nature reviews neuroscience

https://doi.org/10.1038/s41583-022-00670-w

Review article

Check for updates

Neurophysiological mechanisms of error monitoring in human and non-human primates

Zhongzheng Fu [1,2], Amirsaman Sajad [3,4,5], Steven P. Errington [3,4,5], Jeffrey D. Schall [3,5,6,7,8] & Ueli Rutishauser [1,9,10]

Abstract

Performance monitoring is an important executive function that allows us to gain insight into our own behaviour. This remarkable ability relies on the frontal cortex, and its impairment is an aspect of many psychiatric diseases. In recent years, recordings from the macaque and human medial frontal cortex have offered a detailed understanding of the neurophysiological substrate that underlies performance monitoring. Here we review the discovery of single-neuron correlates of error monitoring, a key aspect of performance monitoring, in both species. These neurons are the generators of the error-related negativity, which is a non-invasive biomarker that indexes error detection. We evaluate a set of tasks that allows the synergistic elucidation of the mechanisms of cognitive control across the two species, consider differences in brain anatomy and testing conditions across species, and describe the clinical relevance of these findings for understanding psychopathology. Last, we integrate the body of experimental facts into a theoretical framework that offers a new perspective on how error signals are computed in both species and makes novel, testable predictions.

Sections

Introduction
Medial frontal cortex
The ERN
Intracranial recordings across species
Anatomical differences across species
Tasks for performance monitoring
Proposed conceptual model
Internal models in the MFC
Implementing action error computations
Predictions and suggested experiments
Relationship with existing models
Relevance to psychiatric disorders
Conclusions

[1]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA. [2]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, USA. [3]Center for Integrative & Cognitive Neuroscience, Vanderbilt University, Nashville, TN, USA. [4]Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN, USA. [5]Department of Psychology, Vanderbilt University, Nashville, TN, USA. [6]Centre for Vision Research, York University, Toronto, Ontario, Canada. [7]Vision: Science to Applications (VISTA), York University, Toronto, Ontario, Canada. [8]Department of Biology, Faculty of Science, York University, Toronto, Ontario, Canada. [9]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [10]Center for Neural Science and Medicine, Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA. e-mail: zzbrooksfu@gmail.com; schalljd@yorku.edu; rutishauseru@csmc.edu

Introduction

To survive, all sentient species must adapt to errors[1]. We can, for example, notice that we pressed a wrong keyboard key without external feedback[2], or quickly realize that we have made the wrong turn on our way home or have called an acquaintance by the wrong name. The ability to recognize action errors is essential for learning new tasks, adapting to challenging or changing conditions, and interrupting useless or dangerous courses of action[3]. Error monitoring is the cognitive process by which we rapidly detect situations in which performance deviates from the intended goal. Error monitoring is an element of performance monitoring, which is a critical aspect of cognitive control whereby we achieve our goals[4]. What are the neural processes that endow us with the ability to detect whether we have made an error and thereby to gain insight into our own behaviour?

It has long been recognized that performance monitoring is critical for goal-directed behaviour and cognitive control more broadly[4]. In addition, it is increasingly being appreciated that malfunctioning performance monitoring is a key symptom of some psychiatric disorders, including impulsive behaviour, obsessive–compulsive disorder, addiction and schizophrenia[5,6]. Hence, performance monitoring is a central construct in the Research Domain Criteria framework of the US National Institute of Mental Health[7,8] and a core element in new computational psychiatry-based approaches to mental health[9]. As a result, deciphering the neural mechanisms that underlie performance monitoring has become a major interest in both cognitive neuroscience and clinical neuroscience.

The validity of animal model paradigms for higher-level human cognitive processes — in particular, executive functions such as performance monitoring — remains uncertain[10,11]. Indeed, although much has been learned about the neural mechanisms of cognitive control in animal models (particularly in the macaque), little is known about the relevance of these findings for understanding human cognition and its disruption in psychiatric disorders. Intracranial recordings in humans provide a rare opportunity to directly compare findings in animal models with those in humans. However, such comparisons are possible only if the animal model system closely mirrors human anatomy, physiology and behaviour.

The aim of this Review is to compare the neuronal mechanisms for self-monitoring of action errors across macaques and humans. Recent work involving similar tasks and assessment of neural activity at the single-neuron level in the medial frontal cortex (MFC) has demonstrated that there are many similarities in error-monitoring processes in humans[12] and macaques[13]. At the same time, experiments possible only in humans reveal the relevance of these processes for cognitive control involved in language, cognitive flexibility and neuropsychiatric disorders. These studies provide converging evidence that error neurons in the MFC constitute single-neuron correlates of the error-related negativity (ERN), an important event-related potential that has spawned a large literature in cognitive neuroscience (see later). Informed by these new findings, we synthesize a model for how errors are computed and propose a set of tasks that are most suitable for multispecies investigation. We conclude by proposing that the neural processes in the MFC can be studied synergistically by combining work in humans and macaques to reveal new insights into the neural mechanisms of cognitive control and its impairment by disease.

Medial frontal cortex

Converging evidence from neuroimaging, electrophysiological recordings and lesion studies in humans and non-human primates indicates that the MFC is essential for performance monitoring[4,12–15]. Two medial frontal regions in particular have been intensely studied (Fig. 1): the first is the supplementary motor complex, which is located within the superior frontal gyrus (SFG), anterior to the motor cortex. This complex includes the supplementary motor area (SMA), the pre-supplementary motor area (pre-SMA)[16] and the supplementary eye field (SEF)[13,17]. The second is a collection of areas in both banks of the cingulate sulcus and the cingulate gyrus. Habitual descriptions of the 'anterior cingulate cortex' (ACC), including ours[12,18,19], can be replaced by more precise nomenclature reflecting more refined anatomical and comparative studies, which distinguish the middle cingulate cortex (MCC) ventral to the supplementary motor cortex from the ACC around the genu of the corpus callosum[20–22]. The rostral ACC contributes more to emotion and autonomic regulation, whereas the MCC contributes more to performance monitoring and cognitive control[23], and includes the cingulate motor areas that project to the spinal cord[24].

One of the most robust findings in the MFC in humans and macaques is its vigorous single-neuron spiking response following erroneous actions[12,13,19,25–28]. This signal is also detectable with non-invasive methods such as functional MRI (fMRI)[4] and simultaneous fMRI and scalp electroencephalography (EEG)[29,30] in humans. Error-monitoring signals in both the MCC and the SFG are often but not always associated with performance adjustments, such as post-error slowing[12,13]. These findings bridge monitoring processes and the engagement of subsequent cognitive control. Together with evidence from lesion and electrical stimulation studies[31–33], this body of research indicates that recording from and manipulating the MFC provides a powerful paradigm to probe the neural substrate of error monitoring.

Invasive single-neuron recordings in the MFC can be obtained in both macaques and humans. In macaques, recordings in the MCC, SEF and SMA, and pre-SMA are routinely performed with microelectrodes or silicon probes (Fig. 1a). In humans, recordings in the MFC have been performed in two clinical scenarios (Fig. 1a): first, in patients with epilepsy, areas along the medial wall of the frontal lobe (including the MCC, SMA and pre-SMA) are targeted with depth electrodes to localize focal seizure onset zones or seizure spread patterns[34,35]; second, in patients undergoing awake brain surgery for implantation of a deep brain stimulator or targeted resection, the MFC is targeted using microelectrodes[36,37]. These neurosurgical scenarios provide opportunities to study the contribution of the MFC to performance monitoring in both macaques and humans with very similar experimental techniques. The aspects of performance monitoring best suited for study in each species differ owing to behavioural and technical constraints, with some aspects approachable only in humans and others approachable only in macaques (Table 1). We note that although our focus here is on the role of the SFG and MCC in error monitoring, these brain areas contribute importantly to other cognitive functions — such as signalling response conflict[14,17,38–40] — that are beyond the scope of this Review.

The ERN

The MFC has long been thought to be the source of the ERN (also known as the Ne), which is a brief period of greater negative polarization in the electroencephalogram over the MFC when an error is made relative to correct trial performance[41–46] (Figs. 1b and 2). Although error monitoring has also been investigated with fMRI, the high temporal resolution of EEG offers unique leverage in the investigation of cognitive control processes. Furthermore, the ERN can be measured with a single scalp electrode, thereby providing high translational potential. Consequently, understanding the neural processes that give rise to

the ERN has relevance for studying normal human behaviour and as a potential endophenotype for psychiatric disease.

Although the ERN has become a major workhorse in cognitive neuroscience, until recently little was known about the underlying mechanisms that give rise to it. The ERN is different from sensory-evoked potentials or movement-related potentials that are associated with observable events. Rather, the ERN indexes an internal state that can be inferred only indirectly. Therefore, the improved understanding of the ERN that is now emerging (discussed later) can also provide insights into other event-related potentials, such as the contingent negative variation and the readiness potential. The ERN is closely related to the feedback-related negativity (FRN), which is an event-related potential evoked by externally signalled failures (for example, after a correct response receiving less-than-expected reward or sensory feedback indicating an error). The two signals have overlapping scalp voltage distributions[47], which has led to the hypothesis that both index the computation of prediction errors[48]. We discuss the validity of this claim further later. We refer to the terms 'ERN' and 'FRN' whenever the EEG data are analysed time-locked to the erroneous response onset or feedback onset, respectively. We note that in some studies the term 'feedback ERN' is used instead of 'FRN' (for example, see ref. [49]), which we avoid for clarity.

Understanding how the ERN is generated begins with knowing where it arises. In 1994, using source modelling, researchers showed that a single dipole in the MCC could account for the spatial distribution of the ERN[50]. Confidence in this conclusion cannot be high, though, because the inverse problem of locating dipoles given a spatial distribution of voltages has no unique solution (Fig. 1b). More uncertainty arises because the simplified geometric models of the head that were used in the 1990s were poor approximations of the human brain and head. Another uncertainty overlooked by all early and most current studies arises from variation in sulcal morphology across individuals[51], which necessarily changes the orientation of dipoles (Fig. 1b). Of most relevance for the ERN is the absence or presence of a paracingulate sulcus (PCS), which is located superior to the cingulate sulcus in ~70% of humans[52,53]. Such morphological differences in cortical folding patterns produce differences in EEG voltage distributions; for example, a weaker, briefer N400 versus a stronger, longer N400 when response conflict is detected[54].

Although the inverse problem of source localization of current locations from voltages is ill-posed and offers indefinite results, forward modelling from current locations to voltage distributions offers definite results. Recent work using this approach has demonstrated that a spatial pattern of voltage resembling the ERN can be produced by a simulated dipole in the SEF of macaques[55]. The biophysics underlying the relationships between neuronal spiking, synaptic potentials and EEG signals has many complexities. For example, the relationship between current dipoles in the cerebral cortex and surface EEG signals has most commonly been described only in terms of instantaneous electric fields. However, research has demonstrated that the slower diffusion of ions, possibly mediated by glia[56], can interact with current dipoles in producing the EEG signal. Relative to granular sensory areas, agranular cortical areas have a higher ratio of glial cells to neurons[57], so the contribution of slower ion diffusion to the EEG signal can differ between granular and agranular areas. To our knowledge this has not been investigated.

If both the SFG and the MCC separately signal errors, then the ERN can have at least two sources[12]. Moreover, if error-related signals are produced in different areas within the MCC, then the number of sources increases further. The voltage measured with an electrode on the head will be the superposition of the voltages produced by current dipoles in each cortical area. The specific character of that superposition depends on the geometry of the cortex, which specifies dipole location, orientation and distance relative to surface electrodes. Voltages from two sources can sum or cancel depending on the orientation of the dipoles. This means that the ERN is a manifestation of multiple neural signals arising in different cortical areas[58] (Fig. 1b). Given the similar anatomical location of the pre-SMA in humans and macaques, its biophysical contribution to the ERN sampled at the midline with EEG electrodes is expected to be similar in both species (Fig. 1b). However, we expect that dipoles established in the SEF of macaques will contribute more to the ERN than those in the SEF of humans (Fig. 1b). Previous comparisons of human and macaque MFC functions[59] have not incorporated these details of cortical structure.

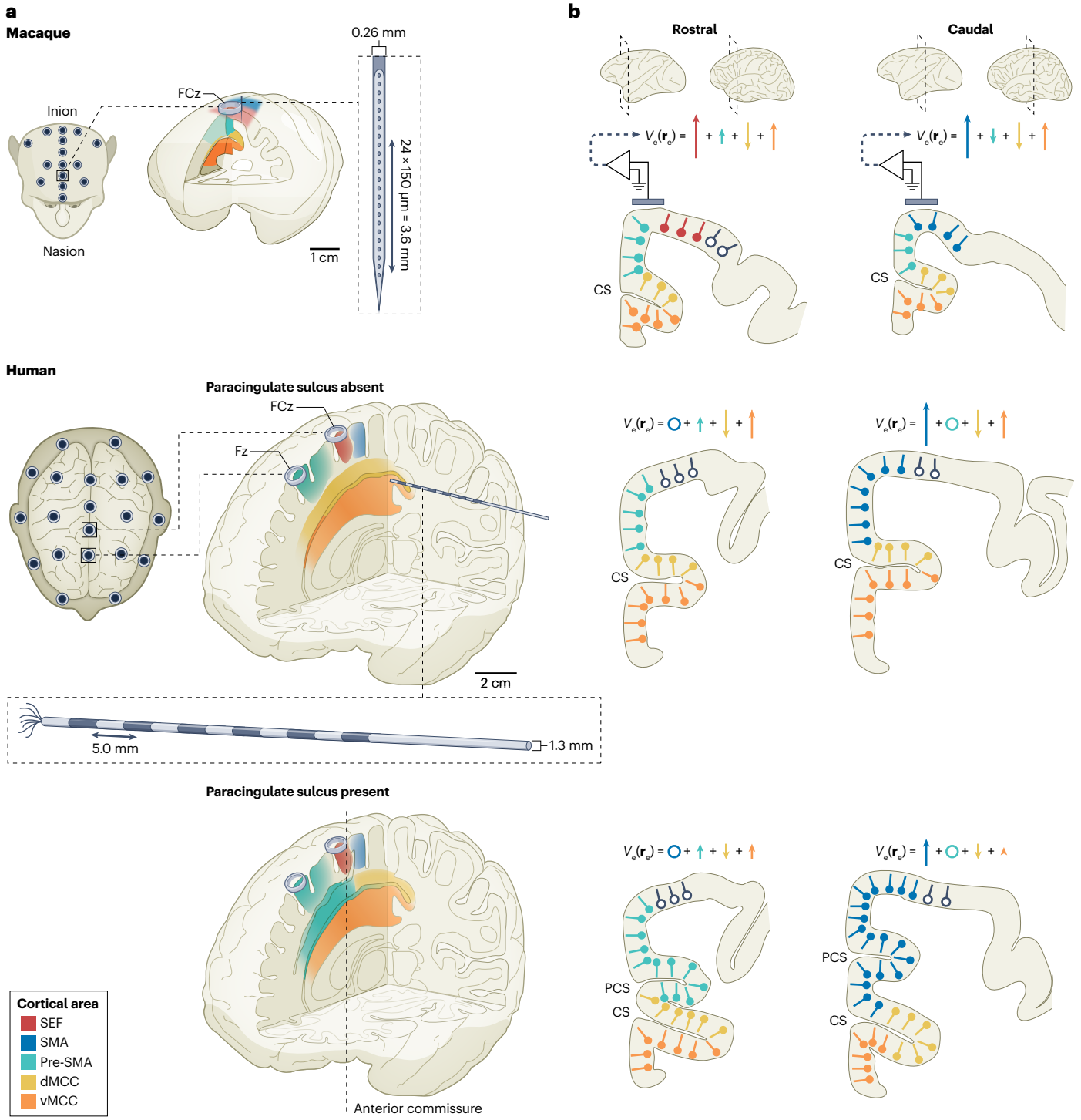
Spectral decomposition of the EEG signal has provided additional insights into the processes generating the ERN. For example, changes in theta-band and delta-band oscillations along the frontal midline coincide with the ERN[60–62], with both phase resetting and changes of power in ongoing theta oscillations contributing to this signal[63,64]. More broadly, theta-band activity in the MFC has been described as a mesoscale neural correlate of cognitive control[65]. Although they provide powerful insights, these data have not revealed the specific contributions of different parts of the MFC to the ERN. They also do not provide a cellular-level explanation of how the changes in power and phase of theta-band EEG arise following the commission of errors.

In humans, detection of errors sometimes evokes error awareness, which is a metacognitive reflection that enables conscious reasoning about errors. The relationship between error awareness and the ERN, and the positive polarization that follows the ERN (known as the Pe component), is complex and remains an active area of research[66,67]. For our purposes it suffices to note that although humans are typically aware of the kind of errors that are accompanied by an ERN in standard tasks, this does not appear to be necessary; an ERN can still be present for unaware errors[68]. We therefore posit that the mechanisms of error detection we describe are not conditional on awareness.

Intracranial recordings across species

A key goal of this Review is to contrast findings across species (Figs. 2 and 3). Single neurons that signal errors have been found in the SEF[13], pre-SMA, SMA[16] and MCC[19] in non-human primates and in the pre-SMA and MCC in humans[12] (Fig. 2a,b). The temporal properties of these neurons are similar across species (Fig. 3c,d): the spike rate is maximally modulated ~100 ms after an erroneous action (saccade, arm movement or finger button press). Two features determine that the error signal reported in these studies is not exclusively of sensory origin. First, the timing of error signals is synchronized with the muscle contraction rather than any sensory events. Second, the error signal arises before the trial outcome is signalled by external feedback about accuracy or the gain or loss of a reward. Also, the magnitude of single-neuron error signals depends weakly, if at all, on the laterality of the motor effector[19,69]. It is unknown whether the error-related modulation of single neurons is invariant across motor effectors. One investigation of the topography of the ERN observed after errors committed with the hand or foot indicated a common source[70], but a more recent study of errors committed with the eyes or the hand and using more sophisticated modelling of current sources described different sources[71].

The scalp-recorded ERN reliably occurs ~100 ms after a self-monitored error in macaques and humans (Fig. 2c,d), despite



differences across individuals, task requirements and effectors (eye, hand, foot or vocal)[46,70–72]. The intracranial ERN (iERN) also has a similar peak onset latency between humans and macaques: ~100 ms in the macaque SEF[25] and ~130 ms in the macaque MCC after an oculomotor response[26]; ~100 ms in the human pre-SMA and ~130 ms in the human MCC after a button press using a finger[12]. In both species, the iERN and the scalp ERN can be estimated reliably on single trials, with the iERN having the highest reliability (Fig. 3a,b). Two separate human studies have found, on a trial-by-trial basis, that the latency and the amplitude of the iERN in the MCC are correlated with those of the iERN in the pre-SMA[12] or the SMA[35]. In addition, the MCC iERN occurs only when there is a preceding SMA iERN in sessions where both areas are recorded

Fig. 1 | Locations of medial frontal lobe areas implicated in performance monitoring and typical recording approaches. a, Cross sections of the macaque (top) and human (middle and bottom) brains highlighting the relative locations of the pre-supplementary motor area (pre-SMA; green), supplementary motor area (SMA; blue), supplementary eye field (SEF; red), dorsal middle cingulate cortex (dMCC; yellow) and ventral middle cingulate cortex (vMCC; orange). Human brains are illustrated with the paracingulate sulcus absent (middle) or present (bottom). For each species, the approach used for neurophysiological sampling is illustrated. In macaques, intracortical neural signals are sampled with multicontact linear electrode arrays inserted nearly vertically, perpendicular to the cortical layers. In humans, intracortical neural signals are sampled with microwire electrodes. In both species, electroencephalography signals can be recorded from electrodes placed on the cranium (grey cylinder). **b**, Biophysical contributions of areas in the medial frontal cortex to the error-related negativity. Rostral (left) and caudal (right) sections are presented to highlight variation in contributions from the different cortical areas. Putative dipoles are distinguished by colour for the SEF (red), pre-SMA (green), dMCC (yellow) and vMCC (orange). Dipole orientation varies with cortical folding. The hypothesized vectorial contributions of dipoles in each area ($\mathbf{r}_e$) to medial frontal electroencephalography voltage (V_e) are portrayed through arrows indicating polarity and strength. CS, cingulate sulcus; Fz, frontal midline EEG electrode; FZc, frontal central EEG electrode; PCS, paracingulate sulcus.

simultaneously[35]. This is yet to be confirmed in macaques. These results strongly suggest a hierarchical relationship between the MCC and the SFG in error processing across species (Fig. 3c,d), with action errors being first detected in the SFG and then communicated to the MCC for updating control forward models (see later)[12,35].

In both macaques and humans, two types of error-related neurons have been identified. In the human MFC, 'type I' neurons produce more spikes on average on trials in which an error was made ('error trials') than on trials with a correct response ('correct trials'), whereas 'type II' neurons do the opposite. The proportion of type II error neurons is notably larger in the pre-SMA (40%) than in the MCC (26%)[12]. A similar distinction has been made in the macaque SFG, with neurons responding more to error trials than to correct trials being referred to as 'error cells' and those with reduced activity in error trials being referred to as 'reward expectation cells'[16]. Note that in the human study, no reward but delayed visual feedback on accuracy was provided, suggesting that type II responses cannot be solely attributed to reward expectation.

Given that the latency of error responses in MFC neurons falls in the same range as that of the scalp ERN and iERN, a key unresolved question is how these two signals are related. This question is of particular importance because little is known about how the ERN is generated even though many computational aspects of this signal are well known[42]. Our macaque and human data[12,13] provide insight into the relation between simultaneously recorded scalp ERN in macaques or iERN in humans and the activity of error neurons. In macaques, simultaneous recording of the ERN and the spiking activity of neurons across all layers of the SEF demonstrates clear associations between the microlevel and macrolevel error signals[13]. Notably, error neurons in layer 2/3 but not those in layer 5/6 demonstrate this relationship. In humans, the firing rate of error neurons is predicted by the amplitude of the iERN in the pre-SMA and the MCC[12]. In both species, this relationship exists exclusively for error neurons, suggesting that it is not merely a generic biophysical one but is specific to error computation. Laminar recordings in the human MCC[73] and in the macaque SEF[13] demonstrate prominent current sinks in layers 2 and 3 during errors, where pyramidal neurons in lower layer 3 and layer 5 extend their dendrites.

One hypothesis for these cross-species findings is that the ERN reflects the highly synchronous and aligned electric dipoles generated within error neurons[48]. In our view, these dipoles are generated by synaptic inputs from the thalamocortical projections and passive return currents at apical locations, as well as inhibitory actions by interneurons and projections from top-down regions. The larger the ERN amplitude at a given location is, the stronger and more synchronous these synaptic inputs may be, and the more effectively these synaptic inputs can drive the firing of error neurons. These results provide the most direct evidence to date that the ERN is jointly generated in the SFG and the MCC in primates (Fig. 1b), thereby providing a solid physiological foundation for the many EEG studies that use this signal to study cognitive control.

A second signal of relevance for error monitoring is response conflict[74,75]. We differentiate between two types of response conflict-related signals: those occurring during action selection while multiple response options are being considered (ex ante), and those occurring after a response has been made (ex post). Whereas the former calls for online control to resolve conflict proactively, the latter is an 'after the fact' evaluative signal. In humans, neurons signalling response conflict ex ante exist in both the MCC and the pre-SMA[18,36,76]. In macaques, no ex ante conflict neurons have been found using the stop-signal task or related conflict tasks, which involve conflict between concurrent go and stop/distractor processes (see below)[19,69,77–79]. In a task involving visually guided saccades in the presence of a salient distractor, some have interpreted neural spiking in the macaque MCC as ex ante response conflict between the goal-compatible and distractor-driven saccade plans[80], but this conclusion is debatable[81]. The existence of ex ante conflict signals may depend on the task used to probe it. Whether neurons in the human MFC signal ex ante conflict in stop-signal tasks remains an open question. By contrast, neurons signalling ex post conflict exist commonly in both species[18,69,77,78]. In humans, a subset of neurons in the pre-SMA and the MCC differentiate between whether a correctly performed action was made during high or low conflict, whereas in macaques this signal arises after successfully cancelling a saccade or an arm movement[69,77,78]. In humans, ex ante and ex post conflicts are signalled by different neurons, and multivariate population firing rate patterns do not generalize across these two periods, suggesting that they represent different types of conflict signals[18]. Although it may seem puzzling why conflict would be signalled ex post, this signal plays a critical role in the model we develop later herein.

Last, we consider how the findings discussed differ from those derived from fMRI and/or scalp EEG. Neither spectral analysis nor fMRI blood oxygenation level-dependent (BOLD)-based analysis has been able to identify that ex ante and ex post conflict signals are distinct because both involve temporal smoothing. Therefore, studies on the relationship between BOLD-based conflict signals and subsequent behavioural adjustments may need to be revised to clarify whether the effect is due to ex ante or ex post conflict signals. Similarly, the insight that errors and ex ante conflict are represented by distinct groups of neurons is uniquely provided by single-neuron studies[12].

Anatomical differences across species

The foregoing comparison between humans and macaques indicates that the error-monitoring system is evolutionarily conserved across primates. At the same time, there are notable differences in

Table 1 | Technical considerations linked to and cognitive processes examined in performance monitoring in humans and macaques

Item	Humans	Macaques
Technical issues		
Acquisition of task rules	Verbal instruction	Non-verbal conditioning
Number of participants	At least 20	Typically 2
Number of sessions or number of trials per individual	Low	High
Motivation	Social reinforcement	Primary reinforcement
Brain geometry	Large, highly folded, idiosyncratic cortex	Small, modestly folded, stereotyped cortex
EEG electrode placement	On scalp	On scalp; on skull beneath scalp and muscles
Typical microelectrode recordings	Brush wire with no layer specificity	Linear electrode arrays allowing layer specificity; square electrode arrays allowing areal sample
Possible causal interventions	Electrical stimulation	Electrical stimulation, pharmacology, lesions, optogenetics
Potential impact of neurological disease and/or medication	Yes	No
Effectors typically studied	Button press with finger	Eye movements, arm movements
Study of individual differences (large samples)	Yes	No
ERP–single neuron correlates	Yes	Yes
Study impact of comorbid neurological disease	Yes	No
Cognitive processes examined		
Outcome registration (reward, points)	Yes	Yes
Error monitoring	Yes	Yes
Expertise level	Naive	High
Task switching or multitask comparisons (domain generality)	Yes	No
Response inhibition or conflict monitoring (slowing)	Yes	Yes
Interference or conflict due to reading	Yes	No
Test of awareness of performance	Yes	No
Time estimation	Yes	Yes
Aversive outcomes	No	Yes

All factors are considered from the point of view of invasive medial frontal cortex recordings. 'Yes' and 'no' entries indicate topics that, to date, have or have not been studied routinely in the given species, respectively. EEG, electroencephalography; ERP, event-related potential.

the mechanisms underlying such monitoring between humans and macaques. Investigation of these differences can reveal new insights into the evolution of cognitive functions.

The human MFC and the macaque MFC exhibit many similar features (Fig. 1). The agranular areas seem largely homologous in organization and function but differ in their location on the cortical surface[82,83]. In both humans and macaques, the SMA is located immediately rostral to the primary motor cortex, and the pre-SMA is located rostral to the SMA (Fig. 1). In both species, the pre-SMA is found on the medial surface. However, whereas much of the SMA and all of the SEF are located on the dorsal convexity in macaques[84], in humans, they are located on the medial wall, with the SEF centred at the paracentral sulcus[85].

The folding pattern of sulci in the human MFC is considerably more variable than it is in the macaque MFC. A notable species difference is the PCS[86]. The PCS is unique to humans and other apes[52,87] but is not present in macaques[82]. As noted earlier herein, although all humans have a cingulate sulcus, ~70% of humans also have a PCS (most commonly in the left hemisphere but not the right hemisphere)[52,53]. The PCS is also more prominent in males[51]. To our knowledge, no invasive studies have specifically compared the neural activity within versus outside the paracingulate gyrus during performance-monitoring tasks. The PCS is important for understanding performance monitoring for the following reasons. First, its presence dictates where specific neural signals are located[86], thereby influencing what the ERN measures (Fig. 1b). Second, the presence and morphology of the PCS are correlated with interindividual differences in metacognitive abilities such as reality monitoring[88,89] or the ability to perform tasks with response conflict[90]. Third, properties of the paracingulate gyrus differ in individuals with psychiatric disorders that impair performance monitoring, in particular schizophrenia[91] and obsessive–compulsive disorder[92].

A final challenge in translating findings between macaques and humans involves the intrinsic composition of the cortical tissue. For example, humans, but not macaques, have large spindle neurons in layer 5 of the cingulate cortex[93]. However, what difference this makes functionally and for the properties of the local field potential remains unknown. Furthermore, in both species, relatively little is known about

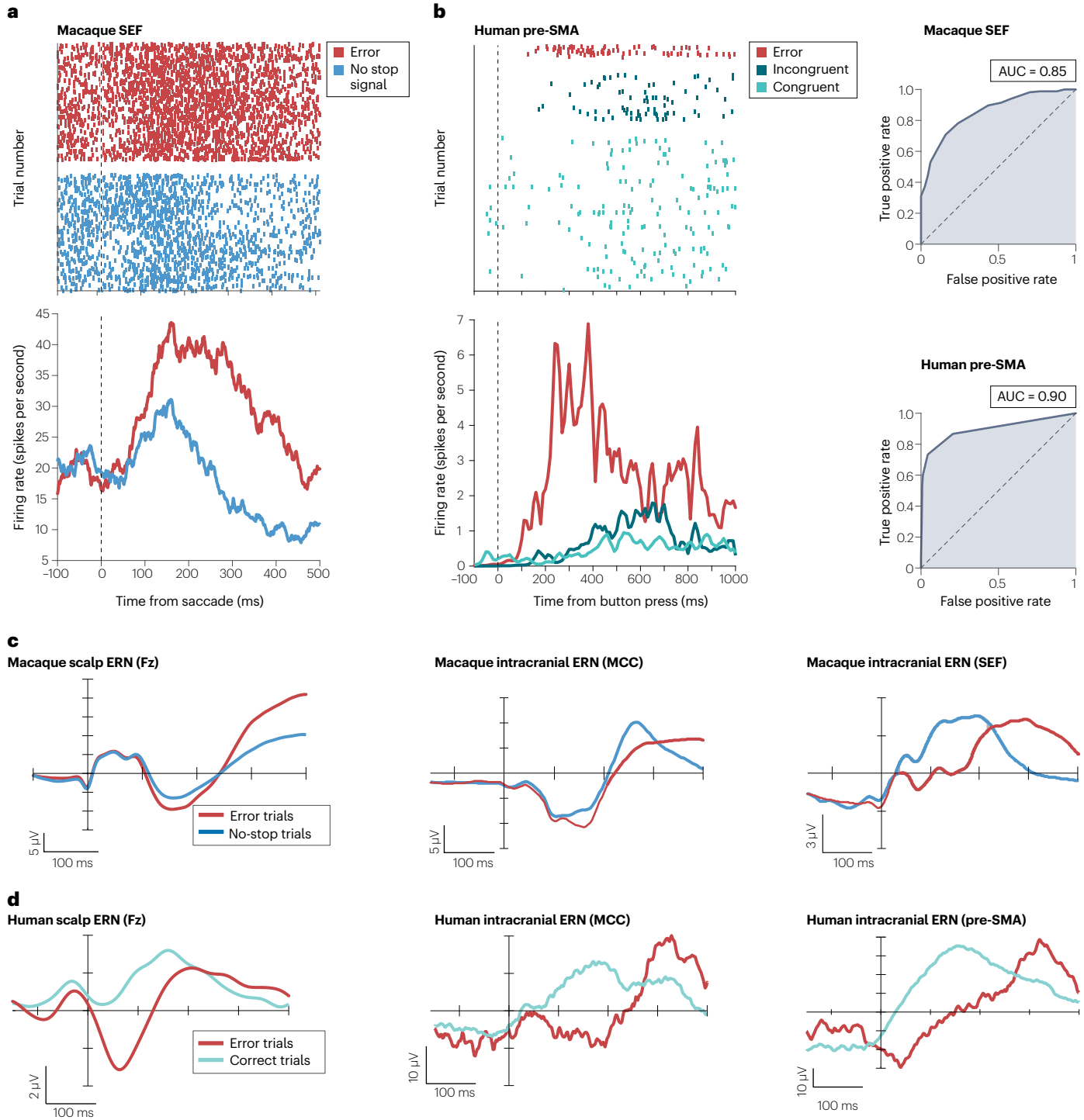


Fig. 2 | Error neurons and error-related negativity in macaques and humans. Each piece of data is shown for both species for visual comparison. **a**,**b**, Examples of single neurons responding to error trials with increased firing rate compared with correct trials. **c**,**d**, Average scalp and intracranial error-related negativity (ERN) from individual sessions, revealing a stronger negativity for error than correct trials in both species. AUC, area under the curve; Fz, frontal midline EEG electrode; MCC, middle cingulate cortex; pre-SMA, pre-supplementary motor area; SEF, supplementary eye field. Part **a** adapted from ref. [13], Springer Nature Limited. Data in part **b** are from ref. [12]. The left panel in part **c** is adapted from ref. [13], Springer Nature Limited. The middle panel in part **c** is adapted with permission from ref. [25], APS. The right panel in part **d** is adapted with permission from ref. [26], APS. The left panel in part **d** is adapted with permission from ref. [12], Elsevier. The data in the middle and right panels in part **d** are from ref. [12].

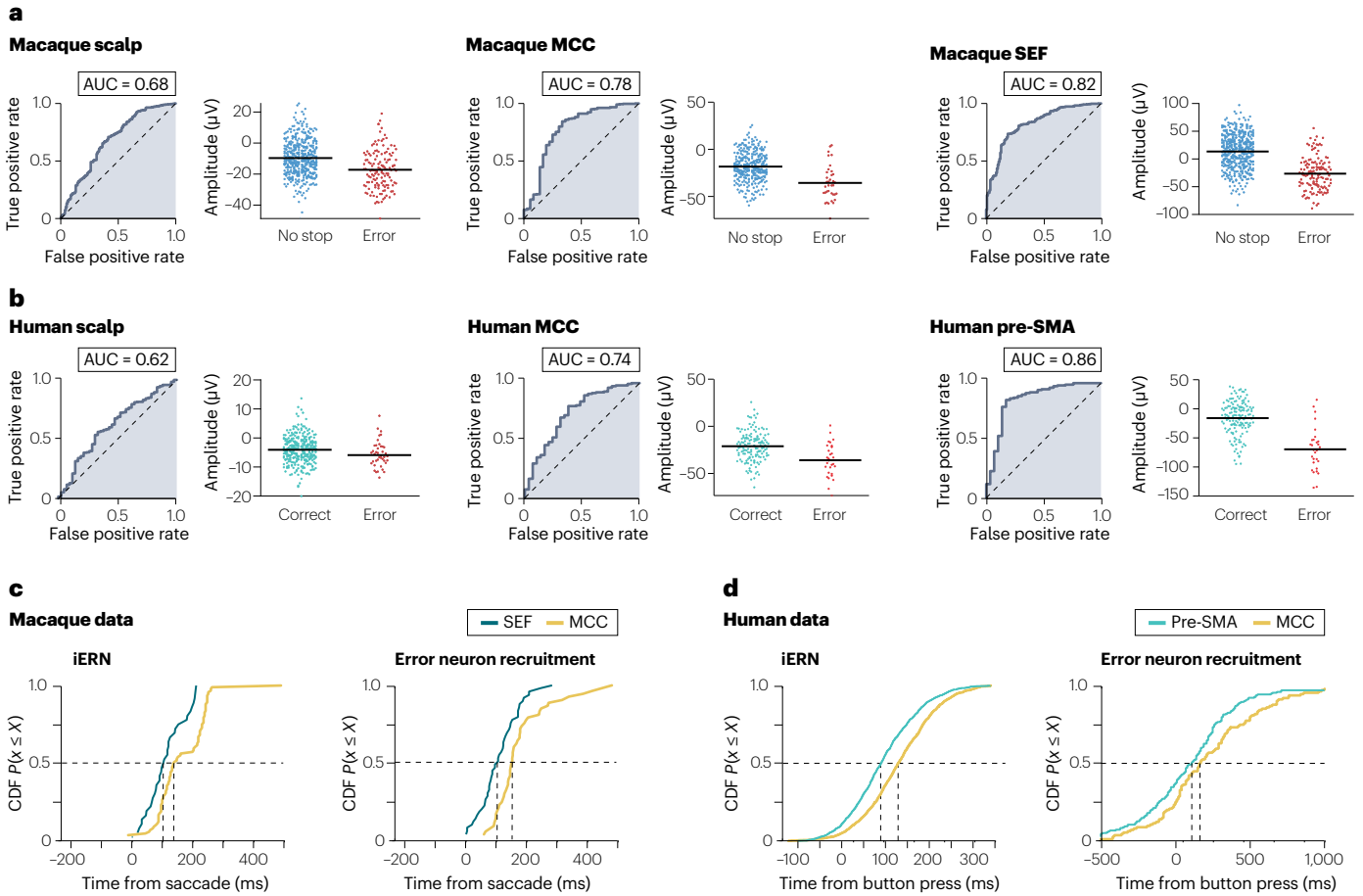


Fig. 3 | Reliability and latency of error responses in macaques and humans. Each piece of data is shown for both species for visual comparison. **a**,**b**, Single-trial reliability of the error-related negativity (ERN) amplitude at the scalp (left) and intracranial (right) level is high in both species. Each plot is from an individual session. **c**,**d**, Single-trial latency of the intracranial ERN (iERN) (left) and error neurons (right). In both species and with both metrics, error signals consistently appear first in the superior frontal gyrus. AUC, area under the curve; CDF, cumulative distribution function; MCC, middle cingulate cortex; pre-SMA, pre-supplementary motor area; SEF, supplementary eye field. Parts **a**,**b** are based on reanalysis of data shown in refs. [12,13]. Single-trial amplitudes were extracted using the method described in ref. [12]. Part **c** is adapted with permission from ref. [26], APS. Part **d** is adapted with permission from ref. [12], Elsevier.

how the functional organization of the cingulate sulcus contributes to performance monitoring[40]. In macaques, the MCC comprises several separate anatomically distinct areas, extending ventrally from the dorsal and ventral banks of the anterior cingulate sulcus, along the medial wall, to the corpus collosum[94]. Anatomical descriptions of the MCC in macaques note that the cytoarchitecture of the dorsal and ventral banks of the cingulate sulcus differ. Whereas the ventral bank of the MCC is identified as area 24, and as homologous to the MCC in humans[95], the dorsal bank is identified as an extension of area F6 caudally or area 9 rostrally (Fig. 1). Therefore, some researchers argue that the dorsal bank should not be considered part of the cingulate cortex proper[95], but others disagree[96,97]. In humans, recordings from the MCC are typically pooled across both the dorsal bank and the ventral bank, with no attempts at analysing the two separately. In macaques, recordings have been done mainly in the dorsal bank[20]. Hence, uncertainty persists about whether the dorsal and ventral banks of the MCC contribute differentially to performance monitoring, and where the boundary is between the dorsal MCC and area F6.

Tasks for performance monitoring

Relating findings across species and neural recording modalities entails understanding the cognitive constructs and demands of different tasks. Establishing equivalences across species, tasks and methods remains a considerable challenge. The validity of such comparisons depends greatly on the details of the tasks used and the instructions given. Guided by models of cognitive control and the Research Domain Criteria framework[7], we consider a set of cognitive constructs that jointly allow an animal to monitor its own behaviour. Different tasks engage these constructs to different degrees (Table 2). Although performance monitoring itself is only one of the eight constructs considered, the other seven constructs are all essential components needed for performance monitoring to be possible. The constructs we consider are as follows: goal maintenance, representing and implementing the instructed goal in a form of working memory[98]; response inhibition, inhibiting or delaying a motor response; stimulus selection, selecting one of several possible stimuli while ignoring others, which often requires feature or spatial attention; response selection, selecting one

of several possible motor responses, which can entail response conflict; performance monitoring, sensitivity to consequences after an action and difficulty during a decision; timekeeping, estimating when an event will occur or reproducing a time interval; stimulus–response mapping, rule specifying which response should be produced in response to which stimulus, which can be more compatible and automatic or more incompatible and challenging; and post-error adjustments, change of performance, typically by delaying responses, to increase accuracy[99].

Unlike experiments in humans, in which participants can verbalize instructions and choices, experiments with macaques must rely on sensory–motor paradigms in which performance is shaped through operant conditioning. In these tasks, participants express choices through an overt action, typically an eye, forelimb or digit movement. Thus, the use of common sensory–motor tasks can bridge the empirical gap between species.

The stop-signal (and change-signal) task examines response inhibition[100] (Fig. 4a,b and Supplementary Fig. 1a,b). After participants adopt a posture of readiness (for example, fixating), a stimulus is presented that requires an immediate response ('go' response; for example, a gaze shift). In some proportion of trials, while the prepotent response is being prepared, a second stimulus (the stop signal) instructs inhibition of the response, which initiates the inhibitory 'stop' processes. In the change-signal version, the second stimulus specifies an alternative response. The delay before the stop signal or change signal is presented after the first stimulus ('stop-signal delay' or 'change-signal delay') is adjusted trial-by-trial by a staircase procedure to allow successful responses on a subset of trials. Performance in the stop-signal task is modelled as the outcome of a race between go and stop processes[101]. Since the go and stop processes lead to incompatible outcomes, the activation of both of these processes results in response conflict[102]. Within this framework, the probability of making an error can be inferred from the stop-signal delay, which serves as a proxy for response conflict for successfully cancelling a trial[26,69,78]. A core feature of the stop-signal (change-signal) task is that, on every trial, there is uncertainty about whether the stop (change) instruction will happen[103]. Because the stop-signal (change-signal) delay is chosen randomly from a certain distribution, participants can develop predictions about the duration of this delay[78]. Hence, another key aspect of these tasks is timekeeping[78,104]. Another timekeeping demand involves maintaining the posture of inhibition (continued fixation of the central stimulus or holding digit or forelimb posture) for a sufficient interval, in order for performance to be qualified as 'correct'. Thus, two kinds of error are possible. The most common type of error is the failure of inhibition through production of the prepotent response despite the stop (change) signal. This type of error, especially in the case of long stop-signal delays, can result from the fact that there is not enough time for the stopping process to finish, even if it is promptly initiated. In addition, participants can determine that a 'go' response is 'correct' only by the absence of the stop signal. The less common type of error is the failure to maintain the posture of inhibition (or execution of the changed response). One key advantage of the stop-signal tasks is that the causes of errors are relatively homogeneous, as mentioned earlier, and are well described by mathematical models[105]. Error rates are well controlled (~50% of stop-signal trials) by adjusting the stop-signal delay using adaptive staircase procedures. A key signature of cognitive control is post-error slowing. In the stop-signal task, post-error slowing is prominent and involves delaying the initiation of a go response after a non-cancelled trial[106].

Go–no-go and anti-saccade tasks are also used to study response inhibition[107] (Supplementary Fig. 1c–f). Unlike the stop-signal (change-signal) task, however, the cue that determines the type of trial and stimulus–response mapping rule is usually presented before the cue instructing response initiation. Therefore, when the imperative stimulus is presented, there is no uncertainty about the expected response. In the go–no-go task, go trials involve a simple response (gaze shift or button press) usually with direct spatial mapping of the stimulus and the response. On no-go trials, participants must maintain the original posture. In the anti-saccade task, the pro-saccade trials involve a simple orienting response with direct spatial mapping to a visual target. The anti-saccade trials involve inhibition of the reflexive pro-saccade and production of a saccade in the opposite direction. Performance in the anti-saccade task differs according to whether the pro-saccade and anti-saccade trials are intermixed or blocked, with intermixed trials resulting in a higher error rate and longer response time[108]. Errors made in either

Table 2 | Cognitive constructs tested with performance-monitoring tasks in humans and macaques

Cognitive construct	Performance-monitoring task							
	Stop-signal task[a]	Change-signal task	Go–no-go task	Anti-saccade task	Simon task	Stroop task	Flanker task	MSIT
Goal maintenance[b]	2	2	2	2	2	2	2	2
Response inhibition[b]	2	2	1	1	1	1	1	1
Stimulus selection[b], spatial and feature-based attention	0	1[c]	0	0	2[d]	2[d]	2[c]	2[c]
Response selection[b,e]	0	2	0	2	2	2	2	2
Performance monitoring[b]	2	2	2	2	2	2	2	2
Timekeeping	1	1	1	1	0	0	0	0
Stimulus–response incompatibility[f]	1	2	1	2	2	2	1	2
Post-error adjustments[g]	2	2	1	1	1	1	1	1

Rating scale: 0, not required to perform the task; 1, required but task not designed to test the cognitive construct indicated; 2, required because task designed to test the cognitive construct indicated. MSIT, multisource interference task. [a]An alternative and equivalent name for the stop-signal task is the 'countermanding task'. [b]Construct that is part of the Research Domain Criteria framework of the US National Institute of Mental Health. [c]Relates to spatial attention. [d]Relates to feature-based attention. [e]Applies only to tasks in which multiple different actions are possible and also referred to as 'response conflict'. [f]In some tasks, whether there is a stimulus–response conflict depends on the instructions (for example, word reading versus reporting ink colour in the Stroop task). [g]Restricted to the next trial (not within-trial correction).

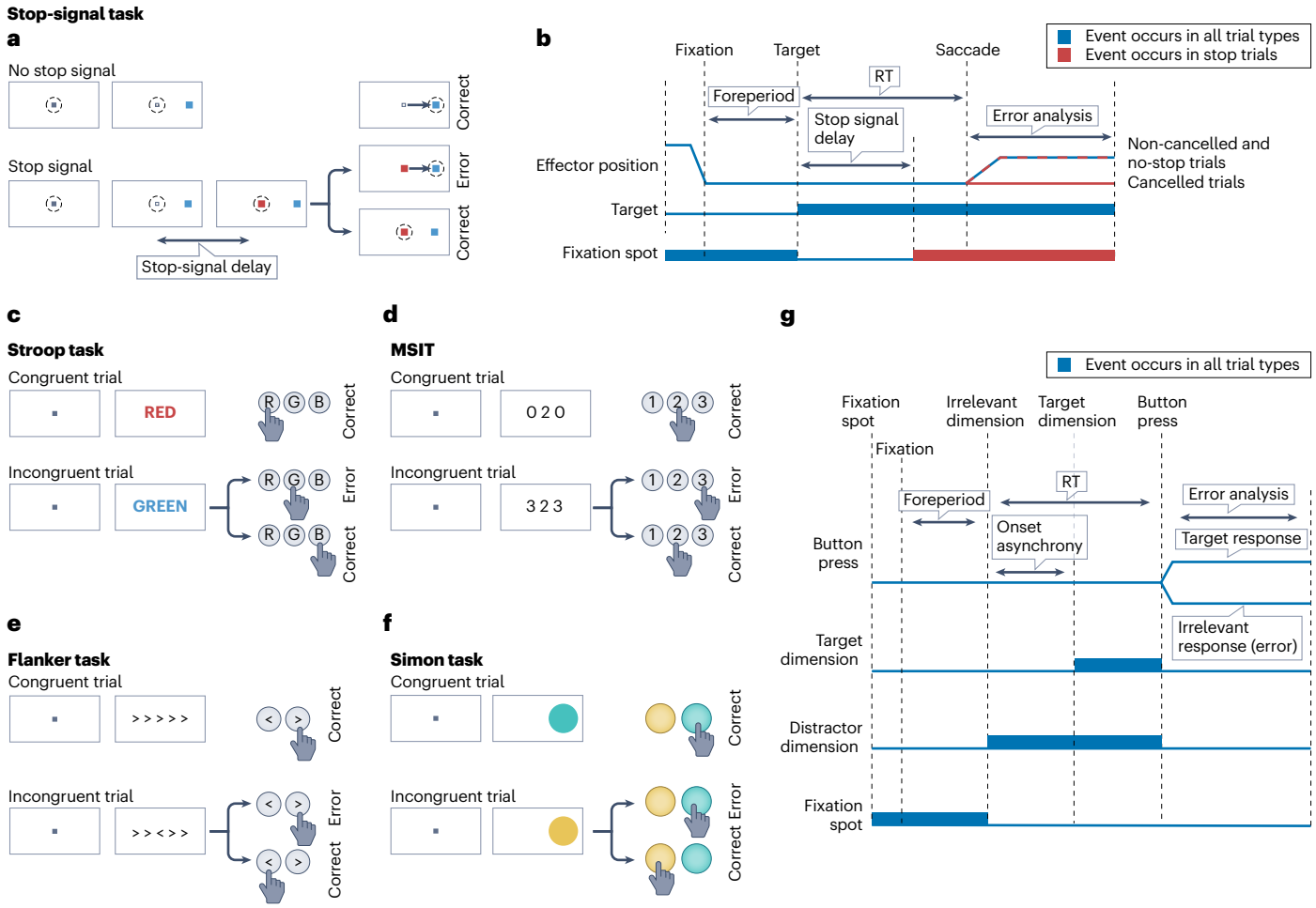


Fig. 4 | Tasks for studying performance monitoring in macaques and humans. **a**–**f**, Summary of tasks that have commonly been used to study performance monitoring; except those requiring reading (**c**,**e**), the tasks are suitable for both macaques and humans. For each task, the sequence of screens shown to the participant (left side) and the timing of critical events during a trial (right side) are shown. **a**,**b**, Stop-signal task, which requires participants to stop a movement when the stop signal is shown (red). **c**–**f**, Screens for the four human tasks: Stroop task, flanker task, multisource interference task (MSIT) and Simon task, in which participants are prompted to make a response by button press indicating the word colour, unique number, central target orientation or identity of the target, respectively. **g**, Timing applicable to all tasks shown in parts **c**–**f**. See Supplementary Fig. 1 for an illustration of the other three commonly used tasks: change-signal task, go–no-go task and anti-saccade task. Table 2 shows the cognitive constructs engaged by these tasks.

the go–no-go task or the anti-saccade task are due to a failure in incorporating the stimulus–response mapping rule and/or a failure in response inhibition[107]. Of note, participants can make an error by failing to maintain the effector position for a predetermined duration before the trial is considered successful in no-go trials in the go–no-go task.

The Simon task, the Stroop task, the flanker task and the multisource interference task are used to examine response selection, response inhibition, stimulus–response mapping and performance monitoring in humans (Fig. 4c–g). In these tasks, a single visual stimulus encodes multiple feature dimensions (spatial location, colour and/or word meaning), each of which can prompt a different response. Participants are instructed to respond as quickly as possible to one of the feature dimensions while ignoring the others. This is difficult when the feature dimension to be ignored is mapped to a prepotent, even habitual response. The source of the prepotent response is spatial (Simon task)[109], reading (Stroop task)[110], irrelevant visual distractors that attract attention (flanker task)[111] or mixtures thereof (multisource interference task)[112]. Trials on which the distractor and target dimensions point to the same response (congruent) engender a minimal level of conflict. A mismatch between the two dimensions (incongruent) engenders response conflict between the target and distractor action representations. Selecting the response associated with the relevant dimension requires inhibiting the prepotent response[75,113], which leads to correct responses with longer response times. Errors in these tasks arise endogenously: participants always have all the information needed to determine the correct response. For the same reasons, errors can be avoided by trading speed for accuracy. Given this property, error rates in these tasks are more difficult to control and generally lower. All four tasks result in post-error slowing.

Errors in the Stroop task and the multisource interference task are due to interference between high-level cognitive processes (related to language; for example, reading a word or number) only

available in humans. Nevertheless, Stroop-like effects can be produced in macaques using either extensive training on a particular stimulus–response mapping or through numerosity-based competition[114–116]. The extensive training required to produce these effects in macaques highlights one of the key strengths of Stroop-like tasks in humans; that is, they produce interference and errors immediately, showing that the underlying mechanisms do not require training.

Proposed conceptual model

We adopt the framework of forward–inverse modelling to conceptualize the processes underlying performance monitoring and cognitive control[117,118] (Fig. 5) and to explain the computation and roles of different types of ex post performance-monitoring signals. In motor control, the forward model estimates the state of the motor effectors (for example, position, velocity and acceleration) on the basis of delayed and noisy sensory feedback and predicts their future states given a motor command[119,120]. The inverse model generates a motor command in a feedforward manner given a desired movement trajectory and the effector's current state[119,120]. Here we generalize these concepts to the network-level neural dynamics involved in cognitive control and performance monitoring to propose a mechanism by which ex post performance-monitoring signals are computed.

Viewed within the actor–critic framework[121,122], the critic comprises the action and control forward models, whereas the actor comprises the feedback controllers, control inverse models and action selection mechanisms (Fig. 5). Each of the different actions competing for final motor output (whether compatible or incompatible with the goal) is represented by a corresponding action forward model (Fig. 5). After an action is selected, it is sent for execution as a motor command and back to the critic as a corollary discharge. To compute whether an action error occurred, this corollary discharge is compared against the output of the forward model that predicts the goal-compatible action. To compute the ex post conflict signal, the same corollary discharge is compared with the output of the forward model that predicts the goal-incompatible action. The control forward model learns to predict whether the current control settings will result in action errors and/or ex post conflict, on the basis of task representations (possible actions) and a corollary discharge of the 'control command' generated by the controllers. A control prediction error is generated when either an action error or ex post conflict occurs unexpectedly, signalling that the current control settings are inappropriate or inadequate for the current task.

We characterize the control command in terms of the identity and intensity of the mechanisms used to influence action selection. It is composed of feedback control generated by feedback controllers and feedforward control generated by the 'control inverse model' (Fig. 5). Response inhibition triggered by the stop signal is an example of within-trial feedback control, whereas post-error slowing[123,124], post-conflict slowing[125] and conflict sequence effects[75,126] are examples of between-trial feedback control. We posit that feedback control involves global adjustments of motor readiness, arousal and attention that influence

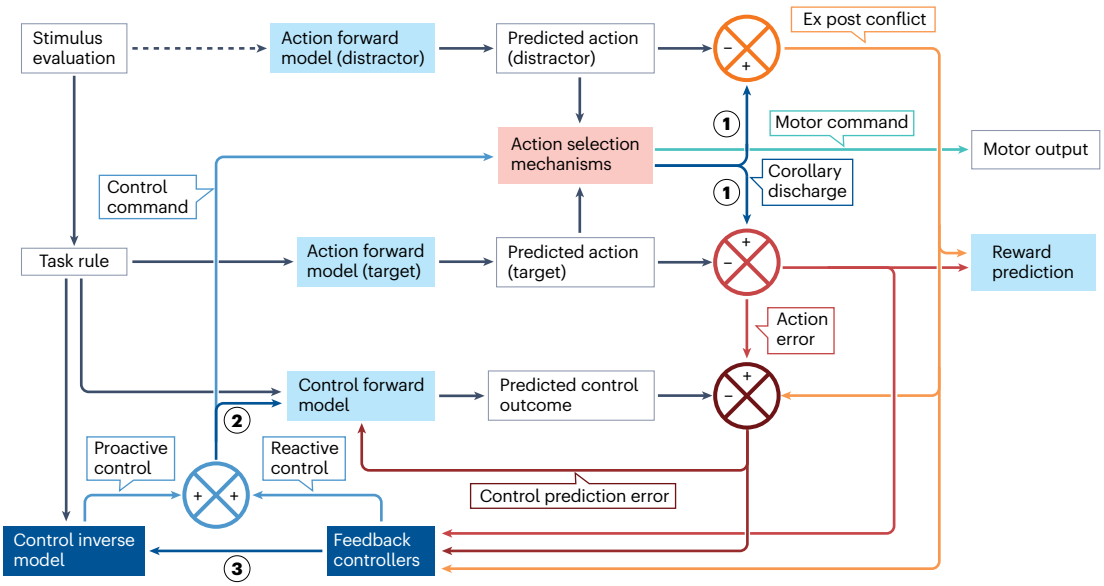


Fig. 5 | Conceptual framework for action error computation. The available actions under a given task rule and stimulus are predicted by action forward models (light blue). This includes both the correct response (target) and the incorrect response (distractor). The action selection process (red box) then chooses between one of the possible actions. Action selection is modulated by the control command (blue line), which is composed of proactive and reactive components (blue). The feedback controllers use performance-monitoring information from the prior trial to provide reactive control, and the control inverse model provides proactive control based on the task rules. Three kinds of performance-monitoring signals are computed (none of which depends on external feedback). The red cross computes action error signals by comparing the selected action, conveyed as corollary discharge 1, with the predicted goal-compatible action. The orange cross computes ex post conflict signals that are the result of comparing the selected action, conveyed as corollary discharge 1, with the predicted goal-incompatible action. The dark red cross computes control prediction error by comparing the predicted control outcome and the actual control outcome (error, conflict); it can also recruit feedback control. Action errors and ex-post conflict are used to predict the occurrence of reward. The control forward model predicts whether the current control settings, conveyed as corollary discharge 2, will result in an action error and/or ex post conflict. The extent to which feedback control was recruited is provided to the control inverse model as corollary discharge 3. Light blue boxes are forward models (predictors), dark blue boxes are controllers and the red box is action selection. Dark blue arrows are corollary discharges.

subsequent task performance only transiently and non-specifically. The inverse model generates the proactive control signal on the basis of the desired control outcome (for example, a point on the speed–accuracy curve) and its knowledge of the dynamics of the selection process and the statistical structures of the identity and intensity of control demand in a task. An important type of control that can be influenced by either control process is adjusting the time used for stimulus evaluation and action selection. We posit that this form of control is subserved by time estimation processes in the MFC[127,128] and response inhibition through the hyperdirect pathway (a monosynaptic axonal connection between cortical areas and parts of the subthalamic nucleus)[103,129–132]. For a new task, a control inverse model is initially not available, and the reactive feedback controller supplies the entire control command. As the control inverse model is adapting, the output of the feedback controller is gradually reduced as training proceeds. This distinction between proactive control and reactive control is motivated by the dual mechanism of the control framework[133]; our novel contribution is to assign these two processes to the control inverse model and feedback controllers, respectively, and to propose that the output from the feedback controllers serves as a training signal for the control inverse models.

Internal models in the MFC

We next consider how the outlined model might be instantiated across different parts of the brain (Fig. 6a). We propose that the SFG first computes action error signals and ex post conflict signals, and then conveys them to the MCC (Fig. 6a). These ex post signals are used to recruit feedback control, which in turn trains the control inverse model in the MCC. When the SFG error signals occur more often, or with more intensity than predicted by the MCC, the MCC uses the control prediction error to revise control forward and inverse models. We posit that action selection is accomplished in the basal ganglia, with the outcome reported back to the SFG and MCC as a corollary discharge through the thalamus.

We distinguish between the specification and implementation of cognitive control, with the MFC being concerned primarily with the former[134,135]. We propose that the MCC instantiates the control forward model (which computes control prediction error), whereas the SFG instantiates the action forward models (which compute action error and ex post conflict). Although many of these predictions remain to be tested, these hypotheses are motivated by the following findings, which supplement those discussed earlier herein.

Evidence for inverse models and feedback controllers comes from findings indicating that the MFC encodes variables that reflect aspects of the currently active control settings[18]. In humans, MFC neurons encode conflict probability estimated from the history of experienced conflict[18]. When two types of conflict co-occur in the same task, the probability for each type is represented separately, suggesting that these neurons are not a reflection of generic arousal signals. The conflict probability encoded in the MFC predicts reaction time and error likelihood, and it explains significantly more variance in reaction times than conflict in the immediately preceding trial alone (a proxy for feedback control). The activity of neurons encoding conflict probability changes from one trial to the next in proportion to the degree of conflict experienced on the current trial; we view this form of updating as evidence for acquisition of the control inverse model[18]. A similar conflict probability signal in the MFC can also be decoded using fMRI[136,137]. Similarly, the activity of macaque SEF neurons before trial onset predicts whether animals are emphasizing speed or accuracy in a visual search task, which is evidence for proactive control settings[122]. The strength of the correlation between iERN amplitude and error neuron firing rate in the MCC predicts the extent of post-error slowing, thereby providing evidence that the MCC is involved in specifying feedback control[12]. Further causal evidence for the SFG's involvement in specifying or implementing control comes from microstimulation studies. Electrical microstimulation of the SEF increases performance accuracy in the stop-signal task by delaying saccade reaction times but does not delay saccades in simple visually guided saccade tasks[106]. Microstimulation of the pre-SMA in a cued-switching saccade task increases performance accuracy after response rule switching by delaying saccades[138], an effect possibly mediated through the hyperdirect pathway of the pre-SMA to the subthalamic nucleus[139]. Thus, experimental evidence indicates that the MFC plays a role in implementing the inverse models and the feedback controller.

Evidence for forward models comes from examining representations of choices and ex post conflict signals. First, during a task in which participants make either a perceptual or a memory-based 'yes' versus 'no' decision, neurons with a stimulus onset-triggered response in the pre-SMA predicted the subsequent choice regardless of whether it was true or false[34]. This encoding of choice is task dependent but effector (button press or saccade) independent, suggesting that these neurons represent a mapping between the stimulus/internal representation and the 'goal-appropriate' response based on the current task rule (Fig. 5). Similarly, in a study of value-based decision making, pre-SMA neurons also predicted choices well before action execution[140]. These data indicate that during action selection, the SFG represents target and distractor responses instead of actual movement (for example, which button is pressed) as expected from a forward model. Second, in macaques performing the stop-signal task, neurons in the SFG that encode ex post conflict signals are defined by their correlation with the probability of non-cancelled saccades[69,77,78]. This is compatible with representations of forward models because before the stop signal occurs, the action forward model representing the 'go' action makes a prediction that a saccade motor plan will be selected for output, with its probability increasing as a function of time. The longer the stop-signal delay, the greater the probability is for the 'go' saccade to occur and the larger the prediction error is when the 'go' action is successfully cancelled, thereby resulting in an ex post conflict signal[102].

Lastly, we consider the distinct and common roles of the SFG and the MCC. In both humans and macaques, error signals appear first in the SFG and then in the MCC. This indicates a hierarchical relationship: the MCC learns to predict the internal error signals computed by the SFG to adjust control settings for future trials[141]. An unexpected action error can signal insufficient control settings, which necessitates the recruitment of feedback control, or incorrect control settings (for example, a change of task or reward environment), which necessitates the revision of control forward models. In macaques, simultaneous recordings in the MCC and the SFG during a time estimation task, in which errors are informed by a loss of reward, provide support for this conclusion[141]. Although error signals were present in both areas, only baseline activity in the MCC showed signatures of a control inverse model by predicting the animals' switching to an alternative response rule. In the same task, electrical microstimulation in the SFG increased activity in the MCC, but only when the appropriate response rule was inferred rather than instructed. These results support the view that directional flow from the SFG to the MCC is a key aspect of bridging monitoring with control, with the SFG implementing the forward model and the MCC implementing the inverse model. We note that the purpose of the proposed model is to explain only error detection instead of the many other functions of the MFC[14,17,38–40].

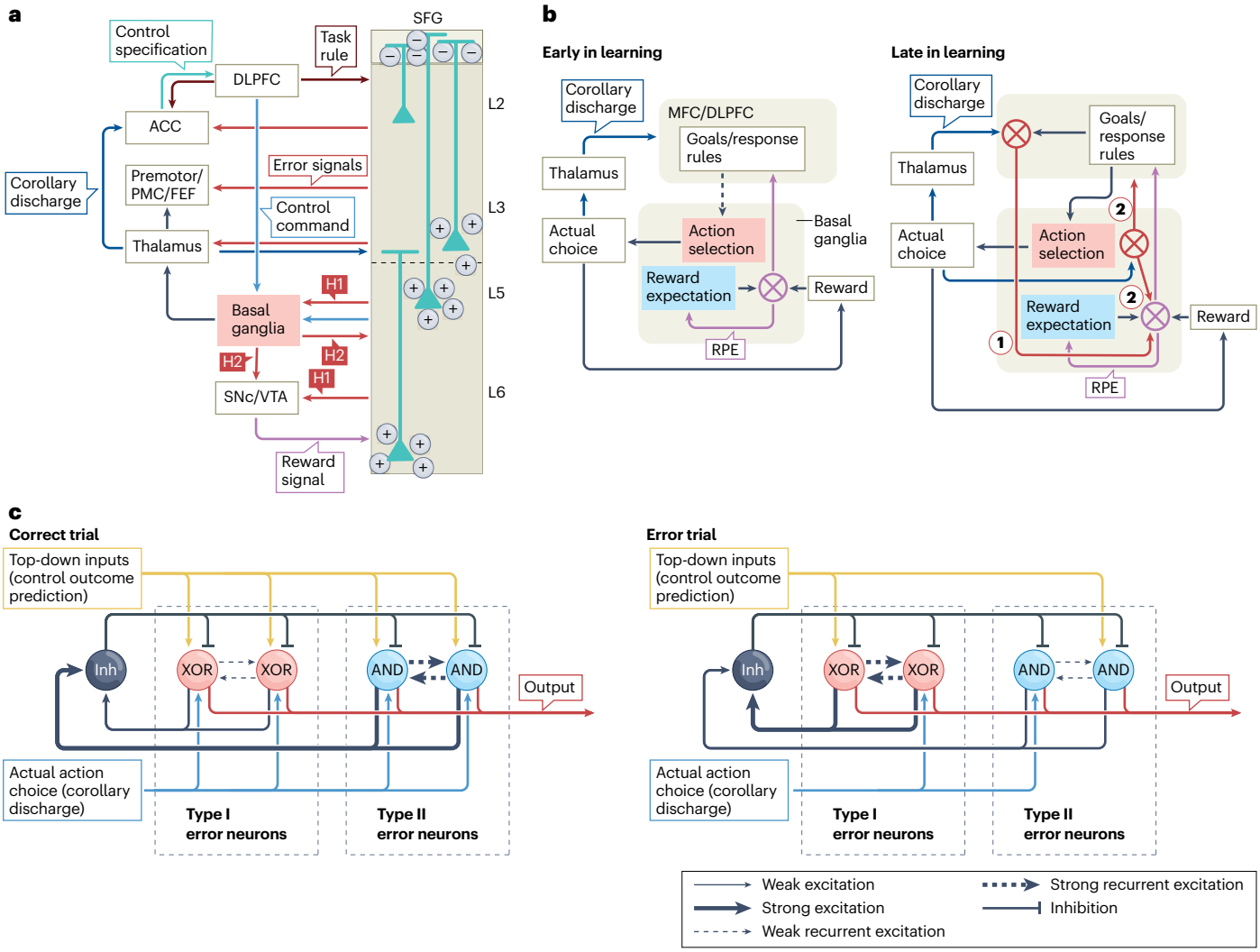


Fig. 6 | Circuit model for action error computation. a, Putative microcircuitry model of action error computation. The middle cingulate cortex (MCC) receives error signals from the superior frontal gyrus (SFG) and instantiates the control forward models and the specification part of the control inverse model. The inverse models that generate control commands are part of the MCC and the dorsolateral prefrontal cortex (DLPFC). The basal ganglia receive control commands and help to select an action, which is conveyed back to the SFG and the MCC as a corollary discharge. The illustrated pyramidal neurons are error neurons in the SFG. Red arrows indicate flow of error information. The positive and negative symbols represent electric charges contributing to the dipole influencing the electroencephalography signal upon error neuron activation. **b**, Error computation at different stages of learning. When a new task is being learned, learning is driven by reward prediction errors (RPEs), which are conveyed by midbrain dopaminergic neurons to train cortical representations (left). After choice and response rules have been learned, errors are either computed locally in the medial frontal cortex (MFC) (hypothesis H1, our proposed model) or in the basal ganglia and are passed on to the MFC (hypothesis H2, the reinforcement learning–error-related negativity model). **c**, Winner-take-all computation of action errors. Type I error neurons increase (decrease) spike rates in error (correct) trials, implemented as anti-coincidence detectors. Type II error neurons decrease (increase) spike rates in error (correct) trials, implemented as coincidence detectors. The top-down inputs represent the predicted action, and the thalamic inputs represent the action choice conveyed as a corollary discharge. Neurons labelled 'AND', 'XOR' (exclusive or) and 'Inh' are coincidence detectors, anticoincidence detectors and inhibitory, respectively. ACC, anterior cingulate cortex; FEF, frontal eye field; L, layer; PMC, primary motor cortex; premotor, premotor cortex; SNc, substantia nigra pars compacta; VTA, ventral tegmental area.

Implementing action error computations

We envision that the computational processes that underlie action errors interact with other parts of the brain (Fig. 6). The action selection process is a biased competition that occurs in cortico–basal ganglia–thalamocortical loops[142–145] and is influenced by top-down inputs that originate in the dorsolateral prefrontal cortex and represent behavioural goals[142]. The conclusion of this competition or selection process produces a motor output, together with a corollary discharge that conveys the action choice back to the SFG (as well as other parts of the cortex, including the MCC[146]; but note that the thalamic nuclei that project to the SFG are different from those that project to the MCC) via projections from the thalamus to deep layer 3. This is based on the

finding that the neurons at the layer 3–layer 5 border signal errors earlier than do others[13]. There, the actual action choice is compared with the predicted action choice generated by the action forward models. On the basis of whether these two signals match or do not match, an error signal is generated, which then propagates to layer 2 and layer 6, conveyed to various upstream and downstream targets (Fig. 6a), such as the MCC. Although most of the experimental evidence supporting our model is obtained directly from agranular SEF data[147,148], some of the assumptions of the underlying microcircuitry are based on work on granular neocortex[149] and remain to be confirmed for agranular cortex.

The experimental findings discussed above inform the following potential circuit-level implementation. In the SFG, the predicted action choice and the actual action choice (thalamic inputs) are represented by multivariate patterns of neural activity, and the matching of these patterns may provide a mechanism for detecting errors. Type I neurons may implement anti-coincidence detection (active when there is a mismatch and inactive when there is a match or no inputs; that is, an exclusive or (XOR) operation), whereas type II neurons implement coincidence detection[150] (active when there is a match and inactive when there is a mismatch or no input; that is, an AND operation). These two processes might be implemented biophysically at the single-neuron level through a combination of strong afterhyperpolarization and/or nonlinear dendritic computations[151], with the two types of neuron competing for activation through a soft winner-takes-all process mediated by shared inhibition[152,153] (Fig. 6c).

Across the cortical sheet of the MFC, dipoles created by synchronous thalamic inputs and the resultant returning current at the apical dendrites, as well as top-down inhibition mediated by intralaminar inhibitory neurons, are aligned, and summate to generate the ERN (Fig. 1b). These dipoles are possibly strengthened through the activation of calcium spikes initiated in the layer 5 pyramidal neurons[55] (also see ref. [48]). In humans, most error neurons are of type I (74% in the MCC and 60% in the pre-SMA), indicating that they may contribute more to the dipoles that give rise to the ERN than type II error neurons. However, the relationships between the scalp ERN, the iERN measured in different parts of the MFC and single-neuron activity are complex and little understood (similarly for BOLD signals), leaving it an open question to what extent type I and type II error neurons (and the resulting iERN) differentially contribute to these macroscale metrics of error monitoring.

If error computations are subserved by specific cortical microcircuitry, we would expect that error neurons compute errors across different tasks. In support of this, in the human MFC, most error neurons are domain general, with the others being dependent on the task performed[18]. Alternatively, error neurons might emerge from learning of specific tasks. Further studies are needed to investigate how error signalling changes as a function of learning several tasks to investigate this prediction.

Predictions and suggested experiments

In our model, action errors are computed on the basis of a corollary discharge conveying the action choice rather than the proprioceptive feedback generated by executing the action to the MFC (this hypothesis is mentioned in one of the first articles describing the ERN[46]). We argue that the corollary discharge is provided through thalamic input to layer 3/5 of the SFG. This assumption predicts that disrupting thalamic input will disrupt action error computation but disrupting proprioceptive feedback will not. In support of this, the ERN is abolished and error detection is compromised in patients with lesions in the ventral anterior nucleus and the ventral lateral anterior nucleus of the thalamus[154]. In primates, both the MCC[146,155] and the SFG[156] are innervated by neurons in these nuclei of the motor thalamus. Therefore, lesions in the ventral anterior nucleus or the ventral lateral anterior nucleus are expected to compromise the corollary discharge to both the MCC and the SFG, leading to profound deficits in error signalling, which is demonstrated in this lesion study[154]. Transient inhibition of the ventral anterior nucleus and the ventral lateral anterior nucleus results in increased error rates in inhibiting saccades[157]. This suggests that these nuclei are important for response competition and thus are potential sources of the corollary discharge that reports the results of this competition. The SFG, in addition, is innervated strongly by neurons in the mediodorsal nucleus. Consistent with this, the ERN in an anti-saccade task is attenuated by lesions to this nucleus[158]. By contrast, the ERN was largely normal in a deafferented patient without sensory feedback[159]. A critical experiment will be to examine the effects of transiently disrupting neuronal activity in these nuclei on the activity of error neurons. A case study of a patient with an extensive unilateral MFC lesion provided direct behavioural support for a role of a corollary discharge in learning from errors. In that study, learning from external feedback was impaired only when the individual was not able to observe the action he had executed, in which case visual feedback was not available and the patient thus had to rely on the corollary discharge to attribute outcomes to a specific action[160].

Lesioning the SFG (but not the MCC) should impair action forward models, resulting in a loss of action error and ex post conflict signals. As a result, controllers cannot be recruited to ensure the achievement of goals, and behaviours are dominated by habitual actions[161]. Lesioning the MCC, by contrast, would not abolish error detection but would rather compromise the ability to use the error signals generated by the SFG to acquire control forward and inverse models. As a result, such lesions would result in an over-reliance on feedback control, a failure to engage proactive control and a failure to learn from internal error signals.

Action errors are different from the types of 'action outcome' that may serve to update internal models of the external environment rather than of cognitive control. We therefore predict that action errors can recruit feedback control but exogenously signalled action outcomes do not. Results from a human behavioural study of fast typing provide support for this prediction[2]. A proportion of action errors made by typists were covertly corrected, whereas some correctly typed contents were covertly switched to typographical errors. Although the typists took ownership of both the corrected errors and the inserted errors and thus subjectively experienced agency for both (thus qualified as 'action outcomes'), only the corrected errors triggered post-error slowing.

We expect that the MCC signals estimates of the 'volatility' of the rates of commiting action error or experiencing response conflict. Volatility is defined here as the rate of change over time[137,162]. The stop-signal task is well suited to test this as the target error rate can be externally controlled. In this task, in a high-volatility block, the target error rate would vary frequently. Similarly, in a response conflict task, in a high-volatility condition the probability that a given trial contains conflict would vary frequently. We expect that the error signals in the MCC depend on volatility, whereas those in the SFG do not. This experiment can thus dissociate action error signals from control prediction error signals: the former should not depend on volatility manipulations, whereas the latter should.

We posit that action errors are represented separately from reward prediction errors, and that it is the cells that signal action errors but not those signalling reward prediction errors that predict the amplitude of the ERN. This prediction is borne out in the macaque SEF, where error

neurons are distinct from neurons signalling gain or loss of rewards and the ERN amplitude is predicted by the former, but not the latter[13]. In addition, in a difficult visual search task that dissociates choice errors from reward prediction errors, distinct neurons in the SEF signalled each error type[122]. In the macaque pre-SMA and SMA, most error neurons do not respond to unexpected reward[16]. By contrast, many (not all) of the action error neurons in the macaque MCC also signalled unexpected omission of rewards[19], suggesting that the MCC error neurons are more abstract[163], and that the SFG action error signals may represent one of the many inputs to these neurons. This prediction remains to be tested in humans.

As an action error predicts loss of reward, it must be conveyed to the midbrain dopaminergic system to serve as an input for computing reward prediction error. We expect simultaneous recordings from MFC error neurons and midbrain dopaminergic neurons to show that the responses of these two groups of neuron are correlated on a trial-to-trial basis in terms of magnitude and latency, that MFC error neurons respond earlier than dopaminergic error-signalling neurons and that disrupting error signals in the MFC abolishes error signals in dopaminergic neurons. This proposal is backed by established evidence that dopaminergic neurons integrate cortical inputs[164]. By contrast, if dopaminergic error neurons fire earlier than MFC error neurons, the reinforcement learning (RL)–ERN model[48], which posits that action errors are merely prediction errors, would be supported (see the discussion later for details of this model and further justification of this experiment).

It is important to establish the flow of error information from the SFG to the MCC. In the MCC, we predict that feedforward input from the SFG provides error signals to the superficial layers, which would be visible as superficial sinks in laminar recordings and earlier responses by layer 2/3 error cells relative to those in deeper layers. Disrupting action error signalling in the SFG should abolish these sinks and sources in the MCC as well as error neuron firing in the MCC.

We hypothesize that the MCC instantiates the control inverse models, whereas the SFG instantiates the feedback controller. This proposal predicts that disrupting MCC function will disrupt the acquisition of control inverse models, thereby increasing reliance on the feedback controller, whereas disrupting SFG ex post signalling will disrupt both the feedback controller and the control inverse model. The magnitude of the ex ante conflict signal would be expected to track such causal manipulations. A well-learned control inverse model (for a particular task) could generate the feedforward control commands that anticipate conflict, effectively minimizing the conflict experienced during action selection. Consistent with this, prolonged practice greatly diminishes the Stroop effect[110]. We predict that the responses of neurons signalling ex ante conflict in the MFC become weaker as participants gain more practice in a task involving response conflict. Compatible with this interpretation, a human neuroimaging study[165] showed that conflict signals were greatly diminished when conflict was predicted by a preceding visual cue.

The control inverse model should allow participants to proactively switch between multiple sets of control settings. One setup for testing different control settings is to change the speed–accuracy trade-off by instruction, as we did recently in macaques[122,166]. The expectation would be that MFC neurons accomplishing the control inverse model, such as the recently described neurons signalling conflict probability[18], change their activity systematically following such cues. Again, we expect that ex ante conflict neurons would be modulated on the basis of the control settings for otherwise identical stimuli.

Computing the control command requires domain-specific control signals (for example, specific control signals targeting the specific type of conflict experienced). However, before the control inverse model is acquired, the feedback controller generates generic controls signal (for example, global changes in the level of arousal and/or excitation in the motor system). Consistent with this, human MFC neurons represent both domain-general and task-specific monitoring signals[18]. A key experiment will be to examine the temporal dynamics of these two types of signal, with the expectation that for a new task, domain-general signals are available immediately, whereas task-specific signals emerge gradually with task exposure and practice.

Theoretical models indicate that the largest source of the scalp ERN is the cingulate sulcus, owing to the orientation of the neuronal dipoles (Fig. 1b). If so, the scalp ERN should primarily reflect error signals in the MCC, which as we describe here appear later than those in the SFG. However, in hemispheres with a PCS or at more caudal locations, dipoles within the SFG are well positioned to impact the scalp ERN as well. This is compatible with the findings we reviewed that activity in the SFG is predictive of the scalp ERN amplitude and latency. It will be important to determine experimentally the extent to which dipoles in different parts of the MFC differentially contribute to the scalp ERN using simultaneous intracranial recordings. In addition, a critical open question is how performance-monitoring signals other than action errors, such as ex post conflict and control prediction errors, influence the scalp ERN.

Relationship with existing models

Three major models have been proposed to conceptualize how action errors are computed: mismatch theory, response conflict theory and reinforcement learning–ERN theory (RL–ERN). Our model borrows key ideas from mismatch theory[46,167,168]. This theory proposes that action errors are computed by comparing a delayed representation of the correct response and the committed erroneous response. The representation of the 'correct' response in this instance is derived from continued bottom-up stimulus evaluation and the mapping of the stimulus to response representations. A key common feature shared between mismatch theory and our model is that in order to compute action errors, both the executed action and the 'correct' action are represented. However, in our model, the correct action representation is a prediction made by the action forward models rather than a result of continuous stimulus evaluation.

Response conflict theory proposes that the MCC error signal is computed as the conflict between the representations of the executed erroneous action and the delayed correct action that is computed by continued stimulus evaluation[74]. Predictions of this theory include the following: if neurons signalling action errors are conflict detectors, they should also signal conflict during action selection (ex ante conflict) and following action selection on correct trials (ex post conflict); and given that reaction times differ significantly between congruent and incongruent errors[12], which suggest that their levels of conflict differ, error neuron responses should also reflect this difference. Recordings from the human MFC do not support these predictions. Most error neurons do not signal ex ante conflict or ex post conflict in correct trials, and most ex ante conflict neurons do not signal errors. Most ex ante conflict neurons also do not signal conflict after an error is made[12,18]. In macaques, similarly, error neurons and ex post conflict neurons in the SEF constitute separate groups[69,78]. Together, these data argue that it is unlikely that action errors are computed as a conflict signal.

Glossary

Action errors
Errors made owing to a failure of cognitive control, so that goal-directed responses yield to more automatic responses.

Control command
The identity and intensity of control mechanisms used in a given trial, including contributions from the control inverse model and the feedback controller.

Cognitive control
Ability to adjust behaviour according to the desired goal and to avoid repeating errors through detection of conflict, errors and success.

Corollary discharge
Copy of efferent signals that is provided to sensory or higher-level cognitive structures to enable forward models.

Error-related negativity
(ERN). Evoked potential that is measured over the medial frontal cortex shortly after an action error.

Forward models
Internal models that predict the future state of the motor effector or action selection process on the basis of the current state and a copy of a control or motor command (obtained as a corollary discharge).

Feedback controllers
Controllers that are activated by internally generated feedback (such as action errors). They provide control in a reactive manner.

Inverse models
Internal models that compute which control command is most likely to lead to a desired future state of the system on the basis of knowledge of the dynamics of the system. Used to provide control in a proactive manner.

Medial frontal cortex
(MFC). Areas along the midline of the frontal lobe, including the pre-supplementary motor area, supplementary eye field and middle cingulate cortex.

Middle cingulate cortex
(MCC). A region in the medial frontal cortex, located ventral to the primary motor cortex, supplementary eye field, supplementary motor area, and pre-supplementary motor area in Brodmann area 24. This area contributes to diverse functions, including performance monitoring and cognitive control.

Paracingulate sulcus
(PCS). A sulcus found in some people but not in macaques that is aligned parallel to and located dorsal to the cingulate sulcus in the medial frontal cortex.

Performance monitoring
Ability to monitor one's own performance, either by self-monitoring or on the basis of external feedback.

Post-error slowing
A delay of response time in the trial following commission of an error response that is found commonly but not uniformly in various choice tasks.

Pre-supplementary motor area
(pre-SMA). An area in the medial frontal lobe, located anterior to the supplementary motor area in Brodmann area 6aβ, that contributes to cognitive control of action.

Superior frontal gyrus
(SFG). Dorsomedial gyrus of the frontal lobe situated along the midline. It includes the medial part of Brodmann area 6, within which the pre-supplementary motor area, supplementary motor area and frontal eye field are located.

Supplementary motor area
(SMA). An area in the medial frontal lobe, located anterior to the primary motor cortex in Brodmann area 6aα, that contributes to the production of self-generated actions.

Supplementary eye field
(SEF). An area in the superior frontal gyrus. It is distinguished from surrounding cortical areas by denser connections with visual and ocular motor structures and by its functional relationship with eye movements.

RL–ERN theory proposes that the ERN is a result of a negative reward prediction error caused by action errors, which is signalled by a pause in firing of midbrain dopaminergic neurons[48]. In this model, this pause in firing leads to a transient reduction in dopamine levels near the apical dendrites of pyramidal neurons in the MFC. The resulting disinhibition then causes these neurons to fire synchronously to generate the ERN. One motivation for this theory is that MCC lesions compromise the ERN but do not abolish behavioural signatures of error detection[33], which its proponents take as evidence that the error signal is computed elsewhere. More broadly, this theory provides a plausible mechanism for the midbrain dopaminergic system to train task learning in the prefrontal cortex[169]. Although this theory offers a plausible explanation for the FRN that follows external negative feedback when a new task is being learned (Fig. 6b), more recent data suggest action errors can be independently computed by the MFC to generate the ERN in a well-learned task without external feedback[12,13,18,19,106] (Fig. 6b). First, we showed that both the SFG and the MCC participate in ERN generation. The fact that MCC lesions do not abolish error detection may therefore be a result of intact error processing in the SFG. Moreover, because error neurons are widespread in the MFC[12,18] (random sampling from the area in human recordings yielded >20% error neurons), it is unlikely that partial lesions abolish all error-related signals. In addition, chronic lesions instead of acute lesions may be compensated by functional remapping, which may limit the validity of lesion studies[170]. Second, one macaque study using the stop-signal task showed that in the case of an action error, some midbrain dopaminergic neurons increased rather than decreased their activity as would be expected for a negative reward prediction error[171] (in this study, most of the responsive neurons were in the substantia nigra and only a few were in the ventral tegmental area). The absence of a pause in firing indicates that dopamine levels in the MFC are unlikely to decrease rapidly during errors. Third, conduction velocity in the axons of dopaminergic neurons is slow (~0.55 m s^{-1} (refs. [172–174])). Given the putative length of dopaminergic projections (in primates, the length of the midbrain–MFC pathway is ~56 mm), we estimate that ~100 ms elapses between initiation of a dopaminergic neuron spike and its arrival in its terminal in the MFC[78]. Translating reductions in extracellular levels of dopamine into neural activity is expected to induce additional delays (see also ref. [4]). This latency estimate still applies even if action errors are conveyed to the MFC by co-release of glutamate by dopaminergic neurons instead of a pause in firing[175]. Comparisons across studies indicate that the onset latencies of error signals in dopaminergic neurons and MFC neurons are similar[39,171]. Therefore, midbrain dopaminergic neurons are unlikely to be the source of MFC error signals. Rather, we posit that the ERN is a result of error computation processes within the MFC rather than the result of error signals conveyed by dopaminergic neurons to the MFC. Further experiments are needed to directly compare these two different hypotheses (outlined in Fig. 6b).

Several related models[176–179] argue that errors are detected as violations of action outcome predictions[180]. The predicted response outcome model argues that the MCC learns to predict the outcome and its timing of all possible actions in performing a task[176]. A related model of the MCC trains a recurrent neural network to predict action outcomes in action sequences[179]. A hierarchical version of the predicted response outcome model further assumes that there is a rostral–caudal abstraction gradient within the MCC with respect to error computations, with more rostral parts predicting the action-outcome prediction errors generated by more caudal parts[178]. The idea of hierarchical computation of prediction errors is broadly consistent with the predictive coding framework[118,181].

Relevance to psychiatric disorders

The ERN demonstrates potential as a biomarker for gaining insight into psychiatric disorders that implicate dysfunctional cognitive control, such as obsessive–compulsive disorder, attention deficit hyperactivity disorder, schizophrenia and autism spectrum disorders. For example, the amplitude of the ERN is increased in people with obsessive–compulsive disorder and reduced in individuals with schizophrenia compared with unaffected controls[182–184]. Modified recruitment of the error-monitoring system is apparent also in individuals who are merely at genetic risk of developing these psychiatric disorders[185]. This property is specific to the ERN: EEG signals elicited by external performance feedback do not differ significantly between patients with dysfunctional cognitive control and unaffected individuals[186,187]. Apart from these trait-like properties, the ERN also indexes cognitive states: ERN amplitude is reduced by sleep deprivation, alcohol consumption and substance abuse[188–190]. Therefore, the ERN provides an index of the overall integrity of the cognitive control system.

Owing to its co-variance with psychiatric disease, the ERN is currently the strongest candidate for a biomarker that indexes cognitive control abnormalities[191]. Compared with functional imaging, EEG measurements are more economical and procedurally more feasible and easier to collect in young children, and measure the electrical activity that underlies neuronal activity more directly. However, without the ability to relate specific aspects of the underlying neural processes to properties of the ERN, correlating ERN properties with disease states provides little insight into the underlying mechanisms. As we have reviewed here, this difficulty is now partially resolved for the ERN: we now know that specific types of cell within the primate MFC (localized in specific layers in macaques) exhibit activity that is correlated strongly on individual trials with the ERN. These cells have been identified in both humans and macaques, thereby establishing a closely linked animal model system within which the same biomarker as in humans is measurable and that can perform some of the very same tasks that humans can. A key next step is to identify the transcriptomic and epigenomic profiles of error neurons (for example, see ref. [192]), which would contribute to determining how error neurons could be targeted specifically by electrical stimulation and/or pharmacological manipulations. The macaque model is particularly well suited to test potential neuropharmacological manipulations of error monitoring. The links between behaviour, large-scale activity patterns, single-cell activity and parameters in models of psychopathology that the ERN–single neuron relationship establishes may enable the development and validation of new computational psychiatry models[193].

Conclusions

We have reviewed the parallels between error monitoring in macaques and humans. The apparent functional convergence in this process in these species is reflected in the temporal dynamics of single-neuron responses and the ERN, despite differences in tasks and between species. The spiking activity of error neurons predicts the amplitude of both the iERN and the scalp-measured ERN in both species, providing direct evidence that the ERN is generated in the MCC and the SFG. Many psychiatric disorders implicate dysfunctional performance monitoring and cognitive control. The cross-species approach described here not only validates macaques as an important model for such cognitive control-related psychiatric disorders but also provides the data motivating an overarching framework for how performance monitoring signals are computed.

The species-comparison approach in addition yields theoretical insights. It reveals that goal-directed action control can be analysed within the framework of forward–inverse modelling. This framework shows how action errors can be computed by comparing predictions made by forward models with the outcome of action selection conveyed by a corollary discharge. We have discussed algorithmic and microcircuitry models for computing action errors that incorporate the laminar information made available by macaque data. Computation of action errors can be implemented by a soft winner-takes-all network that comprises coincidence and anti-incidence detectors and also qualitatively accounts for neuronal data in both species.

The model of how errors are computed and implemented that we propose bridges multiple gaps between scales of observation, from single neurons and cognitive functions producing behaviour to scalp voltage measurements. Therefore, this model can be an important benchmark that synthesizes our current understanding of the neural computations underlying error monitoring and cognitive control. Equally importantly, this model makes specific predictions in the form of novel hypotheses that remain to be tested.

Published online: 27 January 2023

References

1. Rabbitt, P. M. Errors and error correction in choice-response tasks. *J. Exp. Psychol.* **71**, 264–272 (1966).
The original description of slowing after errors, revealing the existence of and a means to investigate performance monitoring.
2. Logan, G. D. & Crump, M. J. Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science* **330**, 683–686 (2010).
3. Wessel, J. R. An adaptive orienting theory of error processing. *Psychophysiology* **55**, e13041 (2018).
4. Ullsperger, M., Danielmeier, C. & Jocham, G. Neurophysiology of performance monitoring and adaptive behavior. *Physiol. Rev.* **94**, 35–79 (2014).
5. Ullsperger, M. Performance monitoring in neurological and psychiatric patients. *Int. J. Psychophysiol.* **59**, 59–69 (2006).
6. Gillan, C. M., Fineberg, N. A. & Robbins, T. W. A trans-diagnostic perspective on obsessive-compulsive disorder. *Psychol. Med.* **47**, 1528–1548 (2017).
7. Insel, T. R. The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. *Am. J. Psychiatry* **171**, 395–397 (2014).
8. Holroyd, C. B. & Umemoto, A. The research domain criteria framework: the case for anterior cingulate cortex. *Neurosci. Biobehav. Rev.* **71**, 418–443 (2016).
9. Loosen, A. M. & Hauser, T. U. Towards a computational psychiatry of juvenile obsessive-compulsive disorder. *Neurosci. Biobehav. Rev.* **118**, 631–642 (2020).
10. Nestler, E. J. & Hyman, S. E. Animal models of neuropsychiatric disorders. *Nat. Neurosci.* **13**, 1161–1169 (2010).
11. Robble, M. A. et al. Concordant neurophysiological signatures of cognitive control in humans and rats. *Neuropsychopharmacology* **46**, 252–1262 (2021).
12. Fu, Z. et al. Single-neuron correlates of error monitoring and post-error adjustments in human medial frontal cortex. *Neuron* **101**, 165–177.e165 (2019).
Shows that single neurons in the human MFC that signal errors respond first in the pre-SMA and later in the dorsal ACC, and in both areas the amplitude of the ERN predicts firing rates of error neurons.
13. Sajad, A., Godlove, D. C. & Schall, J. D. Cortical microcircuitry of performance monitoring. *Nat. Neurosci.* **22**, 265–274 (2019).
Describes the organization of error-related neurons across the layers of a medial frontal area and their association with the ERN.
14. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016).

15. Ridderinkhof, K. R., Ullsperger, M., Crone, E. A. & Nieuwenhuis, S. The role of the medial frontal cortex in cognitive control. *Science* **306**, 443–447 (2004).
16. Scangos, K. W., Aronberg, R. & Stuphorn, V. Performance monitoring by presupplementary and supplementary motor area during an arm movement countermanding task. *J. Neurophysiol.* **109**, 1928–1939 (2013).
Demonstrates that the pre-SMA and the SMA in macaques do not enable reactive response inhibition but instead signal errors.
17. Nachev, P., Kennard, C. & Husain, M. Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* **9**, 856–869 (2008).
18. Fu, Z. et al. The geometry of domain-general performance monitoring in the human medial frontal cortex. *Science* **376**, eabm9922 (2022).
Reveals that neurons in the human MFC encode conflict probability, cognitive conflict and errors in a domain-general and compositional manner across two tasks, thereby revealing multiple signals needed by our proposed model.
19. Ito, S., Stuphorn, V., Brown, J. W. & Schall, J. D. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* **302**, 120–122 (2003).
Demonstrates error-related neural spiking in the cingulate cortex of macaques during the stop-signal task.
20. Procyk, E. et al. Midcingulate motor map and feedback detection: converging data from humans and monkeys. *Cereb. Cortex* **26**, 467–476 (2016).
21. Vogt, B. A. Midcingulate cortex: structure, connections, homologies, functions and diseases. *J. Chem. Neuroanat.* **74**, 28–46 (2016).
22. Neubert, F. X., Mars, R. B., Sallet, J. & Rushworth, M. F. Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl Acad. Sci. USA* **112**, e2695–e2704 (2015).
23. Shackman, A. J. et al. The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* **12**, 154–167 (2011).
24. Picard, N. & Strick, P. L. Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* **6**, 342–353 (1996).
25. Emeric, E. E. et al. Performance monitoring local field potentials in the medial frontal cortex of primates: anterior cingulate cortex. *J. Neurophysiol.* **99**, 759–772 (2008).
26. Emeric, E. E., Leslie, M., Pouget, P. & Schall, J. D. Performance monitoring local field potentials in the medial frontal cortex of primates: supplementary eye field. *J. Neurophysiol.* **104**, 1523–1537 (2010).
27. Godlove, D. C. et al. Event-related potentials elicited by errors during the stop-signal task. I. Macaque monkeys. *J. Neurosci.* **31**, 15640–15649 (2011).
Demonstration of the macaque homologue of the ERN.
28. Schall, J. D., Stuphorn, V. & Brown, J. W. Monitoring and control of action by the frontal lobes. *Neuron* **36**, 309–322 (2002).
29. Debener, S. et al. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* **25**, 11730–11737 (2005).
Shows that the amplitude of the scalp-measured ERN co-varies trial-by-trial with the fMRI BOLD signal in the MFC.
30. Iannaccone, R. et al. Conflict monitoring and error processing: new insights from simultaneous EEG-fMRI. *Neuroimage* **105**, 395–407 (2015).
31. Swick, D. & Turken, U. Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proc. Natl Acad. Sci. Usa.* **99**, 16354–16359 (2002).
32. Turken, A. U. & Swick, D. Response selection in the human anterior cingulate cortex. *Nat. Neurosci.* **2**, 920–924 (1999).
33. Stemmer, B., Segalowitz, S. J., Witzke, W. & Schonle, P. W. Error detection in patients with lesions to the medial prefrontal cortex: an ERP study. *Neuropsychologia* **42**, 118–130 (2004).
34. Minxha, J., Adolphs, R., Fusi, S., Mamelak, A. N. & Rutishauser, U. Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* **368**, eaba3313 (2020).
Reveals encoding of task-specific but modality-independent choice signals in the human MFC at the single-cell level.
35. Bonini, F. et al. Action monitoring and medial frontal cortex: leading role of supplementary motor area. *Science* **343**, 888–891 (2014).
36. Sheth, S. A. et al. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* **488**, 218–221 (2012).
37. Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R. & Eskandar, E. N. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat. Neurosci.* **7**, 1370–1375 (2004).
38. Kolling, N. et al. Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* **19**, 1280–1285 (2016).
A review of evidence supporting the MCC's role in decision making and, more broadly, acquiring action policies in a changing environment.
39. Ebitz, R. B. & Hayden, B. Y. Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nat. Neurosci.* **19**, 1278–1279 (2016).
40. Heilbronner, S. R. & Hayden, B. Y. Dorsal anterior cingulate cortex: a bottom-up view. *Annu. Rev. Neurosci.* **39**, 149–170 (2016).
41. Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E. & Donchin, E. The error-related negativity. *Perspect. Psychol. Sci.* **13**, 200–204 (2018).
42. Gehring, W. J., Liu, Y., Orr, J. M. & Carp, J. in *The Oxford Handbook of Event-Related Potential Components* (eds Kappenman, E. S. & Luck, S. J.) (Oxford University Press, 2012).
A comprehensive and authoritative review of the ERN in humans.
43. Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. EPIC Ninth International Conference on Event-Related Potentials of the Brain, Noordwijk, The Netherlands 28 May–3 June 1989. *Electroencephalogr. Clin. Neurophysiol.* **42** (Suppl. 1), 1–393 (1991).
44. Gehring, W. J., Coles, M. G., Meyer, D. E. & Donchin, E. The error-related negativity: an event-related brain potential accompanying errors. *Psychophysiology* **27**, S34 (1990).
45. Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* **78**, 447–455 (1991).
46. Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E. & Donchin, E. A neural system for error detection and compensation. *Psychol. Sci.* **4**, 385–390 (1993).
47. Miltner, W. H., Braun, C. H. & Coles, M. G. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *J. Cogn. Neurosci.* **9**, 788–798 (1997).
48. Holroyd, C. B. & Coles, M. G. H. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
A computational model of error detection which proposes that the ERN is generated by transmission of the error signal carried by dopaminergic inputs to the MFC due to a transient pause of dopamine release and consequently disinhibition of pyramidal neurons.
49. Holroyd, C. B., Pakzad-Vaezi, K. L. & Krigolson, O. E. The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* **45**, 688–697 (2008).
50. Dehaene, S., Posner, M. I. & Tucker, D. M. Localization of a neural system for error detection and compensation. *Psychol. Sci.* **5**, 303–305 (1994).
The first attempt to locate the current dipole producing the ERN.
51. Yucel, M. et al. Hemispheric and gender-related differences in the gross morphology of the anterior cingulate/paracingulate cortex in normal volunteers: an MRI morphometric study. *Cereb. Cortex* **11**, 17–25 (2001).
52. Vogt, B. A., Nimchinsky, E. A., Vogt, L. J. & Hof, P. R. Human cingulate cortex: surface features, flat maps, and cytoarchitecture. *J. Comp. Neurol.* **359**, 490–506 (1995).
Describes the composition of areas within the MFC of humans that may contribute to error monitoring.
53. Amiez, C., Wilson, C. R. E. & Procyk, E. Variations of cingulate sulcal organization and link with cognitive performance. *Sci. Rep.* **8**, 13988 (2018).
54. Huster, R. J. et al. Effects of anterior cingulate fissurization on cognitive control during Stroop interference. *Hum. Brain Mapp.* **30**, 1279–1289 (2009).
55. Herrera, B., Sajad, A., Woodman, G. F., Schall, J. D. & Riera, J. J. A minimal biophysical model of neocortical pyramidal cells: implications for frontal cortex microcircuitry and field potential generation. *J. Neurosci.* **40**, 8513–8529 (2020).
56. Halnes, G., Maki-Marttunen, T., Pettersen, K. H., Andreassen, O. A. & Einevoll, G. T. Ion diffusion may introduce spurious current sources in current-source density (CSD) analysis. *J. Neurophysiol.* **118**, 114–120 (2017).
57. Turner, E. C. et al. Distributions of cells and neurons across the cortical sheet in old world macaques. *Brain Behav. Evol.* **88**, 1–13 (2016).
58. Cohen, M. X. Where does EEG come from and what does it mean. *Trends Neurosci.* **40**, 208–218 (2017).
59. Cole, M. W., Yeung, N., Freiwald, W. A. & Botvinick, M. Cingulate cortex: diverging data from humans and monkeys. *Trends Neurosci.* **32**, 566–574 (2009).
60. Cavanagh, J. F., Zambrano-Vazquez, L. & Allen, J. J. B. Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology* **49**, 220–238 (2012).
61. Yordanova, J., Falkenstein, M., Hohnsbein, J. & Kolev, V. Parallel systems of error processing in the brain. *Neuroimage* **22**, 590–602 (2004).
62. van Driel, J., Ridderinkhof, K. R. & Cohen, M. X. Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics. *J. Neurosci.* **32**, 16795–16806 (2012).
63. Luu, P., Tucker, D. M. & Makeig, S. Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clin. Neurophysiol.* **115**, 1821–1835 (2004).
64. Trujillo, L. T. & Allen, J. J. B. Theta EEG dynamics of the error-related negativity. *Clin. Neurophysiol.* **118**, 645–668 (2007).
65. Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414–421 (2014).
66. Hewig, J., Coles, M. G., Trippe, R. H., Hecht, H. & Miltner, W. H. Dissociation of Pe and ERN/Ne in the conscious recognition of an error. *Psychophysiology* **48**, 1390–1396 (2011).
67. Di Gregorio, F., Maier, M. E. & Steinhauser, M. Errors can elicit an error positivity in the absence of an error negativity: evidence for independent systems of human error monitoring. *NeuroImage* **172**, 427–436 (2018).
68. Wessel, J. R. Error awareness and the error-related negativity: evaluating the first decade of evidence. *Front. Hum. Neurosci.* **6**, 88 (2012).
69. Stuphorn, V., Taylor, T. L. & Schall, J. D. Performance monitoring by the supplementary eye field. *Nature* **408**, 857–860 (2000).
70. Holroyd, C. B., Dien, J. & Coles, M. G. H. Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neurosci. Lett.* **242**, 65–68 (1998).
71. Reinhart, R. M., Carlisle, N. B., Kang, M. S. & Woodman, G. F. Event-related potentials elicited by errors during the stop-signal task. II: human effector-specific error responses. *J. Neurophysiol.* **107**, 2794–2807 (2012).

72. Phillips, J. M. & Everling, S. Event-related potentials associated with performance monitoring in non-human primates. *Neuroimage* **97**, 308–320 (2014).
73. Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K. & Halgren, E. Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *J. Neurosci.* **25**, 604–613 (2005).
74. Yeung, N., Botvinick, M. M. & Cohen, J. D. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* **111**, 931–959 (2004).
75. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
76. Smith, E. H. et al. Widespread temporal coding of cognitive control in the human prefrontal cortex. *Nat. Neurosci.* **22**, 1883–1891 (2019).
77. Scangos, K. W. & Stuphorn, V. Medial frontal cortex motivates but does not control movement initiation in the countermanding task. *J. Neurosci.* **30**, 1968–1982 (2010).
78. Sajad, A., Errington, S. P. & Schall, J. D. Functional architecture of executive control and associated event-related potentials in macaques. *Nat. Commun.* **13**, 6279 (2022).
79. Nakamura K., Roesch M. R. & Olson C. R. Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J. Neurophysiol.* **93**, 884–908 (2005).
80. Ebitz, R. B. & Platt, M. L. Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal. *Neuron* **85**, 628–640 (2015).
81. Shenhav, A. & Botvinick, M. Uncovering a missing link in anterior cingulate research. *Neuron* **85**, 455–457 (2015).
82. Amiez, C. et al. Sulcal organization in the medial frontal cortex provides insights into primate brain evolution. *Nat. Commun.* **10**, 3437 (2019).
Analysis of the variable presence of a PCS in humans and apes and absence in monkeys.
83. Sallet, J. et al. The organization of dorsal frontal cortex in humans and macaques. *J. Neurosci.* **33**, 12255–12274 (2013).
84. Schall, J. D. et al. in *Evolutionary Neuroscience* 2nd edn (ed Kaas, J. H.) 861–890 (Academic Press, 2020).
85. Grosbras, M. H., Lobel, E., Van de Moortele, P. F., LeBihan, D. & Berthoz, A. An anatomical landmark for the supplementary eye fields in human revealed with functional magnetic resonance imaging. *Cereb. Cortex* **9**, 705–711 (1999).
86. Amiez, C. et al. The location of feedback-related activity in the midcingulate cortex is predicted by local morphology. *J. Neurosci.* **33**, 2217–2228 (2013).
87. Amiez, C. et al. Chimpanzee histology and functional brain imaging show that the paracingulate sulcus is not human-specific. *Commun. Biol.* **4**, 54 (2021).
88. Buda, M., Fornito, A., Bergstrom, Z. M. & Simons, J. S. A specific brain structural basis for individual differences in reality monitoring. *J. Neurosci.* **31**, 14308–14313 (2011).
89. Simons, J. S., Garrison, J. R. & Johnson, M. K. Brain mechanisms of reality monitoring. *Trends Cogn. Sci.* **21**, 462–473 (2017).
90. Huster, R. J., Enriquez-Geppert, S., Pantev, C. & Bruchmann, M. Variations in midcingulate morphology are related to ERP indices of cognitive control. *Brain Struct. Funct.* **219**, 49–60 (2014).
91. Garrison, J. R. et al. Paracingulate sulcus morphology is associated with hallucinations in the human brain. *Nat. Commun.* **6**, 8956 (2015).
92. Shim, G. et al. Reduced cortical folding of the anterior cingulate cortex in obsessive-compulsive disorder. *J. Psychiatry Neurosci.* **34**, 443–449 (2009).
93. Nimchinsky, E. A., Vogt, B. A., Morrison, J. H. & Hof, P. R. Spindle neurons of the human anterior cingulate cortex. *J. Comp. Neurol.* **355**, 27–37 (1995).
94. Barbas, H. & Pandya, D. N. Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* **286**, 353–375 (1989).
95. Vogt, B. A., Vogt, L., Farber, N. B. & Bush, G. Architecture and neurocytology of monkey cingulate gyrus. *J. Comp. Neurol.* **485**, 218–239 (2005).
96. Matelli, M., Luppino, G. & Rizzolatti, G. Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *J. Comp. Neurol.* **311**, 445–462 (1991).
97. Petrides, M. Comparative architectonic analysis of the human and the macaque frontal cortex. *Handb. Neuropsychol.* **11**, 17–58 (1994).
98. Paxton, J. L., Barch, D. M., Racine, C. A. & Braver, T. S. Cognitive control, goal maintenance, and prefrontal function in healthy aging. *Cereb. Cortex* **18**, 1010–1028 (2008).
99. Danielmeier, C. & Ullsperger, M. Post-error adjustments. *Front. Psychol.* **2**, 233 (2011).
100. Verbruggen, F. et al. A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *Elife* **8**, e46323 (2019).
101. Camalier, C. R. et al. Dynamics of saccade target selection: race model analysis of double step and search step saccade production in human and macaque. *Vis. Res.* **47**, 2187–2211 (2007).
102. Schall, J. D. & Boucher, L. Executive control of gaze by the frontal lobes. *Cogn. Affect. Behav. Neurosci.* **7**, 396–412 (2007).
103. Mosher, C. P., Mamelak, A. N., Malekmohammadi, M., Pouratian, N. & Rutishauser, U. Distinct roles of dorsal and ventral subthalamic neurons in action selection and cancellation. *Neuron* **109**, 869–881 e866 (2021).
104. Coull, J. T., Cheng, R. K. & Meck, W. H. Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* **36**, 3–25 (2011).
105. Logan, G. D., Van Zandt, T., Verbruggen, F. & Wagenmakers, E. J. On the ability to inhibit thought and action: general and special theories of an act of control. *Psychol. Rev.* **121**, 66–95 (2014).
106. Stuphorn, V. & Schall, J. D. Executive control of countermanding saccades by the supplementary eye field. *Nat. Neurosci.* **9**, 925–931 (2006).
107. Menon, V., Adleman, N. E., White, C. D., Glover, G. H. & Reiss, A. L. Error-related brain activation during a Go/NoGo response inhibition task. *Hum. Brain Mapp.* **12**, 131–143 (2001).
108. Rivaud-Pechoux, S., Vidailhet, M., Brandel, J. P. & Gaymard, B. Mixing pro- and antisaccades in patients with parkinsonian syndromes. *Brain* **130**, 256–264 (2007).
109. Simon, J. R. in *Advances in Psychology* Vol. 65 (eds Proctor, R. W. & Reeve, T. G.) 31–86 (Elsevier, 1990).
110. MacLeod, C. M. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* **109**, 163–203 (1991).
111. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).
112. Bush, G., Shin, L. M., Holmes, J., Rosen, B. R. & Vogt, B. A. The Multi-Source Interference Task: validation study with fMRI in individual subjects. *Mol. Psychiatry* **8**, 60–70 (2003).
113. Mostofsky, S. H. & Simmonds, D. J. Response inhibition and response selection: two sides of the same coin. *J. Cogn. Neurosci.* **20**, 751–761 (2008).
114. Washburn, D. A. The Stroop effect at 80: the competition between stimulus control and cognitive control. *J. Exp. Anal. Behav.* **105**, 3–13 (2016).
115. Lauwereyns, J. et al. Interference from irrelevant features on visual discrimination by macaques (*Macaca fuscata*): a behavioral analogue of the human Stroop effect. *J. Exp. Psychol. Anim. Behav. Process.* **26**, 352–357 (2000).
116. Michelet, T. et al. Electrophysiological correlates of a versatile executive control system in the monkey anterior cingulate cortex. *Cereb. Cortex* **26**, 1684–1697 (2016).
117. Bastos, A. M. et al. Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
118. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
119. Bhushan, N. & Shadmehr, R. Computational nature of human adaptive control during learning of reaching movements in force fields. *Biol. Cybern.* **81**, 39–60 (1999).
120. McNamee, D. & Wolpert, D. M. Internal models in biological control. *Annu. Rev. Control. Robot. Auton. Syst.* **2**, 339–364 (2019).
121. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
122. Reppert, T. R., Heitz, R. P. & Schall, J. D. Neural mechanisms for executive control of speed-accuracy tradeoff. Preprint at *bioRxiv* https://doi.org/10.1101/773549 (2019).
123. Dutilh, G. et al. Testing theories of post-error slowing. *Atten. Percept. Psychophys.* **74**, 454–465 (2012).
124. Laming, D. Choice reaction performance following an error. *Acta Psychol.* **43**, 199–224 (1979).
125. Verguts, T., Notebaert, W., Kunde, W. & Wuhr, P. Post-conflict slowing: cognitive adaptation after conflict processing. *Psychon. Bull. Rev.* **18**, 76–82 (2011).
126. Egner, T. Congruency sequence effects and cognitive control. *Cogn. Affect. Behav. Neurosci.* **7**, 380–390 (2007).
127. Sohn, H., Narain, D., Meirhaeghe, N. & Jazayeri, M. Bayesian computation through cortical latent dynamics. *Neuron* **103**, 934–947.e5 (2019).
128. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).
129. Frank, M. J., Samanta, J., Moustafa, A. A. & Sherman, S. J. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* **318**, 1309–1312 (2007).
130. Chen, W. et al. Prefrontal-subthalamic hyperdirect pathway modulates movement inhibition in humans. *Neuron* **106**, 579–588 e573 (2020).
131. Cavanagh, J. F. et al. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat. Neurosci.* **14**, 1462–1467 (2011).
132. Aron, A. R. & Poldrack, R. A. Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. *J. Neurosci.* **26**, 2424–2433 (2006).
133. Braver, T. S. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* **16**, 106–113 (2012).
An influential theoretical framework that divides cognitive control into a proactive mode and a reactive mode and proposes the differential roles of the PFC in subserving these two modes of cognitive control processes.
134. Kerns, J. G. et al. Anterior cingulate conflict monitoring and adjustments in control. *Science* **303**, 1023–1026 (2004).
135. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
Offers a comprehensive theory of the role of the ACC in cognitive control.
136. Brown, J. W. & Braver, T. S. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005).
137. Jiang, J., Beck, J., Heller, K. & Egner, T. An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nat. Commun.* **6**, 8165 (2015).
138. Isoda, M. & Hikosaka, O. Switching from automatic to controlled action by monkey medial frontal cortex. *Nat. Neurosci.* **10**, 240–248 (2007).
139. Isoda, M. & Hikosaka, O. Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. *J. Neurosci.* **28**, 7209–7218 (2008).
140. Aquino, T. G., Cockburn, J., Mamelak, A. N., Rutishauser, U. & O'Doherty, J. P. Neurons in human pre-supplementary motor area encode key computations for value-based choice. *bioRxiv* https://doi.org/10.1101/2021.10.27.466000 (2021).
141. Sarafyazd, M. & Jazayeri, M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911 (2019).
Provides neuronal evidence that the macaque MFC is responsible for hierarchical reasoning about an error (that is, determining whether a loss of reward is caused by perceptual errors or a change in the response rule).
142. Maia, T. V. & Frank, M. J. From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* **14**, 154–162 (2011).

143. Klaus, A., Alves da Silva, J. & Costa, R. M. What, if, and when to move: basal ganglia circuits and self-paced action initiation. *Annu. Rev. Neurosci.* **42**, 459–483 (2019).
144. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol. Cybern.* **84**, 401–410 (2001).
145. Wiecki, T. V. & Frank, M. J. A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychol. Rev.* **120**, 329–355 (2013).
146. Hatanaka, N. et al. Thalamocortical and intracortical connections of monkey cingulate motor areas. *J. Comp. Neurol.* **462**, 121–138 (2003).
147. Godlove, D. C., Maier, A., Woodman, G. F. & Schall, J. D. Microcircuitry of agranular frontal cortex: testing the generality of the canonical cortical microcircuit. *J. Neurosci.* **34**, 5355–5369 (2014).
148. Ninomiya, T., Dougherty, K., Godlove, D. C., Schall, J. D. & Maier, A. Microcircuitry of agranular frontal cortex: contrasting laminar connectivity between occipital and frontal areas. *J. Neurophysiol.* **113**, 3242–3255 (2015).
149. Douglas, R. J. & Martin, K. A. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* **27**, 419–451 (2004).
150. Larkum, M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151 (2013).
151. Gidon, A. et al. Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science* **367**, 83–87 (2020).
152. Rutishauser, U., Slotine, J. J. & Douglas, R. Computation in dynamically bounded asymmetric systems. *PLoS Comput. Biol.* **11**, e1004039 (2015).
153. Rutishauser, U., Douglas, R. J. & Slotine, J. J. Collective stability of networks of winner-take-all circuits. *Neural Comput.* **23**, 735–773 (2011).
154. Seifert, S., von Cramon, D. Y., Imperati, D., Tittgemeyer, M. & Ullsperger, M. Thalamocingulate interactions in performance monitoring. *J. Neurosci.* **31**, 3375–3383 (2011).
155. Hoesen, G. W. V., Morecraft, R. J. & Vogt, B. A. in *Neurobiology of Cingulate Cortex and Limbic Thalamus* (eds Vogt, B. A. & Gabriel, M.) 249–284 (Springer, 1993).
156. Matelli, M. & Luppino, G. Thalamic input to mesial and superior area 6 in the macaque monkey. *J. Comp. Neurol.* **372**, 59–87 (1996).
157. Kunimatsu, J. & Tanaka, M. Roles of the primate motor thalamus in the generation of antisaccades. *J. Neurosci.* **30**, 5108–5117 (2010).
158. Peterburs, J. et al. Altered error processing following vascular thalamic damage: evidence from an antisaccade task. *PloS ONE* **6**, e21517 (2011).
159. Allain, S., Hasbroucq, T., Burle, B., Grapperon, J. & Vidal, F. Response monitoring without sensory feedback. *Clin. Neurophysiol.* **115**, 2014–2020 (2004).
160. Amiez, C., Champod, A. S., Wilson, C. R., Procyk, E. & Petrides, M. A unilateral medial frontal cortical lesion impairs trial and error learning without visual control. *Neuropsychologia* **75**, 314–321 (2015).
161. Nachev, P., Wydell, H., O'Neill, K., Husain, M. & Kennard, C. The role of the pre-supplementary motor area in the control of action. *Neuroimage* **36**, T155–T163 (2007).
162. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
163. Krigolson, O. E. & Holroyd, C. B. Hierarchical error processing: different errors, different systems. *Brain Res.* **1155**, 70–80 (2007).
164. Haber, S. N. The place of dopamine in the cortico-basal ganglia circuit. *Neuroscience* **282**, 248–257 (2014).
165. Aarts, E., Roelofs, A. & van Turennout, M. Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *J. Neurosci.* **28**, 4671–4678 (2008).
166. Heitz, R. P. & Schall, J. D. Neural mechanisms of speed-accuracy tradeoff. *Neuron* **76**, 616–628 (2012).
167. Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. *Psychophysiol. Brain Res.* **1**, 192–195 (1990).
168. Brooks, V. B. How does the limbic system assist motor learning? A limbic comparator hypothesis (part 1 of 2). *Brain Behav. Evol.* **29**, 29–41 (1986).
169. Wang, J. X. et al. Prefrontal cortex as a meta-reinforcement learning system. *Nat. Neurosci.* **21**, 860–868 (2018).
170. Jeurissen, D., Shushruth, S., El-Shamayleh, Y., Horwitz, G. D. & Shadlen, M. N. Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. *Neuron* **110**, 1924–1931.e1925 (2022).
171. Ogasawara, T., Nejime, M., Takada, M. & Matsumoto, M. Primate nigrostriatal dopamine system regulates saccadic response inhibition. *Neuron* **100**, 1513–1526.e1514 (2018).
172. Grace, A. A. & Bunney, B. S. Nigral dopamine neurons - intracellular-recording and identification with L-dopa injection and histofluorescence. *Science* **210**, 654–656 (1980).
173. Guyenet, P. G. & Aghajanian, G. K. Antidromic identification of dopaminergic and other output neurons of the rat substantia nigra. *Brain Res.* **150**, 69–84 (1978).
174. Thierry, A. M., Deniau, J. M., Herve, D. & Chevalier, G. Electro-physiological evidence for non-dopaminergic mesocortical and mesolimbic neurons in the rat. *Brain Res.* **201**, 210–214 (1980).
175. Lavin, A. et al. Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. *J. Neurosci.* **25**, 5013–5023 (2005).
176. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14**, 1338–1344 (2011).
A computational model that provides a unifying account of a wide range of neural signals in the MFC.
177. Silvetti, M., Seurinck, R. & Verguts, T. Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* **5**, 75 (2011).
178. Alexander, W. H. & Brown, J. W. Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Comput.* **27**, 2354–2410 (2015).
179. Shahnazian, D. & Holroyd, C. B. Distributed representations of action sequences in anterior cingulate cortex: a recurrent neural network approach. *Psychon. Bull. Rev.* **25**, 302–321 (2018).
180. Vassena, E., Holroyd, C. B. & Alexander, W. H. Computational models of anterior cingulate cortex: at the crossroads between prediction and effort. *Front. Neurosci.* **11**, 316 (2017).
181. Huang, Y. & Rao, R. P. N. Predictive coding. *Wiley Interdiscip. Rev. Cogn. Sci.* **2**, 580–593 (2011).
182. Riesel, A. The erring brain: error-related negativity as an endophenotype for OCD — a review and meta-analysis. *Psychophysiology* **56**, e13348 (2019).
183. Foti, D., Kotov, R., Bromet, E. & Hajcak, G. Beyond the broken error-related negativity: functional and diagnostic correlates of error processing in psychosis. *Biol. Psychiatry* **71**, 864–872 (2012).
184. Luijten, M. et al. Systematic review of ERP and fMRI studies investigating inhibitory control and error processing in people with substance dependence and behavioural addictions. *J. Psychiatry Neurosci.* **39**, 149–169 (2014).
185. Manoach, D. S. & Agam, Y. Neural markers of errors as endophenotypes in neuropsychiatric disorders. *Front. Hum. Neurosci.* **7**, 350 (2013).
186. Kirschner, H. & Klein, T. A. Beyond a blunted ERN - biobehavioral correlates of performance monitoring in schizophrenia. *Neurosci. Biobehav. Rev.* **133**, 104504 (2021).
187. Endrass, T. & Ullsperger, M. Specificity of performance monitoring changes in obsessive-compulsive disorder. *Neurosci. Biobehav. Rev.* **46**, 124–138 (2014).
188. Tsai, L. L., Young, H. Y., Hsieh, S. & Lee, C. S. Impairment of error monitoring following sleep deprivation. *Sleep* **28**, 707–713 (2005).
189. Ridderinkhof, K. R. et al. Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science* **298**, 2209–2211 (2002).
190. Franken, I. H., van Strien, J. W., Franzek, E. J. & van de Wetering, B. J. Error-processing deficits in patients with cocaine dependence. *Biol. Psychol.* **75**, 45–51 (2007).
191. Hajcak, G., Klawohn, J. & Meyer, A. The utility of event-related potentials in clinical psychology. *Annu. Rev. Clin. Psychol.* **15**, 71–95 (2019).
192. Bakken, T. E. et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111–119 (2021).
193. Huys, Q. J., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* **19**, 404–413 (2016).

Acknowledgements
The authors thank R. Adolphs, K. Silm and J. Brown for discussion and C. Holroyd for valuable comments. J.D.S. has been supported by the US National Institutes of Health and by the E. Bronson Ingram Chair in Neuroscience and is currently supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-04592). U.R. has been supported by the US National Institutes of Health (R01MH110831 and U01NS117839) and the NSF (BCS-2219800). A.S. has been supported by a Canadian Institutes of Health Research Postdoctoral Fellowship award.

Author contributions
The authors all contributed to all aspects of preparing the article.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41583-022-00670-w.

Correspondence should be addressed to Zhongzheng Fu, Jeffrey D. Schall or Ueli Rutishauser.

Peer review information *Nature Reviews Neuroscience* thanks M. Ullsperger, C. Carter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2023