus and rostral sulcus (Fig. 2d, red area). The lesion in this region had an impact on grey matter in adjacent areas such as 47/12o (Fig. 2e–j) and ACC (Fig. 2g) – in other words, it had an effect on the same regions that were associated with practised performance

of the object discrimination reversal task. The lesions also dramatically altered the pattern of activity coupling between 47/12o, ACC, and the anterior insula (Fig. 2h–j).

In summary, interactions between 47/12o and ACC are important for object-value discrimination reversal learning. Sallet et al. [24] were unable to see clear differences in impact of either learning object-value discrimination reversal or space-value discrimination reversal learning. Lesion studies, however, suggest that orbital/ventrolateral areas such as 47/12o and ACC are, respectively, more important during object and spatial learning [29]. The changes seen in both areas in Sallet et al.'s study may reflect the fact that regardless of the experimenter's intention, animals attempt to learn both object-dependent and space-dependent reward contingencies even when only type of contingency is important [30]. It is possible that while animals may genuinely be learning object-reward associations in object-value discrimination reversal learning, they may be learning whether to repeat a choice or change a choice. It is known that ACC is especially important when monkeys or humans decide whether or not to repeat a choice or switch to an alternative [31–33]. Such a view of ACC function is consistent with arguments, which we consider below, suggesting that the ACC may guide information seeking and decisions about whether or not to explore or engage in the same behaviour as previously. While several interesting ideas are emerging about the role of a number of other adjacent prefrontal areas, for example, dorsolateral prefrontal cortex [34, 35], here we will particularly focus on 47/12o and ACC and the way in which they contribute to learning and behavioural change. The key point of our arguments and

theories is that these cortices provide an anatomical substrate that links (i) the control of information seeking behaviour with (ii) dynamic processes to link choices and predictive environmental events to outcomes such as food rewards that are important for animals. We next preview the difficulties agents face when regulating their behaviour adaptively in dynamic contexts, and then get more specific about the roles of ACC and 47/12o in behaviour.

REWARD, SURPRISE, AND REWARD UNCERTAINTY INFLUENCE FLEXIBLE BEHAVIOUR

One of the reasons that ACC and 47/12o are so important for the flexible behaviour patterns that are needed for object-value reversal tasks is because they identify the stimuli in the environment to which it is important to attend because they are predictive of subsequent important events, such as food rewards. We know an increasing amount about the neural mechanisms that enable humans and other animals to allocate cognitive resources or “attention” to particular features or spatial locations in the environment. But we are still learning how they ascertain what is worth attending to [36, 37]. One long-established line of argument is that we attend preferentially to objects that are certain or *valid* predictors of important outcomes such as gaining access to food rewards [38]. However, another proposition is that unpredictable or uncertain events preferentially attract attention over certain ones [39]. Evidence suggests that humans and other primates attend to both types of predictors but that they do so at different times during tasks, depending on the sources and types of values and uncertainties [40–42]. In a new situation or when encountering new options, humans and monkeys first seek information to resolve uncertainty – this is facilitated by gaze and attention towards, for example, new or uncertain objects [43–45]. Often, over time, as they learn the objects’ values and predictions, they begin to show attentional bias to the most certain predictors of important outcomes based on their absolute value [43, 46–51].

In naturalistic volatile environments, outcome distributions often change without warning just as they do in reversal tasks or in tasks in which new stimuli are continually encountered (for Discussion [45]). Therefore, humans and monkeys often engage or attend to uncertain stimuli in order to find out more about them, even when these stimuli on average, may be associated with no or lower expected value in the short term [40, 42, 44, 45, 52, 53]. In other words, monkeys and people make choices not just to obtain reward but also to reduce their uncertainty about future choices and outcomes [37, 45, 52, 54–59].

One type of uncertainty is sometimes called irreducible uncertainty. It is a consequence of stochasticity or variance inherent in the environment and it cannot be changed by the human or animal decision maker. For example, a card player picking a card from a full set of playing cards might be able to estimate the probability that the card will bear a number higher than five but there is inherent variance, or risk, in the outcome of the card selection and that will not change as long as the deck remains the same. However, even though this type of risk does not diminish as more information is acquired, we know that it is still tracked; neural activity reflects irreducible risk [45, 60, 61]. In other cases, however, uncertainty may be reducible if further observations can be made and more can be learned about the frequency with which different outcomes follow a choice or predictor. This may occur when environments are volatile [42, 45, 62–64]. For example, if a card player knows that a deck is either stacked with more high number cards or more low numbers then they can work out – become less uncertain about – what type of deck is in front of them as they sample more and more of the cards.

Whether outcome uncertainty in an environment is reducible or irreducible, it is often adaptive to learn as much about it as

possible, particularly when big rewards may be available [45, 52]. When there is reducible uncertainty, it is worth exploring what happens if a new choice is taken in case it might prove to be a better predictor of reward. In the longer term there is instrumental value in exploring what happens when we take choices about which we are uncertain; [42, 64]. However, even if they cannot change the risks associated with irreducible uncertainty, animals are most likely to forage effectively if they have good internal models of various predictors in their environments and the choices they might take. Even when information may have no instrumental and extrinsic value, humans and animals still engage in non-instrumental information seeking [52–54, 57, 59, 65, 66]. They preferentially attend to reward-uncertain objects over reward-certain objects with known values in order to learn about and reduce their reward uncertainty [44, 53].

Outside the laboratory, in order to make informed decisions, it is useful for a decision maker to know not just the average outcome value of a choice but also the outcome variance or risk even if the decision maker cannot change that risk through any immediate choice that it takes (because the risk is irreducibly uncertainty). If the animal builds a good model of its environment and knows which options are risky and which are not then it might be able to use that information in the future. For example, in some situations in the future there might be an opportunity to pursue either an option that the animal has discovered is risky or an option that the animal has discovered is not risky. Depending on the context and factors such as the current metabolic budget, it might be better to pursue the risky option but on other occasions it might be better to pursue a small but non-risky reward goal [45, 67–71] – and early reduction of the current uncertainty or risk may also serve to regulate risk attitude during subsequent choices.

Natural environments are complex and contain many stimuli any of which might have predictive value and be used to guide behaviour. Similarly, at any moment in time, or in quick succession, an animal may make more than one action and so it may be difficult to identify the stimuli or actions that are predictive of the outcomes that follow. One way to deal with this problem is to construct hypotheses or beliefs about a particular choice or predictor and its causal relationships. Such hypotheses then determine the choices that are attempted and the learning that is possible [72–76]. While hypothesis-driven learning may be necessary, it can also be problematic; if people or other animals can only learn about their hypotheses then they may end up with confirmation biases and be unable to learn about predictors that lie beyond the purview of their hypotheses [73].

THE ACC, UNCERTAINTY, AND INFORMATION SEEKING

Evidence that the ACC has a role in instrumental-information seeking, during outcome learning, and during non-instrumental information seeking (motivated by the desire to update one’s beliefs for their own sake) is available for humans, non-human primates, and rodents [40, 46, 52, 65, 77–80]. For example, Trudel et al. [40] taught human participants to choose between pairs of predictive stimuli that they could use to help them find the locations of tokens that ultimately led to monetary rewards. When the session began, participants were more likely to select predictors about which they were uncertain. This was especially the case when participants knew that the current session would be a long one and so that there would be plenty of time in which to use the information that was gained. Activity in ACC reflected the difference in participants’ uncertainties about what each predictor foretold about the value of a choice.

Such information-seeking is not just a feature of human behaviour but is also readily observable in other primates such as macaques. Stoll et al. [80] taught monkeys to perform a simple discrimination task to obtain small juice rewards. In addition, however, they learned to monitor a reward bonus predictor

stimulus that gradually, over time, grew in size. The monkeys could obtain a larger juice reward from this stimulus when it eventually reached its largest size.

Compared with dorsolateral prefrontal cortex activity, activity in ACC neurons was a worse indicator of which choice the monkeys would take in the simple discrimination task but a better predictor of whether they would seek information about the current size of the bonus reward stimulus. These studies provide clear evidence for the role of ACC in instrumental information seeking in which humans and animals reduce their uncertainty about future outcomes or gain information to update their choice strategy.

Intriguingly, ACC activity tracks uncertainty, and relatedly the value of exploring, even when a person is not currently exploring but making exploitative choices aimed at maximising immediate reward [40]. Similarly, ACC activity in humans [32, 81] and monkeys [31] tracks how good it might be to make a different choice to the one that is currently being taken even if the alternative choice is not actually taken. However, when a person does switch into a more exploratory mode of behaviour, uncertainty signals manifest more widely in other brain areas beyond ACC, such as ventromedial prefrontal cortex (vmPFC), in which activity reflects reward expectations during periods of exploitative, immediate reward-maximising behaviour [40]. Some ventral vmPFC neurons may treat certain and uncertain rewards similarly, in a manner suggesting they are coding reward possibility rather than probability, perhaps in the service of promoting exploration [82]. If this signal arises due to the integration of uncertainty and expected value, the ACC could be in large part the source of these modulations [45, 56]. Also, the activity of neurons in several brain regions linked to reward-guided choice selection, such as the amygdala, ventral striatum, and orbitofrontal cortex exhibit exploration-related changes in activity (or are causally related) to behaviours in which monkeys switch into a relatively more exploratory mode of behaviour [42, 64]. Future studies must assess how interactions with different parts of the PFC, particularly with the ACC and vPFC contribute to their functions. Also, how ACC activity mediates neural dynamics in action selection regions of PFC that change their states during exploration vs exploitation, such as the frontal eye field [83], will be an important direction of research.

Non-instrumental information seeking does not directly impact future reward [52], and yet it is also ubiquitous among humans and animals. Many reasons for this have been proposed, among them are that: (i) dynamic updating of beliefs is helpful when contexts change without warning, (ii) relatedly, treating information as rewarding in itself could have become evolutionarily selected because hypotheses and belief formation have utility (as in *i*). For example, even if an animal cannot use the information it acquires in one setting in an instrumental manner it may often, in natural settings, be able to exploit its knowledge of these predictive stimuli in an instrumental manner in the future when it may have the opportunity to seek these stimuli as opposed to others. Interestingly, non-instrumental information seeking and reward expectancy [52], at some levels of the neural architecture, overlap. For example, phasic dopamine activations reflect the values of rewards and also signal the possibility to resolve reward uncertainty earlier than expected even in a task in which the physical rewards are delivered at the same time [66]. Hence, the question arises, are there distinct circuits that control non-instrumental information seeking, motivating our desire to know what the future holds?

To answer this question, White et al. designed an information viewing procedure in which monkeys exhibited uninstructed gaze behaviour that reflected their anticipation and desire for uncertainty-resolving information (Fig. 3). In this procedure, there was no way for the monkeys to use information or their gaze patterns to influence reward magnitude. Certain and uncertain rewards were predicted with visual objects. On some occasions when an animal saw an uncertain predictor object it was followed by a second object that resolved the monkeys' uncertainty about whether reward would be received before the time for reward itself

arrived. During certain trials, similar visual cues appeared and again they were followed by a second object but now the second object provided no further prediction information about the reward amount that would be delivered at time of the outcome. The monkeys' gaze was strongly attracted to uncertain objects, especially in the moments before receiving information to resolve the uncertainty. This attraction was preferentially related to anticipating uncertainty resolving information, not reward value [44].

Integrating this result with other reports in the literature, one concludes that when primates first encounter a visual object, its relative importance (or absolute value, partly based on stored memory) has a great impact on gaze. In addition, reward uncertainty of the object has a particularly powerful effect on gaze, especially before the time of uncertainty resolution. This suggests that the effect of uncertainty (or surprise) on attention is partly related to the animal's expectations about receiving information that will resolve the uncertainty.

Electrophysiology, pharmacology, and computational modelling have revealed that this type of gaze behaviour is dynamically controlled by the ACC and interconnected regions in the basal ganglia, in the dorsal striatum and pallidum, on a moment-by-moment basis. White et al. [44] found that a subset of neurons in these regions predicted the moment of uncertainty resolution. Crucially these neurons ramped to the time of expected information acquisition but exhibited less ramping activity when animals expected certain outcomes or no information predictions (Fig. 3). Most strikingly is that this ramping activity was linked to information seeking gaze shifts on a moment-by-moment basis: the signal increased before gaze shifts towards informative cues and relatively decreased before gaze shifts away from them.

The ACC predicted information-seeking earlier than the basal ganglia suggesting that its information prediction inputs are crucial for the control of information seeking actions, ultimately implemented by the basal ganglia circuitry. Consistent with this idea, inactivations of the basal ganglia regions that received ACC inputs and contained these signals diminished information-seeking behaviour.

Consistent with the supervisory role of ACC in non-instrumental information seeking, a recent study showed that during decision making, information sampling may be relatively more related to the activity of ACC neurons than to neurons in other prefrontal regions such as OFC (area 13) and dlPFC [58].

At various points, we have referred to the need to seek information in order to test hypotheses or form beliefs. Is it reasonable to think that a monkey is really generating hypotheses about which choice to take as opposed to exploring more randomly, or to suggest that they are motivated to form beliefs for their own sake, assigning value to non-instrumental information? We believe that they do and that the ACC is involved in these processes as discussed elsewhere in this review [45, 52, 56, 84].

A number of observations suggest macaques hold hypotheses about which is the better choice to take next. When macaques attempt to identify which one of four locations is the correct one in which to find reward, they do not appear to explore all the locations randomly but instead they appear to assess each possible location systematically one after the other [85]. When there are two possible alternatives to pursue, then ACC activity in the macaque represents the better and worse alternatives in two very different ways that are consistent with the idea that the macaque is focussing on the possibility – the hypothesis that guides their instrumental information seeking – that one alternative rather than the other would be better to explore [31]. Sarafyazd and Jayazeri [86] taught monkeys a task in which errors could occur either as a consequence of misjudging which stimulus had been seen or because of a genuine change in stimulus-response contingencies. ACC neuron activity reflected the evidence for hypotheses concerning stimulus-response changes and not simply error occurrence.

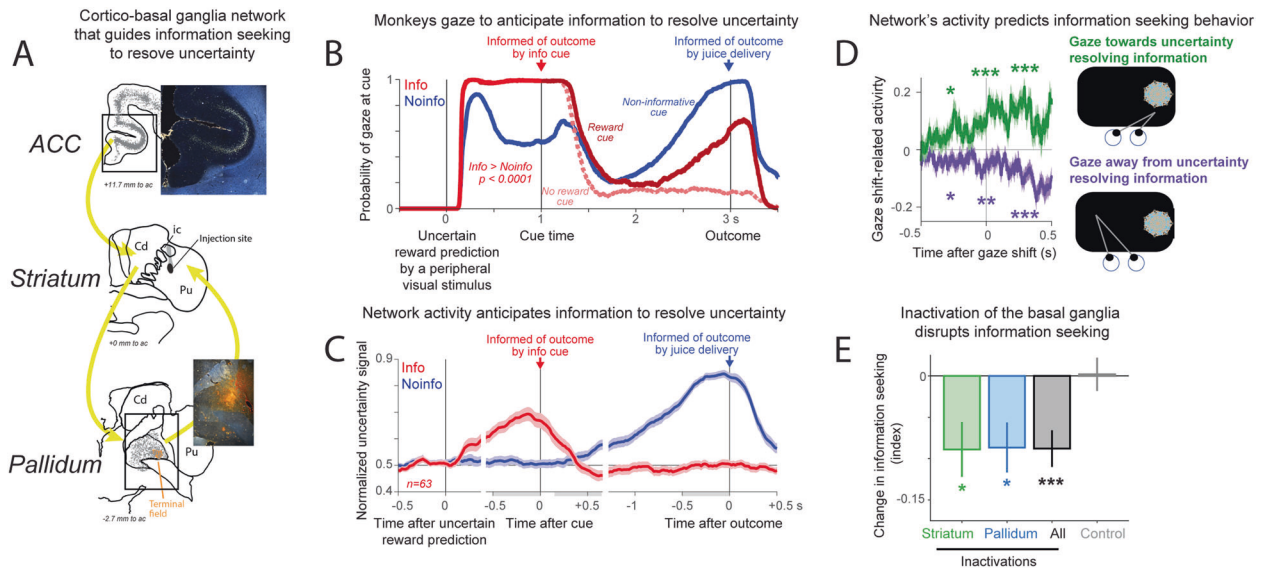


Fig. 3 **A** A primate cortico basal ganglia network for information seeking. **A** Example of bidirectional tracer injection into an uncertainty-related region of the striatum reveals connectivity with anterior cingulate and anterior ventral pallidum. **B** Monkeys' uninstructed gaze behaviour in an information observation task reveals motivation to seek advance information about uncertain rewards. On Info trials (red), a peripheral visual stimulus predicted uncertain rewards. One second after, it was replaced by uncertainty-resolving cue stimuli (see red arrow). Monkeys' gaze on Info trials was attracted to the location of the uncertain prediction in anticipation of receiving informative cues that resolved their uncertainty. After uncertainty resolution in Info trials, gaze is split to trials in which reward will be delivered (dark red) and reward was not delivered (pink). On Noinfo trials (blue), another visual cue also predicted uncertain rewards. Here, the subsequent cue stimuli, shown 1 second after, were not informative and the monkeys resolved their uncertainty at the time of the trial outcome (blue arrow). In Noinfo trials, gaze was particularly attracted to uncertain visual stimuli in anticipation of uncertainty resolution by outcome delivery. Probability of gazing at the stimulus ramped up in anticipation of the uncertain outcome until it became greater than other conditions (50% > 100%; compare blue with dark red). **C** Neural activity across the ACC-Striatum-Pallidum network anticipates uncertainty resolution. Same format as **B**. **D** The networks' activity predicts information-seeking gaze behaviour. Mean time course of gaze shift-related normalised network activity aligned on gaze shifts onto the uncertainty resolving informative stimulus (green) and off the stimulus (purple). Asterisks show significance in time windows before, during, and after the gaze shift. *, **, *** indicate $P < 0.05$, 0.01, 0.001, respectively (signed-rank test). Uncertainty-related activity was significantly enhanced before gaze shifts on the stimulus and reduced before gaze shifts off the stimulus when animals were anticipating information about an uncertain reward. **E** Pharmacological inactivation of the striatum and pallidum regions enriched with information-anticipating neurons disrupted information-seeking relative to saline control, which produced no effect. Asterisks show significance in time windows before, during, and after the gaze shift. *, **, *** indicate $P < 0.05$, 0.01, 0.001, respectively (signed-rank test). Error bars are ± 1 SE. Figure reproduced from Monosov, 2020, Trends in Neuroscience.

As well as reflecting hypotheses about which course of action should be taken, ACC activity also reflects expectations about what the outcome of the course of action might be. For example, some ACC neurons signal expectations (and breaches in those expectations) about rewards while others signal expectations (and breaches in those expectations) about aversive outcomes [46]. We return to the topic of breaches in expectations – prediction errors – in the next section.

Finally, abrupt switches in behaviour that occur when new information is acquired suggest that a hypothesis has been formed and tested and either confirmed or disconfirmed. We know that such changes in behaviour occur in humans [74, 75, 79], monkeys [86], and rodents [77, 78] and in each case they are linked to ACC activity changes.

To summarise, non-instrumental information seeking, requires agents to detect uncertain outcomes, predict the time of uncertainty resolution, and mediate action selection relative to that prediction, to gather uncertainty-resolving information [52, 57]. White et al. found that these functions were reflected in the activities of single neurons in the ACC and in the activities of neurons in the basal ganglia located in regions that receive dense inputs from the ACC [87–89].

ACC AND PREDICTION ERRORS

A key feature of adaptive behaviour is flexible learning (adaptation) on the basis of heterogeneous feedback signals that may

operate across multiple time scales [62, 90–93]. In support of these wide-ranging functions, the brain contains systems for motor adaptation, sensory predictions, and for learning or associating abstract variables, such as subjective values, and policies, with potential choices. This latter type of learning is uniquely dependent on reward prediction errors [94, 95] – or signals reflecting comparisons between expected outcomes and received outcomes. Signals that in some ways resemble reward prediction errors (RPEs) are present in several brain regions including ACC and prefrontal cortex regions such as 47/12o [56, 96–98].

RPE-related signals come in distinct *flavours* [47, 56, 86, 99–101] such as signed and unsigned RPEs, and can be mixed with contextual and task-related variables. Generally, signed RPEs signal how much better or worse an outcome is from the decision maker's expectation, such as most famously observed in the dopamine neurons in the medial substantia nigra (SN) and ventral tegmental area (VTA) [94, 95, 102]. This signal is well suited for updating beliefs about the utility of future actions or the value of an option for subsequent behaviours.

By contrast, unsigned RPEs signal the magnitude of the difference between an outcome and the prior expectation; they indicate how surprising the outcome is given prior expectation rather than the more specific question of how much better or how much worse is the outcome compared to prior expectation. Dopamine neurons, in the caudal lateral part of SN encode a signal that resembles an absolute RPE in some important ways, most salient of which is its rapid phasic increase in response to either

unpredicted rewards or punishments [47, 103]. Broadly, surprise-related signals have also been observed in the basal forebrain [104–108] and potentially in the locus coeruleus [109, 110] which all project to the prefrontal cortex.

Because they signal surprising events, unsigned RPEs are determinants of what features in the environment merit further attention; unsigned RPEs identify stimuli that might be worth learning more about. In line with this suggestion, larger unsigned RPEs predict better memory [111] (but also see [112]) and drive attention [39, 43, 46, 47, 100, 113, 114]. One other useful way to utilise surprise signals is to estimate uncertainty – which can be conceptualised as emerging from the integration of surprise signals over time. Of course, this computation can then guide how we learn (e.g. learning rate) and even modulate synapses – by preparing groups of neurons to learn from future reward value prediction errors [45, 115].

Converging evidence from neurophysiology and imaging suggests that ACC neurons signal motivational variables closely resembling many types of RPEs. However, the ACC RPE signals are uniquely context-sensitive and behaviourally-dependent subserving adaptive behaviour in volatile contexts in which the integration of feedback across multiple time scales and contexts is crucial [90]. Furthermore, the ACC contains groups of neurons that distinctly encode second-order statistical information that is crucial for such adaptive learning behaviour, in particular, how uncertain an agent is about outcomes in its future and how subjectively valuable the resolution of that uncertainty might be and even the source of the uncertainty.

For example, Kennerley et al. [116, 117] analysed ACC neurons' responses during the offer presentation and choice outcome epochs of a decision-making task in which monkeys chose among offers that varied in probability, amount, and effort. They noted that ACC neurons that were excited by increases in chosen offer value during offer presentation, were also most likely to display excitation that scaled with unexpectedly positive probabilistic outcomes suggesting that their offer responses may also be explained by offer-related changes in animals' estimates of the current state value. In other words, the ACC neuron activity in response to a new offer reflected how surprisingly good that offer was in just the same way that ACC neuron activity reflected how surprisingly good the reward outcome was. So, might both offer activations and outcome activations in ACC simply reflect RPEs? We believe that there are many differences between ACC outcome/feedback signals and "standard or classical" RPEs. These differences are important because they give important hints about how the ACC utilises RPEs (from neuromodulatory inputs) and how it guides behaviour.

First, the ACC signals display great context sensitivity, displaying selectivity for the sources of decision uncertainty. When the variability in outcomes was related to several sources of uncertainty, such as volatility due to changes in task rules, and noise [86, 118], ACC signalled prediction errors most prominently when errors were due to changes in task rules (or volatility) [119, 120]. Also, a classic study by Shima and Tanji [121] found that ACC neurons' outcome signals predicted stay-or-switch decisions in an action-outcome association task in which the values of actions changed. Admittedly, while this activity modulation could be due to additional hidden states or dynamic processes associated with behavioural adaptation (such as "attention"), integrating these results with lesion studies [122] and human neuroimaging [123–125], points strongly to the idea that ACC RPE-like outcome signals are tied to behavioural adaption (learning and switching).

To make the differences between standard RPEs and ACC activity clearer, the ACC was functionally contrasted with the lateral habenula, a key source of RPEs in dopamine neurons [126]. During reversal learning of probabilistic action-outcome contingencies, monkeys benefited from the integration of past

outcomes to decide which action to take. ACC neurons reflected this integration: increasing their outcome-related signals as a function of trial number towards behavioural adaption (switches in behaviour). But surprisingly, the lateral habenula tracked mostly single-trial outcomes by encoding negative value reward prediction errors of the current trial's outcome. Consistent with this observation, Kennerley et al. [122] reported that ACC lesions did not prevent monkeys from noticing that a reward outcome had failed to be delivered but instead they disrupted their ability to integrate the feedback information received across trials in order to work out the best choice to take in a changing environment.

The dynamic nature of outcome signals in the ACC has also been observed in tasks resembling foraging. There, some ACC neurons respond to feedback when animals explored choice-reward associations but the same neurons responded to trial starts when the animals changed behavioural model and began to make reward-guided exploitative choices [85].

Such dynamic across-trial modulations of outcome signals in the ACC could reflect the dynamics of a credit assignment process that under volatility ought to be related to evidence accumulation across multiple trials and behavioural states.

Another feature of ACC outcome activity that can facilitate its functions in adaptive behaviour and credit assignment is that many ACC neurons display valence specificity – that is instead of signalling an abstract error in subjective value, divorced from information about the valence of predictions, ACC neurons can signal prediction errors either about rewards, aversive noxious punishments, or both [46]. This means they can signal the source of the change in the agent's experience which may differ dramatically across valence. For example, it has been proposed that though, for example, unexpected reward deliveries and unexpected punishment omissions may both have positive subjective value, the behavioural and internal responses of these distinct RPEs may dramatically differ [127]. Therefore, the source or valence selectivity in ACC outcome signals may further assist the agent in regulating behaviour and emotion. Finally, ACC also contains neurons that signal absolute RPEs [46, 128], while valence-specific outcome signals could regulate valence-specific behavioural and emotional responses, absolute RPEs can drive attention and arousal. Hence, ACC could have highly flexible context-dependent control over reinforcement learning.

A final important point is that the results reviewed thus far indicate that, at the level of circuit or brain area, information predicting and outcome encoding functions are linked by both being present within the ACC. But an important element of flexible control exerted by ACC is that these functions are handled by distinct groups of neurons within ACC [46, 56]. But this fact only becomes apparent when we move beyond measures of the aggregate activity to look at the information encoded in individual neurons.

USING INFORMATION ABOUT THE OUTCOMES OF CHOICES TO GUIDE FURTHER CHOICES IN THE FUTURE

So far we have focused on ACC but we have seen that area 47/12o is part of a circuit that includes ACC that is important for behavioural adaption. In order to understand its activity during behavioural adaptation, Chau et al. [129] used fMRI to record across the whole brain of macaques that were performing either an object discrimination reversal task or probabilistic object-reward association learning task. In the second case, animals learned about three objects that were each associated with different probabilities of reward that changed and drifted over the course of a testing session. Only two of the three possible options were available to choose on any given trial so this meant that animals were forced to change the choices they took from one trial to the next but, at the same time, feedback from choices made in the past was useful for determining which choices to take

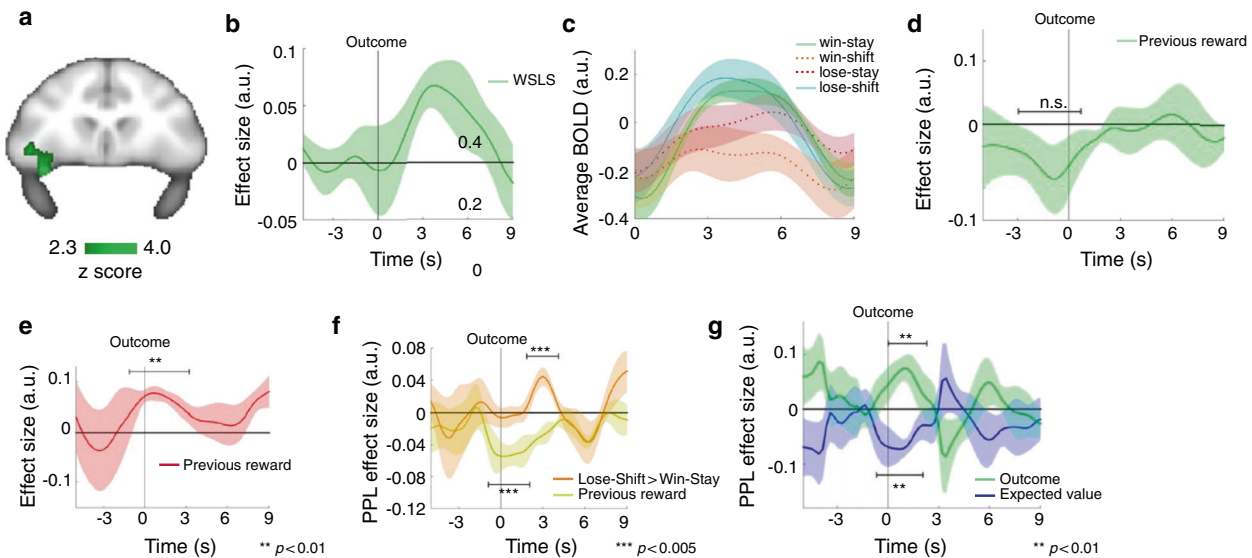


Fig. 4 Activity in 47/12o reflects credit assignment during reversal learning. **a** Activity in area 47/12o reflects whether the outcomes of choices will guide the subsequent decisions macaques take in an adaptive as opposed to maladaptive manner (WLS: win-stay/lose shift versus lose-stay/win shift behaviour). **b** Time course of change in blood oxygen level dependent (BOLD) in 47/12o time-locked to the WLS effect. The haemodynamic response function means that the BOLD signal peaks 3–4 s after the neural activity with which it is associated. **c** The WLS-related BOLD changes are decomposed into the effects associated with each individual behavioural strategy. While 47/12o activity at the time of one choice outcome does not reflect whether or not reward was delivered after the previous choice (**d**), amygdala activity does reflect previous outcomes (**e**). **f** The patterns of activity coupling between 47/12o and amygdala associated with previous reward are negative (yellow). Such a pattern is consistent with the possibility that 47/12o inhibits representations of previous rewards in amygdala. Positive activity coupling between 47/12o and amygdala on lose-shift trials. **g** The same information is not reflected in the patterns of interaction between 47/12o and ventral striatum, but activity coupling between these two areas reflects prediction errors – whether an outcome is better or worse than expected. This is apparent in the fact that activity at the time of an outcome is a negative function of the prior reward expectation and a positive function of actual reward delivery.

in the future. Because changes in reward probability occurred infrequently or gradually, a good guide to which choice to take next was whether or not the same choice had led to reward in the past.

Each time macaques discovered that they had or had not got a juice reward for the choice that they had taken, there was activity across OFC and vIPFC. However, Chau et al. also carried out a second analysis of outcome-related activity; they looked for outcome-related neural activity as a function of whether it would lead to an adaptive pattern of subsequent behaviour. One type of adaptive behaviour occurred when a choice was rewarded and it was subsequently taken again on the next occasion it was offered (win-stay trials). Another type of adaptive behaviour occurred when unrewarded choices were not repeated on the next occasion that they could have been taken (lose-shift). By contrast, repetition of unrewarded choices (lose-stay) and failures to repeat rewarded choices (win-shift), are maladaptive patterns of behaviour. 47/12o activity time-locked to outcomes was significantly higher when behaviour was adaptive as opposed to non-adaptive (win-stay and lose-shift versus win-shift and lose-stay) in both the object discrimination reversal task and in the probabilistic object-reward association learning task (Fig. 4a, b). In summary, whether or not the task is a reversal task is not really important. Reward- and outcome-related activity is widespread but activity in 47/12o at the time of an outcome reflects whether that outcome will be used to guide subsequent decisions next time the same choice is available.

The results suggest that 47/12o is linking a choice to an outcome and its activity determines whether that choice will be taken again in the future. This function may reflect the unusual set of anatomical connections that 47/12o has. On the one hand, it is in receipt of the information that is critical for linking representations of choices and rewards. It receives information about both the component features and the overall form of visual objects

from temporal association cortex area TE and perirhinal cortex [19, 21, 130] but in addition, it is interconnected with ventral striatal, amygdala, and hypothalamic regions associated with reward processing [131–135]. In combination the connections enable 47/12o to forge associations between specific objects or choices and specific occurrences of reward.

Whole-brain imaging methods like fMRI make it possible to examine interactions in activity between 47/12o and some of these other areas at the time of choice outcomes and it is clear that these interaction patterns reflect aspects of adaptive behaviour. For example, activity coupling between 47/12o and amygdala occurs when an absence of reward will lead to lose-shift behaviour (Fig. 4c) while activity coupling between 47/12o and ventral striatum reflects the occurrence of prediction errors – outcomes that are better or worse than expected and which suggest that more learning is needed (Fig. 4d). However, it is also important to note how activity in some of these other areas is also different to 47/12o.

One important difference between 47/12o and amygdala responses to reward is that 47/12o activity especially reflects the outcome that is currently most relevant for behaviour. For example, at the moment that an animal finds out whether or not its choice is being rewarded, 47/12o reflects that particular outcome event. By contrast, at the time that the outcome on one trial might arrive, amygdala activity does not just reflect that event, instead, it also reflects whether or not the previous trial was rewarded. The pattern of activity seen in 47/12o could mediate very precise assignment of the credit for a particular occurrence of reward to a particular choice. By contrast, activity in the amygdala, or at least some parts of it, is sufficient to identify if a choice is made in the presence of reward but it might not be sufficient to establish whether or not the choice itself led to the reward. In many situations, knowing simply that a choice was made in the context of reward is all that is needed because foraging animals

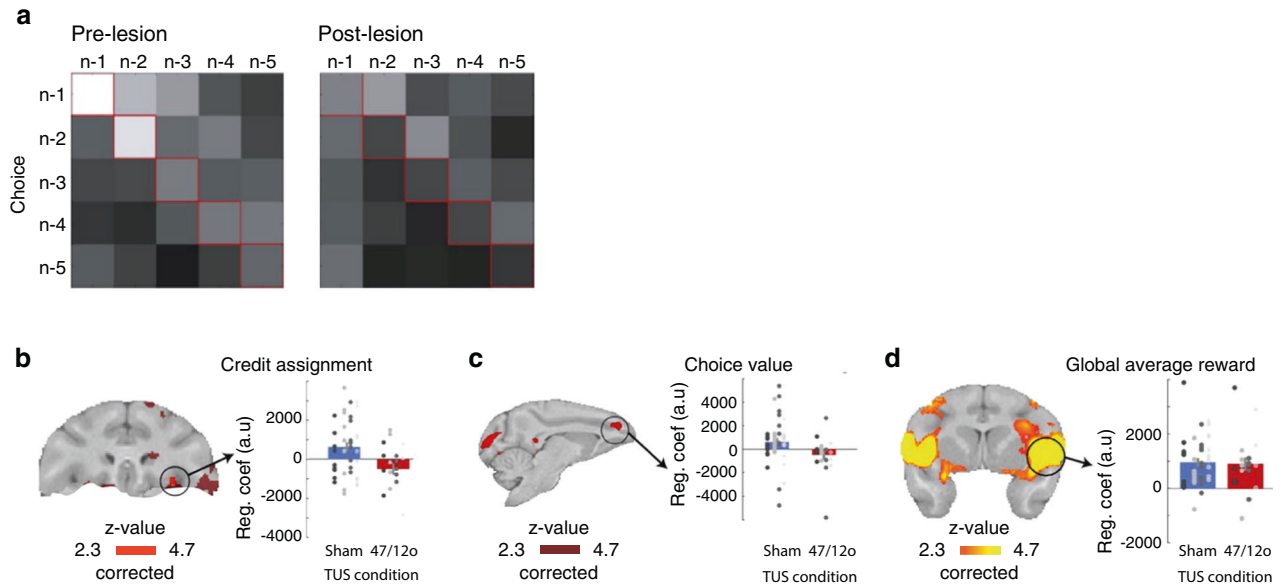


Fig. 5 Lesions or inactivation that affect 47/12o affect credit assignment-related behaviour and neural activity across the brain but leave other reward-related signals intact. **a** The brightness of the squares indicates the weight of influence exerted by past choices and rewards on the next choice made. The diagonal outlined in red indicates conjunctions between choices made and rewards received on the same trial. If these squares are bright then it indicates that if a choice was rewarded, then it is likely to be made again. Off-diagonal effects indicate spread of reward effects. Bright squares above and to the right of the diagonal indicate that choices are more likely to be repeated even if they occurred after the reward was received. Bright squares below the diagonal indicate that choices that were made several trials before a reward was received are likely to be repeated. Normal pre-operative behaviour is shown on the left while the impact of lesions that compromise a number of areas including 47/12o are shown on the right. The lesions diminish the influence of past choice-reward contingencies but leave intact spread of reward effects and the general tendency to repeat choices made at approximately the same time as a reward, albeit on an earlier or later trial. **b** Credit assignment can also often be indexed by whether or not a WSLs strategy is followed. Such a strategy is associated with activity that is prominently in 47/12o. When TUS is directed to this region, however, that credit assignment-related WSLs activity in 47/12o is significantly reduced. **c** Application of TUS to 47/12o also disrupted choice-value encoding in ACC but **(d)** it left global average reward-related activity in anterior insula intact.

often make the same choice repeatedly in a given patch in the environment. In such situations, knowledge of the co-occurrence of reward with a choice is a good guide to the value of the choice. However, in other cases, for example when an animal is alternating between different choices in rapid succession, then precise reward credit assignment is essential for determining exactly which choice led to which outcome and is necessary in order to realise that the choice that led to the reward is more valuable than the choice that followed the reward.

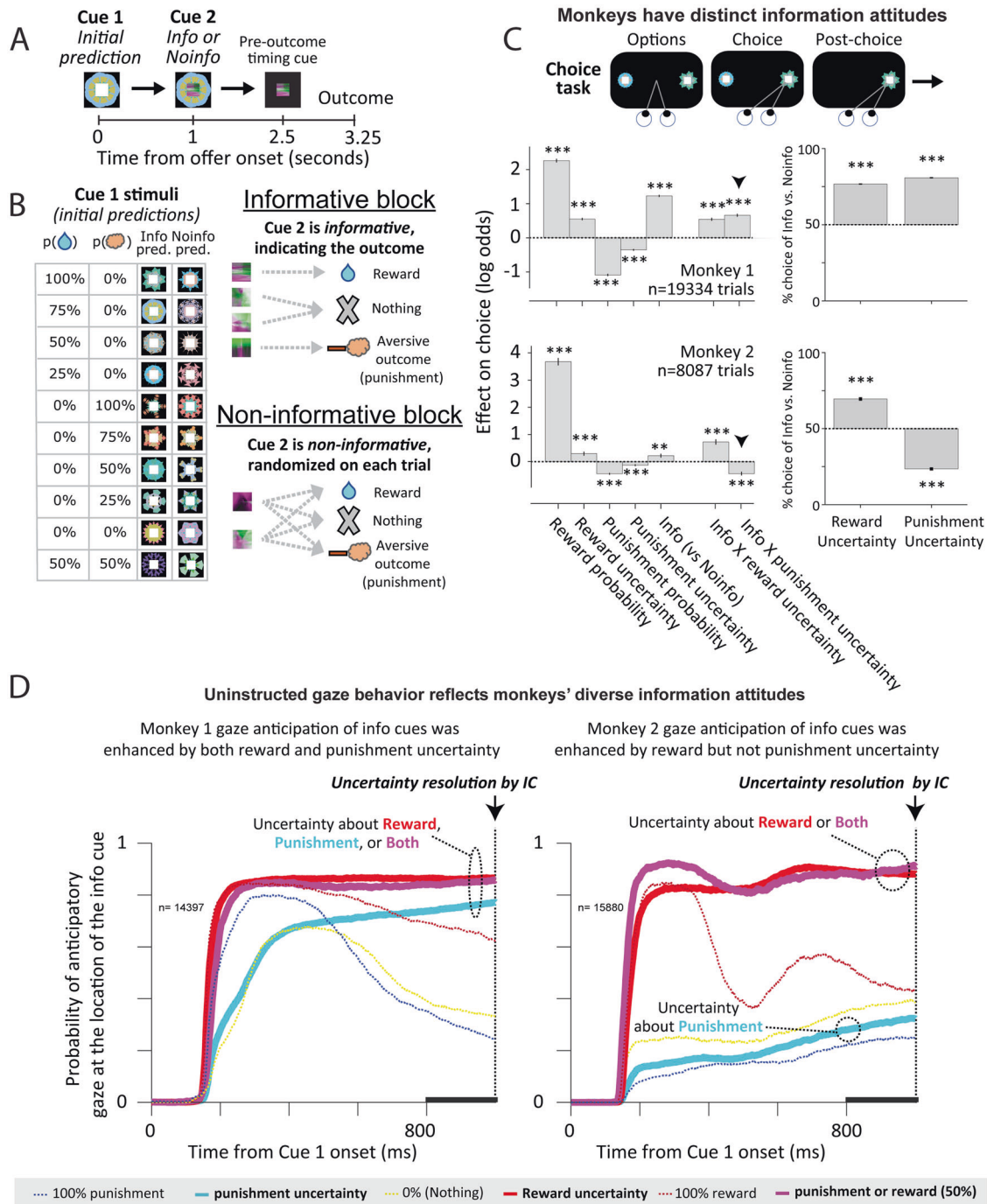
DISRUPTING 47/12O IMPAIRS CREDIT ASSIGNMENT

In order to understand the different ways in which choices and outcomes can be linked an illustration such as the one in Fig. 5a can be useful [136]. The histories of recent rewards and of recent choices are represented on the abscissa and ordinate respectively. The brightness of the squares in the grid indicate the influence that conjunctions of past choices and rewards have on the next choice that is taken. For example, the top left square indicates the influence that the identity of the choice taken on the last trial ($n - 1$), and whether or not a reward was received on the last trial ($n - 1$), will have on whether that choice is repeated. Because the influence is high, the square is bright. The relatively bright diagonal going from the top left to the bottom right (outlined in red) indicates that the identity of the choice taken two trials ago, or three trials ago, and whether it was rewarded also influences which choice will be taken next. The bright diagonal indicates the influence that reward credit assignment to a specific choice has on subsequent decision making. The areas off the diagonal, however, indicate failures in credit assignment. For example, the relatively bright square in the first row, second column indicates that whether or not reward was received two trials ago also influences

whether or not a choice made one trial ago will be repeated. In other words, the more general impact of choice occurring in the context of reward, rather than specific credit assignment, is indicated by the off-diagonal regions. The bright diagonal indicates Thorndike's [137] law of effect: reinforced choices are repeated. The off-diagonal indicates Thorndike's [138] spread of effect: the reinforcing effect of the reward spreads beyond the choice that led to it and extends even to adjacent choices so that they are more likely to be made too.

Lesions that disrupt 47/12o impair credit assignment. The influence of the precise history of choice-reward outcomes on subsequent decisions is diminished and so the diagonal (outlined in red) is less bright (Fig. 5a, right). Establishing that 47/12o is especially important has taken some time. The first study to show an impact on credit assignment employed aspiration lesions that were centred on OFC rather than 47/12o [136] but which were likely to have compromised adjacent areas such as 47/12o [24]. Rudebeck et al. [11] made excitotoxic lesions that spared fibres of passage in either OFC or vIPFC. Area 47/12o was included in the vIPFC lesions but not the OFC lesions and these were the ones that produced the credit assignment deficit. Most recently Folloni, Fouragnan et al. [139] have used transcranial ultrasound stimulation (TUS) to target area 47/12o and shown that this is sufficient to impair credit assignment in the same way while leaving less specific spread of reward effects unaffected (Fig. 5c–e).

Once again it is possible to see evidence of interactions between 47/12o and ACC in the data recorded by Folloni, Fouragnan et al. Choice values may be learned via credit assignment in 47/12o but choice values are also represented in ACC. The way in which they are represented suggests ACC activity encodes how good it would be to switch behaviour to an alternative choice rather than taking the current choice once



again [31]. Activity encoding the relative evidence in favour of a change in behaviour in ACC is, however, disrupted after TUS of 47/12o. This is not because the ultrasound itself is non-specific; activity in brain areas immediately adjacent to the TUS target is not affected by TUS [140–144]. For example, Folloni et al. reported other types of reward-related activity in anterior insula (adjacent to the TUS-targeted area 47/12o) was unaffected by 47/12o TUS. Instead, the pattern of change suggests that choice value-related activity is transferred between 47/12o and ACC to guide the maintenance or changing of a course of behaviour.

If 47/12o is concerned with the representation of specific, contingent associations between choices and rewards, then where in the brain is responsible for spread of reward effects that remain

even when 47/12o is compromised by a lesion or TUS? There are a number of candidates. We have already noted that amygdala activity at the time of a choice outcome does not just reflect whether a particular outcome itself was a reward or non-reward but also whether previous outcomes were rewards [129]. Analogous patterns of activity have been reported in humans by Jocham et al. [145] and Klein-Flügge et al. [146] have reported an analogous spatial rather than temporal spread of reward effect; the human amygdala responds to stimuli that are spatially adjacent to those that, if chosen, lead to reward. The anterior insula and raphe nucleus may be other important areas; activity in the anterior insula reflects the 'global reward state' – the average value of an environment – as opposed to the value of specific

Fig. 6 Investigating information seeking to resolve uncertainty about rewards and punishments. **A** During each trial, 1 second after the appearance of a trial start cue (not shown), Cue 1 appears. After 1 second, Cue 2 appears inside of Cue 1. After 1.5 more seconds, the Cue 1 disappears indicating to the monkey that the outcome will soon be delivered (always in 0.75 more seconds). **B** During the informative block, 10 Cue1s yield 10 different chances of rewards and punishments (air puffs): 100, 75, 50, and 25% chances of reward, 100, 75, 50, and 25% chances of punishment, neither reward or punishment, or 50% chance of either reward or punishment. Cue2 is always informative and indicates the outcome. In the non-informative block 10 other Cue1s appear that yield the same 10 chances of reward and punishment, but are followed by noninformative Cue2s that do not predict the outcome, so uncertainty is not resolved until the outcome is delivered. **C** Two monkeys (top and bottom) chose between the Cue1s and their associated outcomes. Schematic of choice-related epochs of the choice task is shown above. Plotted are the weights from a logistic GLM fit to each monkey's choices based on the attributes of each Cue1, including its outcome probabilities, uncertainty (operationalized as standard deviation), information-predictiveness, and their interactions. Both monkeys were fit with similar patterns of weights, including positive weights of Info x Reward Uncertainty, with one exception: Monkey 1 had a positive weight of Info x Punishment Uncertainty, while Monkey 2 had a negative weight. Error bars are ± 1 SE, **, *** indicate $p < 0.01$, 0.001 . Right: % choice of informative versus non-informative Cue1 for trials where both options had reward uncertainty (left bar) or punishment uncertainty (right bar). **D** During Cue1 and subsequent cue presentations the monkeys were allowed to gaze freely, without their gaze having any influence on the outcome (Fig. 1A; [44]). As in our previous study, we found that during the informative block animals gaze is preferentially attracted to the location of the upcoming informative cue before it appears to resolve reward uncertainty (top, thick red line). Crucially, anticipatory gaze also reflected attitudes toward information to resolve punishment uncertainty. Monkey 1 (left) had prominent information-anticipatory gaze during punishment uncertainty (thick cyan line). In contrast, Monkey 2 only had prominent information-anticipatory gaze during reward uncertainty (right).

choice options [30]. It is such global reward state signals in anterior insula that, as mentioned above, are unaffected by 47/12o TUS [139].

ELECTROPHYSIOLOGY OF AREA 47/12: CREDIT ASSIGNMENT AND INFORMATION SEEKING

While only a few studies have recorded single-neuron activity from area 47/12o in behaving macaque monkeys some evidence is emerging that it provides a link between information seeking and credit assignment. Thus far the data reviewed have outlined the roles of 47/12 in credit assignment. However recent electrophysiological evidence suggests that it also has a prominent role in information seeking and this role may be in part different from that of the ACC.

Jezzini et al. [23] sought to study how information seeking is mediated by preferences for advance information about uncertain rewards and punishments and recorded in 47/12o and ACC.

Most studies thus far have concentrated on reward uncertainty – which necessarily is associated with reward gains or losses [45], but have not studied information attitude about aversive noxious punishments, which are distinct from reward losses and involve distinct effectors, emotional states, and learning strategies [45, 127, 147]. The few studies that studied information seeking about aversive noxious punishments did so independently of information seeking about uncertain rewards and so whether information attitudes about rewards and punishments are behaviourally and neurally linked has been unclear.

For example, behaviourally, we know that humans show variability in their desire for non-instrumental information to resolve uncertainty about future punishments [148]. Some human patients want to know in advance if they are likely to have disease, while others choose to avoid this information prior to possible symptoms [149, 150]. Relatedly, humans display variability in wanting to know whether they will encounter noxious stimuli and aversive or scary images [151–153]. But whether the same people that seek to resolve punishment uncertainty would do so in the case of reward uncertainty and vice versa has been unclear.

To study this in behaving animals, Jezzini et al. [23] trained monkeys to associate the presentation of visual objects with different probabilities of juice and air puff punishments (Fig. 6). Mirroring previous work [44, 53, 66], these initial objects were then followed by secondary cues that reduced the monkeys' uncertainty about the outcomes.

As in studies in human subjects, the monkeys displayed diverse choice preferences or attitudes towards punishment uncertainty resolution (Fig. 6C, D). Attitudes toward punishment and reward

information were not strictly linked to each other – monkeys had similar preferences for reward information but had distinct attitudes towards punishment information. Interestingly, distinct information attitudes were reflected in the ACC-47/12o network, albeit in distinct manners in each region. Both ACC and 47/12o contained single neurons whose activity anticipated the gain of information in a valence-specific-manner – anticipating information about either uncertain punishments or uncertain rewards (Fig. 7). But differentially, 47/12o also contained a subpopulation of neurons that integrated attitudes towards punishment and reward information to encode the overall preference in a bivalent manner. Also, neural activity in each area varied with monkeys' information attitudes on a trial-by-trial basis (Fig. 7-inset). Given these patterns of results, it is reasonable to assume that the ACC-47/12o network can guide information seeking in a flexible manner by integrating outcomes of distinct valences.

Bromberg-Martin and Hikosaka proposed that credit assignment and information seeking maybe intimately linked at the level of neural circuits [66]. The results of Jezzini et al. in combination with the body of work summarised in this review support this hypothesis and highlight the 47/12o as a potential prefrontal substrate. First, the 47/12o seems to possess the capacity to integrate valence-specific motivational signals from the ACC to reflect the total (or subjective) attitude towards information. Second, across the ACC-47/12o network there are overlapping and distinct neurons that signal information anticipation and surprise (after uncertainty is resolved) suggesting that information seeking and credit assignment are in part linked at the level of neural circuit within the ACC-47/12o network.

These functions could be supported by several modulatory inputs and through interactions with the basal ganglia [44]. For example, phasic dopamine activations reflect monkeys' preferences for advance information that will resolve their reward uncertainty and primary reward (or *juice*) prediction errors [66]. This signal may be crucial in teaching the prefrontal cortex to anticipate uncertainty resolution and drive information seeking through cortical-basal ganglia loops. Striatal regions that receive information anticipatory signals from cortex can be modulated by dopaminergic signalling of monkeys' information preferences, while other striatal regions that receive action and reward variables could be relatively more modulated by the canonical value-based predictions encoded by the same dopamine neurons.

SUMMARY AND OUTSTANDING QUESTIONS

Our current working hypothesis is that the ACC carries contextual and historical representations of predictions and internal states

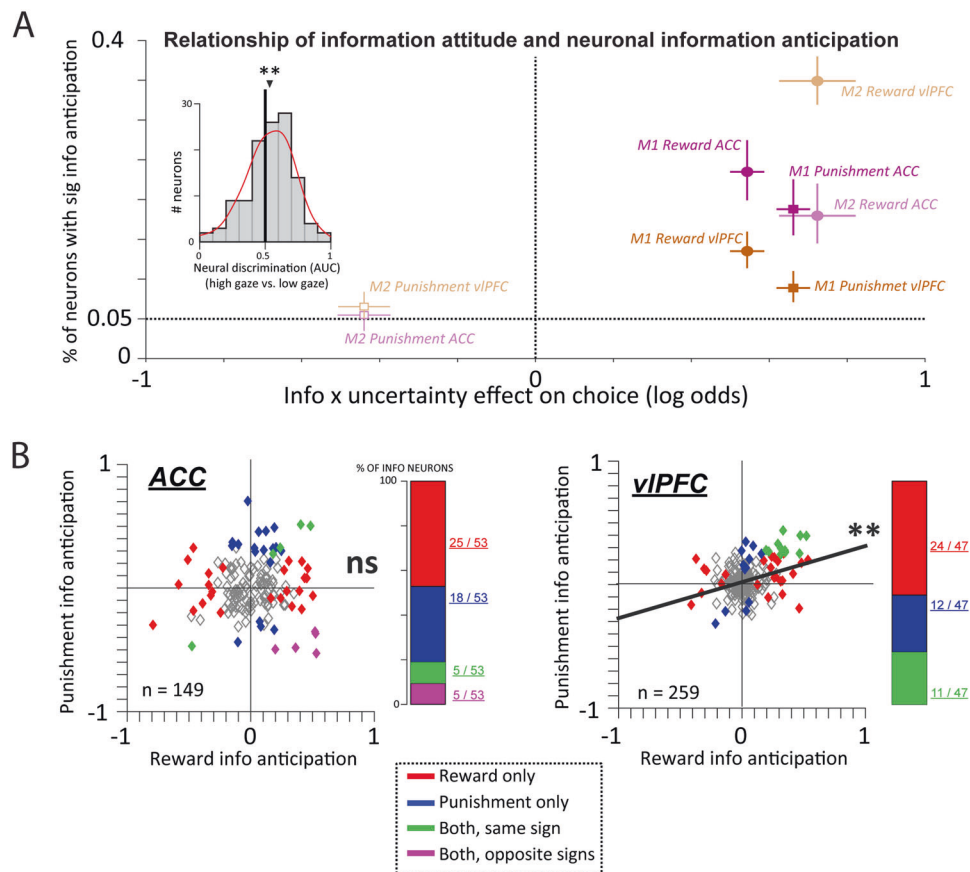


Fig. 7 Individual differences and areal differences between 47/12o and ACC in relation to information seeking. **A** A summary of the relationship of individual monkeys' (M1 and M2) information attitude and neural anticipation of reward and punishment uncertainty resolution in the ACC-47/12o network. The magnitude of choice preference for uncertainty resolution (x-axis) plotted against percentage of neurons with significant information anticipation (y-axis). Data are shown separately for each brain area, monkey, and valence. Inset: Neural discrimination between trials with strong versus weak anticipatory gaze to resolve punishment uncertainty. **B** Information anticipation for reward- (x-axis) and punishment (y-axis) uncertainty resolving information by single neurons (individual dots) in the ACC and 47/12o of Monkey 1 (Fig. 6). The monkey preferred to resolve reward and punishment uncertainty. Both, ACC and 47/12o had more neurons than would be expected by chance that anticipated reward- and punishment- uncertainty resolution, but only in 47/12o the information anticipation processes were linked across rewards and punishments, in a bivalent manner reflecting the subject's attitudes or preferences.

often closely tied (i) to the motivation to resolve uncertainty and (ii) to the assignment of accurate credit to the many events and actions in our lives. These high dimensional control [56] signals are utilised by 47/12o to assign credit to distinct events. However, understanding precisely how this might take place and how exactly information is exchanged between ACC and 47/12o as potential choices are identified, taken, and their outcomes monitored, is currently unknown. It will be important to have a better understanding of just what are the differences and similarities between ACC and 47/12o that facilitate their individual and joint functions. We know that ACC contains neurons that are well-situated to integrate feedback and other information over long time scales [62, 154] and while there is some evidence for similar neurons in the most anterior lateral prefrontal cortex [62] it is not clear whether they are found in 47/12o regions that are important for credit assignment. Moreover, while neuroimaging studies consistently identify the 47/12o in both humans and macaques, whether there are clear differences in neurophysiological activity patterns between 47/12o and either more lateral 47/12o areas or more medial OFC areas is unclear. Also, 47/12o may receive relatively more higher-order sensory inputs than the ACC [19, 21]. But whether it contains any sensory representations remains to be carefully tested. If so, such higher-order sensory representations could enhance its role in credit assignment through a weighted integration of surprising or uncertainty resolving events with visual objects.

Much of what we discussed has been in the realm of rewards – in terms of studies that concentrated on adaptive flexible behaviour in response to reward losses or gains. More work is required to assess the role of 47/12o and ACC under the threat of noxious or aversive punishments which may activate distinct effectors, internal states, and may be driven by distinct neuronal populations in the ACC [23, 44, 46].

A final point is that information anticipation may widely change the cognitive state of the subject to prepare them to learn from prediction errors and surprises [45, 155, 156]. To achieve a wide-ranging change in brain state, one possibility is that the ACC-47/12o network functions with neuromodulatory inputs from, for example, the basal forebrain [45, 100, 105, 157] to the cortical mantle that broadly signal uncertainty and surprise, and have a profound effect on global coordinated computations (i.e., states) [45, 158]. Again, whether this is the case will require careful investigation.

FUNDING AND DISCLOSURE

MFSR is funded by the Medical Research Council (MRC: MR/P024955/1) and Wellcome Trust (WT100973AIA; WT101092MA; 203139/Z/16/Z). IEM is funded by the National Institute of Mental Health under Award Numbers R01MH110594 and R01MH116937 to IEM, and by the McKnight Memory and Cognitive Disorders Award.

REFERENCES

1. Roberts AC. Primate orbitofrontal cortex and adaptive behaviour. *Trends Cogn Sci.* 2006;10:83–90.
2. Passingham RE, Wise SP. *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight.* Oxford: Oxford University Press; 2012.
3. Butters N, Pandya D. Retention of delayed alternation: effect of selective lesions of sulcus principalis. *Science.* 1969;165:1271–3.
4. Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci.* 2004;24:7540–8.
5. Murray EA, Wise SP. Interactions between orbital prefrontal cortex and amygdala: advanced cognition, learned responses and instinctive behaviors. *Curr Opin Neurobiol.* 2010;20:212–20.
6. Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature.* 1996;380:69–72.
7. Rygula R, Walker SC, Clarke HF, Robbins TW, Roberts AC. Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *J Neurosci.* 2010;30:14552–9.
8. Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci.* 2013;16:1140–5.
9. Neubert FX, Mars RB, Sallet J, Rushworth MF. 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.* 2015;112:E2695–704.
10. Rushworth MF, Buckley MJ, Gough PM, Alexander IH, Kyriazis D, McDonald KR, et al. Attentional selection and action selection in the ventral and orbital prefrontal cortex. *J Neurosci.* 2005;25:11628–36.
11. Rudebeck PH, Saunders RC, Lundgren DA, Murray EA. Specialized representations of value in the orbital and ventrolateral prefrontal cortex: desirability versus availability of outcomes. *Neuron.* 2017;95:1208–20. e5
12. Mackey S, Petrides M. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *Eur J Neurosci.* 2010;32:1940–50.
13. Carmichael ST, Price JL. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol.* 1994;346:366–402.
14. Walker EA. A cytoarchitectural study of the prefrontal area of the macaque monkey. *J Comp Neurol.* 1940;73:59–86.
15. Brodmann K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Localisation in the Cerebral Cortex ed. J. A. Barth, translated as Localisation in the Cerebral Cortex by LJ Garey (1994), London: Smith-Gordon; Leipzig; 1909.
16. Petrides M, Pandya DN. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci.* 2002;16:291–310.
17. Petrides M. Lateral prefrontal cortex: architectonic and functional organization. *Philos Trans R Soc Lond B Biol Sci.* 2005;360:781–95.
18. Carmichael ST, Price JL. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 1996;371:179–207.
19. Carmichael ST, Price JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 1995;363:642–64.
20. Webster MJ, Bachevalier J, Ungerleider LG. Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb Cortex.* 1994;4:470–83.
21. Kondo H, Saleem KS, Price JL. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J Comp Neurol.* 2005;493:479–509.
22. Barbas H. Anatomic organization of basoventral and mediadorsal visual recipient prefrontal regions in the rhesus monkey. *J Comp Neurol.* 1988;276:313–42.
23. Jezzini A, Bromberg-Martin ES, Trambaioli LR, Haber SN, Monosov IE. A prefrontal network integrates preferences for advance information about uncertain rewards and punishments. *Neuron.* 2021;S0896-6273:00353-6. <https://doi.org/10.1016/j.neuron.2021.05.013>.
24. Sallet J, Noonan MP, Thomas A, O'Reilly JX, Anderson J, Papageorgiou GK, et al. Behavioral flexibility is associated with changes in structure and function distributed across a frontal cortical network in macaques. *PLoS Biol.* 2020;18:e3000605.
25. Zatorre RJ, Fields RD, Johansen-Berg H. Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat Neurosci.* 2012;15:528–36.
26. Sampaio-Baptista C, Khrapitchev AA, Foxley S, Schlagheck T, Scholz J, Jbabdi S, et al. Motor skill learning induces changes in white matter microstructure and myelination. *J Neurosci.* 2013;33:19499–503.
27. Sampaio-Baptista C, Scholz J, Jenkinson M, Thomas AG, Filippini N, Smit G, et al. Gray matter volume is associated with rate of subsequent skill learning after a long term training intervention. *Neuroimage.* 2014;96:158–66.
28. Kaller MS, Lazari A, Blanco-Duque C, Sampaio-Baptista C, Johansen-Berg H. Myelin plasticity and behaviour-connecting the dots. *Curr Opin Neurobiol.* 2017;47:86–92.
29. Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, et al. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci.* 2008;28:13775–85.
30. Wittmann MK, Fouragnan E, Folloni D, Klein-Flugge MC, Chau BKH, Khamassi M, et al. Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. *Nat Commun.* 2020;11:3771.
31. Fouragnan EF, Chau BKH, Folloni D, Kolling N, Verhagen L, Klein-Flugge M, et al. The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. *Nat Neurosci.* 2019;22:797–808.
32. Kolling N, Behrens TE, Mars RB, Rushworth MF. Neural mechanisms of foraging. *Science.* 2012;336:95–8.
33. Rushworth MF, Kolling N, Sallet J, Mars RB. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol.* 2012;22:946–55.
34. Bartolo R, Averbeck BB. Prefrontal cortex predicts state switches during reversal learning. *Neuron.* 2020;106:1044–1054.e4.
35. Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB. The role of frontal cortical and medial-temporal lobe brain areas in learning a bayesian prior belief on reversals. *J Neurosci.* 2015;35:11751–60.
36. Gottlieb J, Snyder LH. Spatial and non-spatial functions of the parietal cortex. *Curr Opin Neurobiol.* 2010;20:731–40.
37. Gottlieb J, Oudeyer PY. Towards a neuroscience of active sampling and curiosity. *Nat Rev Neurosci.* 2018;19:758–70.
38. Mackintosh NJ. A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological Rev.* 1975;82:276–98.
39. Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev.* 1980;87:532–52.
40. Trudel N, Scholl J, Klein-Flugge MC, Fouragnan E, Tankelevitch L, Wittmann MK, et al. Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. *Nat Hum Behav.* 2021;5:83–98.
41. Wilson RC, Geana A, White JM, Ludvig EA, Cohen JD. Humans use directed and random exploration to solve the explore-exploit dilemma. *J Exp Psychol Gen.* 2014;143:2074–81.
42. Costa VD, Mitz AR, Averbeck BB. Subcortical substrates of explore-exploit decisions in primates. *Neuron.* 2019;103:533–45 e5.
43. Ghazizadeh A, Griggs W, Hikosaka O. Ecological origins of object salience: reward, uncertainty, aversiveness, and novelty. *Front Neurosci.* 2016;10:378.
44. White JK, Bromberg-Martin ES, Heilbronner SR, Zhang K, Pai J, Haber SN, et al. A neural network for information seeking. *Nat Commun.* 2019;10:1–19.
45. Monosov IE. How outcome uncertainty mediates attention, learning, and decision-making. *Trends Neurosci.* 2020;43:795–809.
46. Monosov IE. Anterior cingulate is a source of valence-specific information about value and uncertainty. *Nat Commun.* 2017;8:134.
47. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature.* 2009;459:837–41.
48. Hikosaka O, Yamamoto S, Yasuda M, Kim HF. Why skill matters. *Trends Cogn Sci.* 2013;17:434–41.
49. Gottlieb J, Hayhoe M, Hikosaka O, Rangel A. Attention, reward, and information seeking. *J Neurosci.* 2014;34:15497–504.
50. Ghazizadeh A, Griggs W, Hikosaka O. Object-finding skill created by repeated reward experience. *J Vis.* 2016;16:17.
51. Bichot NP, Heard MT, DeGennaro EM, Desimone R. A source for feature-based attention in the prefrontal cortex. *Neuron.* 2015;88:832–44.
52. Bromberg-Martin ES, Monosov IE. Neural circuitry of information seeking. *Curr Opin Behav Sci.* 2020;35:62–70.
53. Daddaoua N, Lopes M, Gottlieb J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci Rep.* 2016;6:20202.
54. Blanchard TC, Hayden BY, Bromberg-Martin ES. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron.* 2015;85:602–14.
55. Bromberg-Martin ES, Hikosaka O. Lateral habenula neurons signal errors in the prediction of reward information. *Nat Neurosci.* 2011;14:1209–16.
56. Monosov IE, Haber SN, Leuthardt EC, Jezzini A. Anterior cingulate cortex and the control of dynamic behavior in primates. *Curr Biol.* 2020;30:R1442–54.
57. Gottlieb J, Cohanpour M, Li Y, Singletary N, Zabeh E. Curiosity, information demand and attentional priority. *Curr Opin Behav Sci.* 2020;35:83–91.

58. Hunt LT, Malalasekera WN, de Berker AO, Miranda B, Farmer SF, Behrens TE, et al. Triple dissociation of attention and decision computations across prefrontal cortex. *Nat Neurosci*. 2018;21:1471–81.
59. Kobayashi K, Ravaioli S, Baranès A, Woodford M, Gottlieb J. Diverse motives for human curiosity. *Nat Human Behav*. 2019;3:587–95.
60. Preuschoff K, Bossaerts P, Quartz SR. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*. 2006;51:381–90.
61. Preuschoff K, Quartz SR, Bossaerts P. Human insula activation reflects risk prediction errors as well as risk. *J Neurosci*. 2008;28:2745–52.
62. Meder D, Kolling N, Verhagen L, Wittmann MK, Scholl J, Madsen KH, et al. Simultaneous representation of a spectrum of dynamically changing value estimates during decision making. *Nat Commun*. 2017;8:1942.
63. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007;10:1214–21.
64. Costa VD, Averbeck BB. Primate orbitofrontal cortex codes information relevant for managing explore-exploit tradeoffs. *J Neurosci*. 2020;40:2553–61.
65. White JK, Bromberg-Martin ES, Heilbronner SR, Zhang K, Pai J, Haber SN, et al. A neural network for information seeking. *Nat Commun*. 2019;10:5168.
66. Bromberg-Martin ES, Hikosaka O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*. 2009;63:119–26.
67. Real L, Caraco T. Risk and foraging in stochastic environments. *Annu Rev Ecol Syst*. 1986;17:381–90.
68. Caraco T. Energy budgets, risk and foraging preferences in dark-eyed juncos (*Junco hyemalis*). *Behav Ecol Sociobiol*. 1981;8:213–7.
69. Kolling N, Wittmann M, Rushworth MF. Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*. 2014;81:1190–202.
70. Chen X, Stuphorn V. Inactivation of medial frontal cortex changes risk preference. *Curr Biol*. 2018;28:3114–22. e4.
71. Genest W, Stauffer WR, Schultz W. Utility functions predict variance and skewness risk preferences in monkeys. *Proc Natl Acad Sci USA*. 2016;113:8402–7.
72. Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*. 2017;93:451–63.
73. Akaishi R, Kolling N, Brown JW, Rushworth M. Neural mechanisms of credit assignment in a multicue environment. *J Neurosci*. 2016;36:1096–112.
74. Walton ME, Devlin JT, Rushworth MFS. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat Neurosci*. 2004;7:1259–65.
75. Donoso M, Collins AG, Koechlin E. Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science*. 2014;344:1481–6.
76. Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, et al. Reinforcement learning in multidimensional environments relies on attention mechanisms. *J Neurosci*. 2015;35:8145–57.
77. Tervo DG, Proskurin M, Manakov M, Kabra M, Vollmer A, Branson K, et al. Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell*. 2014;159:21–32.
78. Karlsson MP, Tervo DG, Karpova AY. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science*. 2012;338:135–9.
79. O'Reilly JX, Schuffelgen U, Cuell SF, Behrens TE, Mars RB, Rushworth MF. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc Natl Acad Sci U S A*. 2013;110:E3660–9.
80. Stoll FM, Fontanier V, Procyk E. Specific frontal neural dynamics contribute to decisions to check. *Nat Commun*. 2016;7:11990.
81. Boorman ED, Rushworth MF, Behrens T. Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multialternative choice. *J Neurosci*. 2013;33:2242–53.
82. Monosov IE, Hikosaka O. Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *J Neurosci*. 2012;32:10318–30.
83. Ebitz RB, Albarran E, Moore T. Exploration disrupts choice-predictive signals and alters dynamics in prefrontal cortex. *Neuron*. 2018;97:450–61. e9.
84. Bromberg-Martin ES, Sharot T. The value of beliefs. *Neuron*. 2020;106:561–5.
85. Quilodran R, Rothe M, Procyk E. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*. 2008;57:314–25.
86. Sarafyazd M, Jazayeri M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science*. 2019;364:eaav8911.
87. Takada M, Tokuno H, Hamada I, Inase M, Ito Y, Imanishi M, et al. Organization of inputs from cingulate motor areas to basal ganglia in macaque monkey. *Eur J Neurosci*. 2001;14:1633–50.
88. Haber SN, Kim KS, Maily P, Calzavara R. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci*. 2006;26:8368–76.
89. Calzavara R, Maily P, Haber SN. Relationship between the corticostriatal terminals from areas 9 and 46, and those from area 8A, dorsal and rostral premotor cortex and area 24c: an anatomical substrate for cognition to action. *Eur J Neurosci*. 2007;26:2005–24.
90. Wittmann MK, Kolling N, Akaishi R, Chau BK, Brown JW, Nelissen N, et al. Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat Commun*. 2016;7:12327.
91. Cohen JY, Amoroso MW, Uchida N. Serotonergic neurons signal reward and punishment on multiple timescales. *Elife*. 2015;4:e06346.
92. Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci*. 2007;27:8366–77.
93. Bernacchia A, Seo H, Lee D, Wang XJ. A reservoir of time constants for memory traces in cortical neurons. *Nat Neurosci*. 2011;14:366–72.
94. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997;275:1593–9.
95. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*. 1996;16:1936–47.
96. Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci*. 2004;7:887–93.
97. Daw ND, Doya K. The computational neurobiology of learning and reward. *Curr Opin Neurobiol*. 2006;16:199–204.
98. Matsumoto M, Matsumoto K, Abe H, Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci*. 2007;10:647–56.
99. Niv Y, Schoenbaum G. Dialogues on prediction errors. *Trends Cogn Sci*. 2008;12:265–72.
100. Zhang K, Chen CD, Monosov IE. Novelty, salience, and surprise timing are signaled by neurons in the basal forebrain. *Curr Biol*. 2019;29:134–42. e3.
101. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HE, et al. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron*. 2013;80:519–30.
102. Matsumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*. 2007;447:1111–5.
103. Menegas W, Babayan BM, Uchida N, Watabe-Uchida M. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *eLife*. 2017;6:e21886.
104. Ledbetter NM, Chen CD, Monosov IE. Multiple mechanisms for processing reward uncertainty in the primate basal forebrain. *J Neurosci*. 2016;36:7852–64.
105. Hangya B, Ranade SP, Lorenc M, Kepecs A. Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell*. 2015;162:1155–68.
106. Lin SC, Nicolelis MA. Neuronal ensemble bursting in the basal forebrain encodes salience irrespective of valence. *Neuron*. 2008;59:138–49.
107. Avila I, Lin SC. Distinct neuronal populations in the basal forebrain encode motivational salience and movement. *Front Behav Neurosci*. 2014;8:421.
108. Raver SM, Lin S-C. Basal forebrain motivational salience signal enhances cortical processing and decision speed. *Front Behav Neurosci*. 2015;9:277.
109. Boubert S, Richmond BJ. Relation of locus coeruleus neurons in monkeys to Pavlovian and operant behaviors. *J Neurophysiol*. 2009;101:898–911.
110. Boubert S, Richmond BJ. Sensitivity of locus coeruleus neurons to reward value for goal-directed actions. *J Neurosci*. 2015;35:4005–14.
111. Rouhani N, Norman KA, Niv Y. Dissociable effects of surprising rewards on learning and memory. *J Exp Psychol Learn Mem Cogn*. 2018;44:1430–43.
112. Jang AI, Nassar MR, Dillon DG, Frank MJ. Positive reward prediction errors during decision-making strengthen memory encoding. *Nat Hum Behav*. 2019;3:719–32.
113. Baruni JK, Lau B, Salzman CD. Reward expectation differentially modulates attentional behavior and activity in visual area V4. *Nat Neurosci*. 2015;18:1656–63.
114. Brydson DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR. Attention for learning signals in anterior cingulate cortex. *J Neurosci*. 2011;31:18266–74.
115. Soltani A, Izquierdo A. Adaptive learning under expected and unexpected uncertainty. *Nat Rev Neurosci*. 2019;20:635–44.
116. Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci*. 2009;21:1162–78.
117. Kennerley SW, Behrens TE, Wallis JD. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci*. 2011;14:1581–9.
118. Li YS, Nassar MR, Kable JW, Gold JL. Individual neurons in the cingulate cortex encode action monitoring, not selection, during adaptive decision-making. *J Neurosci*. 2019;39:6668–83.
119. Kao CH, Lee S, Gold JL, Kable JW. Neural encoding of task-dependent errors during adaptive learning. *Elife*. 2020;9:e58809.
120. Kao CH, Khambhati AN, Bassett DS, Nassar MR, McGuire JT, Gold JL, et al. Functional brain network reconfiguration during learning in a dynamic environment. *Nat Commun*. 2020;11:1682.
121. Shima K, Tanji J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*. 1998;282:1335–38.

122. Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. *Nature Neurosci.* 2006;9:940–47.
123. Kolling N, Behrens T, Wittmann MK, Rushworth M. Multiple signals in anterior cingulate cortex. *Curr Opin Neurobiol.* 2016;37:36–43.
124. Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MFS. Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci.* 2016;19:1280–5.
125. Ullsperger M, Fischer AG, Nigbur R, Endrass T. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci.* 2014;18:259–67.
126. Kawai T, Yamada H, Sato N, Takada M, Matsumoto M. Roles of the lateral habenula and anterior cingulate cortex in negative outcome monitoring and behavioral adjustment in nonhuman primates. *Neuron.* 2015;88:792–804.
127. Seymour B, Singer T, Dolan R. The neurobiology of punishment. *Nat Rev Neurosci.* 2007;8:300–11.
128. Hayden BY, Heilbronner SR, Pearson JM, Platt ML. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci.* 2011;31:4178–87.
129. Chau BK, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, et al. Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron.* 2015;87:1106–18.
130. Saleem KS, Kondo H, Price JL. Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *J Comp Neurol.* 2008;506:659–93.
131. Carmichael ST, Price JL. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol.* 1995;363:615–41.
132. Ferry AT, Ongur D, An X, Price JL. Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. *J Comp Neurol.* 2000;425:447–70.
133. Ongur D, An X, Price JL. Orbital and medial prefrontal cortical projections to the hypothalamus in macaque monkeys. *J Comp Neurol.* 1998;401:480–505.
134. Amaral DG, Price CJ, Pitkanen A, Carmichael ST. Anatomical organization of the primate amygdaloid complex. In: Aggleton JP, editor *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*. New York: Wiley-Liss; 1992. p. 1–66.
135. Ghashghaei HT, Barbas H. Neural interaction between the basal forebrain and functionally distinct prefrontal cortices in the rhesus monkey. *Neuroscience.* 2001;103:593–614.
136. Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron.* 2010;65:927–39.
137. Thorndike EL. *Animal Intelligence: Experimental Studies*. New York: Macmillan; 1911.
138. Thorndike EL. A proof of the law of effect. *Science.* 1933;77:173–75.
139. Folloni D, Fouragnan E, Wittmann MK, Roumazeilles L, Tankelevitch L, Verhagen L, et al. Ultrasound modulation of macaque prefrontal cortex selectively alters credit assignment-related activity and behavior submitted.
140. Folloni D, Verhagen L, Mars RB, Fouragnan E, Constans C, Aubry JF, et al. Manipulation of Subcortical and Deep Cortical Activity in the Primate Brain Using Transcranial Focused Ultrasound Stimulation. *Neuron.* 2019;101:1109–16 e5.
141. Khalighinejad N, Bongioanni A, Verhagen L, Folloni D, Attali D, Aubry JF, et al. A basal forebrain-cingulate circuit in macaques decides it is time to act. *Neuron.* 2020;105:370–84 e8.
142. Verhagen L, Gallea C, Folloni D, Constans C, Jensen DE, Ahnine H, et al. Offline impact of transcranial focused ultrasound on cortical activation in primates. *Elife.* 2019;8:e40541.
143. Deffieux T, Younan Y, Wattiez N, Tanter M, Pouget P, Aubry JF. Low-intensity focused ultrasound modulates monkey visuomotor behavior. *Curr Biol.* 2013;23:2430–3.
144. Kubanek J, Brown JW, Ye P, Butts Pauly K, Moore T, Newsome W. Remote, brain region-specific control of choice behavior with ultrasonic waves. *Sci Adv.* 2020;6:eaa4193.
145. Jocham G, Brodersen KH, Constantinescu AO, Kahn MC, Ianni AM, Walton ME, et al. Reward-guided learning with and without causal attribution. *Neuron.* 2016;90:177–90.
146. Klein-Flugge MC, Wittmann MK, Shpektor A, Jensen DEA, Rushworth MFS. Multiple associative structures created by reinforcement and incidental statistical learning mechanisms. *Nat Commun.* 2019;10:4835.
147. Barberini CL, Morrison SE, Saez A, Lau B, Salzman CD. Complexity and competition in appetitive and aversive neural circuits. *Front Neurosci.* 2012;6:170.
148. Miller SM. Monitoring and blunting - validation of a questionnaire to assess styles of information seeking under threat. *J Personal Soc Psychol.* 1987;52:345–53.
149. Miller SM. Monitoring versus blunting styles of coping with cancer influence the information patients want and need about their disease. *Implic Cancer Screen Manag Cancer.* 1995;76:167–77.
150. Lerman C, Hughes C, Lemon SJ, Main D, Snyder C, Durham C, et al. What you don't know can hurt you: adverse psychologic effects in members of BRCA1-linked and BRCA2-linked families who decline genetic testing. *J Clin Oncol.* 1998;16:1650–4.
151. Oosterwijk S, Snoek L, Tekoppele J, Engelbert LH, Scholte HS. Choosing to view morbid information involves reward circuitry. *Sci Rep.* 2020;10:1–13.
152. Oosterwijk S. Choosing the negative: a behavioral demonstration of morbid curiosity. *PLoS one.* 2017;12:e0178399.
153. Niehoff E, Oosterwijk S. To know, to feel, to share? Exploring the motives that drive curiosity for negative content. *Curr Opin Behav Sci.* 2020;35:56–61.
154. Spitman M, Seo H, Lee D, Soltani A. Multiple timescales of neural dynamics and integration of task-relevant signals across cortex. *Proc Natl Acad Sci USA.* 2020;117:22522–31.
155. Farashahi S, Donahue CH, Khorsand P, Seo H, Lee D, Soltani A. Metaplasticity as a neural substrate for adaptive learning and choice under uncertainty. *Neuron.* 2017;94:401–14. e6.
156. Khorsand P, Soltani A. Optimal structure of metaplasticity for adaptive learning. *PLoS Comput Biol.* 2017;13:e1005630.
157. Monosov IE, Leopold DA, Hikosaka O. Neurons in the primate medial basal forebrain signal combined information about reward uncertainty, value, and punishment anticipation. *J Neurosci.* 2015;35:7443–59.
158. Turchi J, Chang C, Frank QY, Russ BE, David KY, Cortes CR, et al. The basal forebrain regulates global resting-state fMRI fluctuations. *Neuron.* 2018;97:940–52. e4.

ACKNOWLEDGEMENTS

Both authors express their gratitude to all past and current members of their laboratories.

AUTHOR CONTRIBUTIONS

Both authors prepared the manuscript together.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to I.E.M. or M.F.S.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.