Surprise, value and control in anterior cingulate cortex during speeded decision-making

Vassena, E.; Deraeve, J.; Alexander, William H.

2020, Article / Letter to editor (Nature Human Behaviour, 4, 4, (2020), pp. 412-422)

Doi link to publisher: https://doi.org/10.1038/s41562-019-0801-5

ARTICLES
https://doi.org/10.1038/s41562-019-0801-5
nature human behaviour

Surprise, value and control in anterior cingulate cortex during speeded decision-making

Eliana Vassena[1,2], James Deraeve[2] and William H. Alexander[2,3*]

Activity in the dorsal anterior cingulate cortex (dACC) is observed across a variety of contexts, and its function remains intensely debated in the field of cognitive neuroscience. While traditional views emphasize its role in inhibitory control (suppressing prepotent, incorrect actions), recent proposals suggest a more active role in motivated control (invigorating actions to obtain rewards). Lagging behind empirical findings, formal models of dACC function primarily focus on inhibitory control, highlighting surprise, choice difficulty and value of control as key computations. Although successful in explaining dACC involvement in inhibitory control, it remains unclear whether these mechanisms generalize to motivated control. In this study, we derive predictions from three prominent accounts of dACC and test these with functional magnetic resonance imaging during value-based decision-making under time pressure. We find that the single mechanism of surprise best accounts for activity in dACC during a task requiring response invigoration, suggesting surprise signalling as a shared driver of inhibitory and motivated control.

Activity in the dorsal anterior cingulate cortex (dACC) and surrounding regions in the medial prefrontal cortex (mPFC) is routinely observed in neuroimaging studies of cognitive control and decision-making[1]. Consequently, a number of theoretical and computational accounts have been developed in the past two decades to describe the role and function of dACC in cognitive control[2]. Generally, cognitive control entails the need to suppress an incorrect, prepotent response to generate a correct, but less automatic, response. Early computational models of dACC function therefore principally addressed tasks thought to involve response selection and inhibition, assigning the region roles in signalling behavioural error, detecting and resolving response conflict[3], selecting appropriate motor responses[4] or predicting the likelihood of an incorrect response[5].

Although cognitive control research has traditionally focused on response inhibition, recent work has highlighted the interactions between control and motivation[2,6]. Motivation refers to the drive to pursue specific behavioural goals to obtain desired outcomes (such as rewards (extrinsic motivation) or other states perceived as rewarding (intrinsic motivation)). This line of work defines motivation as the "invigorating impact, on both behaviour and cognition, of prospective reward (both extrinsic reward such as money and instrinsic reward tied to the satisfaction of self-relevant behavioural goals"[7]. Here, we refer to motivated control as the process promoting successful selection and invigoration of a behavioural response leading to a valuable desired outcome. In this view, exerting control is costly, but also valuable, as it allows the securing of a prospective reward[8,9]. Individuals consistently tend to avoid exerting mental effort when possible[10], and preparing for a cognitively effort-demanding task is associated with increased dACC activity[11,12]. Interestingly, activity in the same region appears to correlate with the expectation of higher reward following task completion[13]. The overlap of cognitive effort and reward signals within dACC has led to development of new accounts of the region's function, assigning it a role in computing benefits and costs of actions, and integrating these in a 'net-value' driving adaptive behavioural selection in situations involving exertion of cognitive control or physical effort[9]. In line with these findings, an influential theoretical framework has been proposed, the expected value of control account (EVC; ref. [14]). The EVC posits that activity in dACC reflects 'expected value of control'—a trade-off between cost and benefits resulting in the selection of an optimal control signal.

In parallel, a growing amount of evidence supports a key role of the dACC in tracking the likelihood of events (such as responses and outcomes given a certain stimulus), and computing the discrepancy between predicted and actual events (that is, prediction errors), formalized in the predicted response outcome model (PRO; ref. [15]), which posits that the dACC signal reflects an ongoing comparison between expected and observed events: any unexpected, and therefore 'surprising', event will produce increased activity in dACC, and this signal contributes to updating of future predictions. This type of surprise could be termed epistemic control: dACC activity reflects predictive signals and error signals when predictions are not met. These signals may trigger behavioural adaptation when necessary[16–20]. This line of work successfully explains classical inhibitory control effects as a function of likelihood of responses and outcomes (errors, incongruent options and non-prepotent responses are generally less likely, and therefore surprising). More recently, this approach has also been applied to motivated control, suggesting that motivationally relevant variables (for example, effort requirements or potential reward amounts) may be monitored in a similar fashion. In this framework, deciding to engage in effortful behaviour is generally less likely and therefore the choice to engage is associated with greater dACC activity (PRO–effort)[21]. Even in the absence of subsequent invigoration, the choice itself to accept more effortful tasks is infrequent (humans are generally effort-avoidant), and would therefore elicit increased dACC activity (which, in epistemic control terms, reflects the likelihood of engaging in a task given its specific effort and reward properties). The PRO–effort proposal outlines how a likelihood-monitoring account of dACC function holds potential for generalization to motivated invigoration of behaviour. However, this account is yet to be tested against the wide array of effort-related effects on behaviour and brain activity in the mPFC–dACC, especially considering that previous work

[1]Donders Institute for Brain Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands. [2]Department of Experimental Psychology, Ghent University, Ghent, Belgium. [3]Center for Complex Systems & Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA. *e-mail: walexander@fau.edu

implicated dissociable mPFC–dACC sub-regions in different components of effortful behaviour[22].

Finally, a third perspective on dACC function arises from the value-based decision-making literature. Activity in this region increases when making a choice between two options is difficult—both options entail a similar value (choice difficulty (CD) account[23]). This account extends the classical response conflict theory of cognitive control idea[3] to the decision domain: when a choice needs to be made under time pressure, and both available options have similar value, control is required to ensure that one of the two options is selected, thereby 'resolving the conflict'. Selection between similarly valued options poses a more challenging problem than options with clear value differences, and dACC activity appears to reflect this difficulty. Recent work investigating the relationship between electrophysiological markers of conflict and surprise in a decision-making context (mid-frontal theta power, putatively related to dACC activity), indeed shows increased dACC activity for both more conflicted and surprising decisions[24].

While the theoretical frameworks described above capture several behavioural and neuroimaging effects observed during inhibitory control experiments, it has yet to be shown that their predictions apply in the context of motivated control—that is, situations in which a response requires selective invigoration. Motivated control is frequently confounded with value-driven responses[6], making it difficult to disentangle behaviour arising from prepotent, value-based processes versus behaviour that is selectively invigorated to realize a prospective reward. Although value-based choice may not require explicit control requirements due to the increased incentive or salience of valuable options, it has been demonstrated that the addition of time pressure to such tasks produces effects consistent with cognitive control, including changes in the speed–accuracy trade-off[5] and adjustments in decision parameters underlying choice behaviour[25,26]. Specifically, increases in time pressure are correlated with decreases in the decision boundary[27] of a diffusion process, a variable that is associated with the exertion of control[28] and distinct from other factors, such as drift rate, that may influence response without recruiting control.

In this study, we test predictions of three prominent models of cognitive control using a speeded value-based decision-making task in which the need to invigorate a response is dissociated from reward incentive, and can thus distinguish motivated control from value-driven responding. Model predictions are tested using functional magnetic resonance imaging (fMRI) data recorded from human subjects performing the speeded value-based decision-making task. In each trial, participants had a limited amount of time (750 ms) to choose between options presented on the left and right sides of the screen; each option was composed of two fractal images, displayed one above the other, and each image was associated with a fixed number of points that could be earned. The number of points for an image ranged from 10 to 80, in increments of 10 (for a total of 8 unique images), and subjects were trained in the association between each image and the points associated with it before scanning (Supplementary Methods). Once an option was selected, one of the two images was randomly selected and the participants earned the number of points associated with that image. Participants were instructed to maximize their earnings, and informed that points earned during the experiment would be converted to money at the end. The fractal images for each option were chosen at random from the set of eight images on each trial—in some trials (hard) the expected value of each option could be similar, while in other trials (easy) there was a large value difference in the expected value of each option. As in typical cognitive control experiments, we anticipated that hard trials would require the deployment of additional control processes relative to easy trials. Specifically, we expected that hard trials would require motivated control to generate a response within the response deadline.

Results

Theoretical predictions from three competing accounts. We derived predictions from the three models of dACC described above using published descriptions of the models and the conditions to which they apply (see Methods). Under the CD model, the dACC should track the similarity of options: activity should be maximal for options with similar values, and decrease as the value difference increases. Predictions of the EVC model were derived in two ways. Using only the equations presented in ref. [14], the EVC model (EVC1) predicts that the optimal control signal should be opposite to the pattern predicted by the CD model: dACC activity should be minimal when options are similar (exerting control is not effective, as both options are equally valuable), and increase with increasing value difference between options (reflecting increased value for exerting control as the value of one option increases). If we include the idea (described but not formalized in ref. [14]) that, as decisions become easier due to large value differences, the 'added value' of additional control decreases, the EVC model (EVC2) predicts an M-shaped pattern. The optimal control signal for similar options is low (because in EVC1, additional control does not increase the value of outcomes), as is the control signal for extremely different values (no control is required to make a choice). Finally, the PRO model predicts a W-shaped pattern in dACC activity, derived from the negative component of prediction error that is negative surprise[15]. Population equivalents are observed throughout the brain when actual events differ from expected or intended events. For options with similar values, evidence for choosing either option is approximately equal (that is, a person may equally intend to select either option) and therefore either choice entails surprise related to the unmet partial intention to select the alternate option. Surprise signals also apply at the onset of the trials when options are revealed: in trials with extreme differences in the value between options, surprise derives from the unexpected deviation of option values from their long-run average (that is, one option is better or worse then the average). The W-pattern predicted by the PRO model thus derives from standard population equivalent calculations following the presentation of options and generation of a response (Supplementary Fig. 3 and Methods).

A convenient way to summarize these competing predictions is as polynomial equations (Fig. 1a), which transcend model-specific implementation details. The predictions of both CD and EVC1 can be characterized as quadratic polynomials[16,23,29], the difference between the two being the sign on the coefficient for the quadratic term (U-shaped for EVC1, inverted U for CD). The predictions of the PRO model and EVC2 are best characterized by a quartic (fourth-order) polynomial, also with opposite signs on the coefficient of the quadratic term (negative for the PRO model, positive for EVC2). We can thus distinguish amongst the model predictions as follows (Fig. 1b). First, comparisons of Akaike's information criterion (AIC) values for fits of polynomial equations to neural data provide an estimate of whether the data are better explained by quadratic (CD and EVC1) or quartic (PRO and EVC2) curves. Second, the sign on the quadratic term of the best-fit polynomial equation distinguishes between remaining models (negative (CD and PRO) or positive (EVC1 and EVC2; Fig. 1b).

A second advantage of using polynomial equations is that they are not dependent on assumptions regarding the centre of the data. In value-based choice, individual subjects may be biased to prefer, for example, left over right responses, or an easier rather than a harder decision. During analysis, this bias emerges as a 'value bonus' for one type of response over another and requires explicit estimation of the magnitude of the bonus to compensate for subjective indifference between choices[23]. Since polynomial equations do not include assumptions about the source of bias as do cognitive models (including, for example, the CD model), the subjective centre of the distribution is simply another parameter to be fit. Additionally,

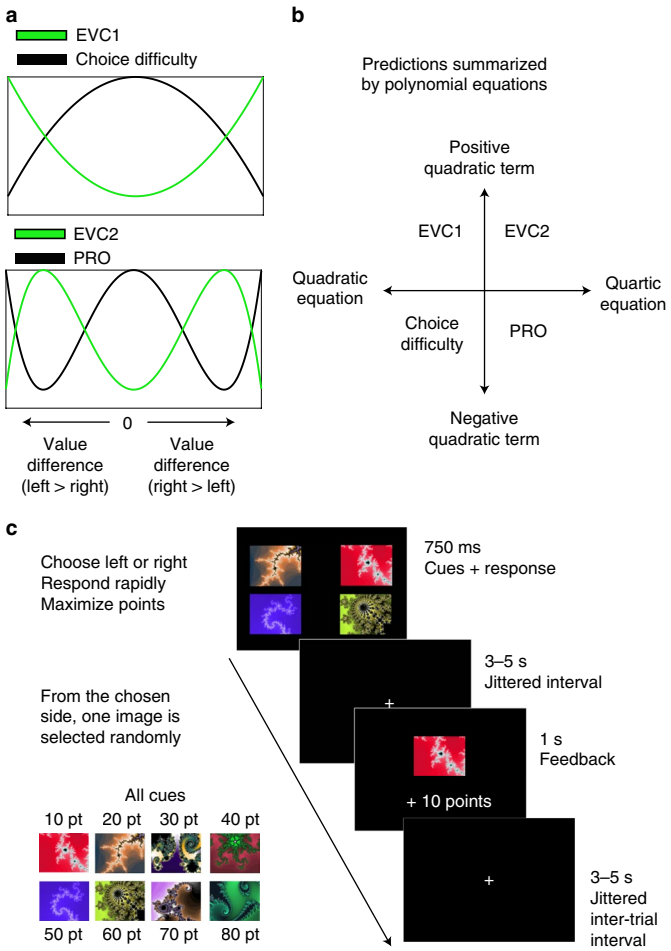


Fig. 1 | Theoretical predictions and experimental paradigm. a, Qualitative predictions of the models for dACC activity. CD and EVC1 predict a quadratic curve, PRO predicts a W-shaped curve, and EVC2 predicts an M-shaped curve. **b**, Summary of analysis rationale. Each model makes predictions that can be discriminated along two dimensions. **c**, The speeded decision-making task. Subjects choose between two options (left or right), with a 50% chance of receiving the points (pt) associated with each of the images for the selected option. Some of these decisions are easy (high-value difference between left and right sides) but some are hard (low-value difference between left and right sides).

although we do not predict strong linear or cubic effects, we have included linear and cubic polynomial functions in our model set to ensure that quartic and quadratic models with Akaike weights >0.999 are not due to significant linear or cubic terms rather than quartic or quadratic terms.

Behavioural strategies and markers of control. Predictions of the computational models of dACC were tested by having human subjects perform a speeded value-based decision task while undergoing fMRI (Fig. 1c). All models under consideration assumed that subjects make choices to maximize some measure of value; in deriving predictions of dACC activity (Fig. 1a) we assumed that subjects would maximize expected value—that is, the average value of possible outcomes for a given option. We additionally assumed that, besides behavioural error induced by control demands, value was the only factor driving choice behaviour. To test these assumptions, we performed three logistic regression analyses on subjects' choice (left vs right) behaviour. All logistic regressions contained a constant predictor variable indicating left or right bias that was not significantly different from 0 ($\beta = -0.21$, $t(21) = -1.32$, $P = 0.2$, odds ratio (OR) = 0.81, 95% confidence interval (CI) = 0.60–1.11).

In the first analysis (Fig. 2a), we tested whether subjects made choices consistent with maximizing expected value, with the minimum and maximum value of each option as predictor variables ($n = 4$). The results of this analysis suggest that our assumption that choice behaviour is dependent on expected value is not correct, although both minimum and maximum possible outcomes contribute to left-vs-right choices (left maximum option: $\beta = 0.96$, $t(21) = 5.11$, $P < 0.001$, OR = 2.61, 95% CI = 1.80–3.76; left minimum option: $\beta = 0.24$, $t(21) = 3.36$, $P = 0.003$, OR = 1.27, 95% CI = 1.11–1.47; right maximum option: $\beta = -0.91$, $t(21) = -7.86$, $P < 0.001$. OR = 0.40, 95% CI = 0.32–0.51; right minimum option: $\beta = -0.14$, $t(21) = -2.30$, $P = 0.032$, OR = 0.87, 95% CI = 0.77–0.98). Estimated β values for the maximum possible outcome are higher than for the minimum outcome for both left ($t(21) = 3.58$, $P = 0.002$, Cohen's $d = 0.76$, 95% CI = 0.30–1.13) and right choices ($t(21) = -5.04$, $P < 0.001$, Cohen's $d = 1.07$, 95% CI = −0.45 to −1.09)—that is, although subjects incorporated information from all possible outcomes of their choice, they appeared to weigh the maximum possible outcome of a choice more heavily.

In the second logistic regression analysis (Fig. 2b), we tested whether subject choices were informed by a tendency to both minimize risk and maximize value. Humans are known to be risk-averse, and tend to select options that either minimize the probability of losing (or receiving a low-value outcome) or minimize the outcome variance. In our regression analysis, we used the difference between the low and high outcomes for each option as predictor variables, as well as the expected value of each option, for a total of four predictor variables (expected value was used in place of the maximum and minimum values, since the former is collinear with the difference between maximum and minimum). We found (unsurprisingly) that the expected value of an option drove choices for that option (left expected value: $\beta = 1.06$, $t(21) = 5.98$, $P < 0.001$, OR = 2.89, 95% CI = 2.04–4.09; right expected value: $\beta = -0.92$, $t(21) = -10.04$, $P < 0.001$, OR = 0.40, 95% CI = 0.33–0.48). Additionally, the variance of the option was positively associated (left variance: $\beta = 0.34$, $t(21) = 3.72$, $P = 0.0013$, OR = 1.40, 95% CI = 1.17–1.68; right variance: $\beta = -0.37$, $t(21) = -4.81$, $P < 0.001$, OR = 0.69, 95% CI = 0.59–0.80) with choices for that option (Fig. 2b). This finding is consistent with a choice strategy that weighs the maximum possible outcome of an option more heavily than the minimum possible outcome, and inconsistent with a strategy that involves avoiding potential negative consequences or option variance. To verify that subjects were not engaged in risk-seeking behaviour, as suggested by the significant variance predictor results from the second logistic regression, a third logistic regression was performed in which two predictor variables were included for the maximum outcome of each option, as well as two predictor variables for the variance of each option (Fig. 2c). Here we found that the variance of an option had a slight negative influence on the probability of selecting that option (left variance: $\beta = -0.23$, $t(21) = -3.40$, $P = 0.002$, OR = 0.79, 95% CI = 0.70–0.91; right variance: $\beta = 0.14$, $t(21) = 2.26$, $P = 0.034$, OR = 1.14, 95% CI = 1.02–1.29)—that is, subjects were mildly risk-averse. The results of our logistic regression analyses suggest that subjects incorporated all available information in their choices, but weighed the maximum value more heavily. Simulations of the models (Supplementary Methods) using this strategy demonstrate that model-derived predictions of dACC activity (Fig. 1a) remain valid.

While the predictions of the PRO and CD models are not explicitly dependent on the engagement of cognitive control[23,30], those of the EVC model are contingent on the recruitment of such control processes[14]—that is, the predictions derived from the EVC model in our simulations are valid only if it can be demonstrated that our speeded decision-making task requires control. While previous statements by the authors of EVC specifically indicate that time

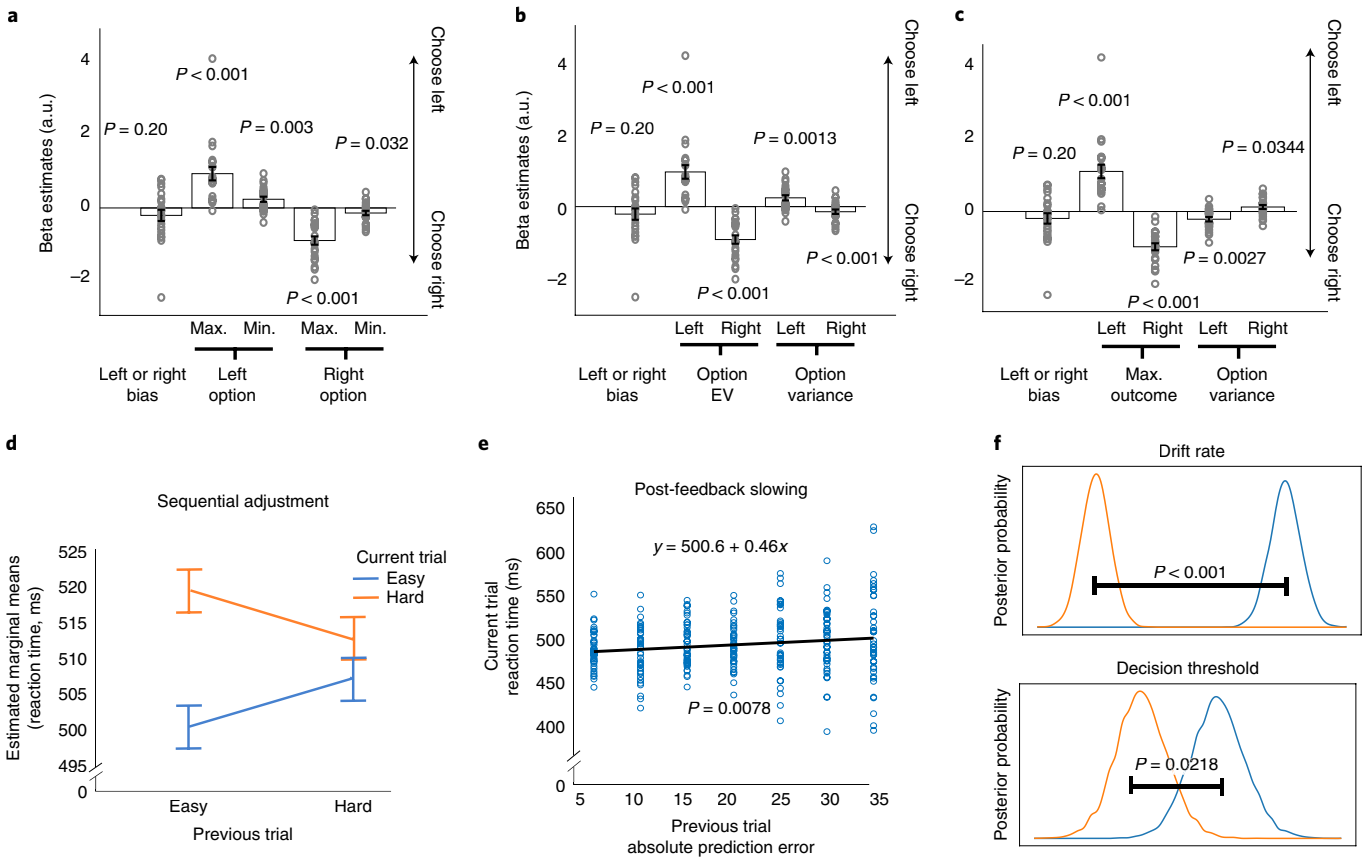


Fig. 2 | Behavioural strategies and reactive and proactive control processes. a–c, Logistic regression against left-vs-right choice data indicates that subjects adopted a strategy that weighted the maximum (Max.) available outcome of an option more than the minimum (Min.) possible outcome. This strategy is observed when the maximum and minimum values of each option are used as predictor variables (**a**), as well as when the difference between maximum and minimum values and expected values is used to predict choice behaviour (**b**). Subjects are slightly risk-averse when considering only the maximum value of each option (**c**). EV, expected value; a.u., arbitrary units. Error bars are s.e.m. See additional simulations in Supplementary Methods. **d**, In cognitive control, repetitions of trials with high and low control demands tend to produce more rapid RTs relative to trials with markedly different control demands from previous trials. Our behavioural data reproduce the effect of trial sequences (conflict adaptation), a widely replicated effect in studies of cognitive control. **e**, Unsigned feedback prediction errors contributed to longer RTs on trials following immediately, consistent with post-surprise slowing accounts of reactive control[32]. **f**, Drift-diffusion modelling of decision parameters suggests both bottom-up and top-down influences on responses. Estimated drift rates (top) are higher for trials in which the difference in expected value between the two options is higher than for those with similarly valued options (posterior probability of equal drift rates < 0.001), suggesting that identification of high-valued options is easier. Estimates for the decision threshold parameter (bottom) are lower for hard trials than for easy trials (posterior probability of equal decision thresholds = 0.022), implying that control is required in hard trials in adjusting the decision threshold to ensure a timely response. Decision threshold adjustment is a marker of proactive control processes[14,28].

pressure is a control-demanding manipulation[14], it is not automatic that our implementation successfully engaged control in our subjects. Therefore, to assess whether our time-pressure manipulation recruited cognitive control processes, we analysed subjects' behavioural data (accuracy and response times) to identify markers of cognitive control. Results from a two-way analysis of covariance (Fig. 2c; data satisfied assumptions for normality and homogeneity of variance) with current trial type (hard vs easy) and previous trial type (hard vs easy), and controlling for individual subject response times (RTs), show a significant interaction ($f(1,83) = 5.18$, $P = 0.025$, partial $\eta^2 = 0.059$), indicating that RTs for the current trial were influenced by the previous trial type and consistent with sequential adjustment (conflict adaptation) effects[3], a common behavioural marker in cognitive control experiments thought to involve reactive control processes[31]. Further evidence for reactive control in our data is provided by regression of RTs against the unsigned feedback prediction error from the trial immediately preceding. Post-error slowing, like sequential adjustment, is thought to indicate reactive adjustments in control[3] and may reflect novelty or surprise as opposed simply to affectively negative feedback[32]. Estimated β parameters for RT data regressed against absolute feedback prediction errors show a positive relationship between the magnitude of an unsigned prediction error and reaction time (Fig. 2d; $t(305) = 2.68$, $P = 0.0078$, $\eta^2 = 0.023$, $\beta = 0.46 \pm 0.17$), indicating that large deviations (positive or negative) between the reward received and the average reward expected contributed to longer RTs, in line with reactive control.

We additionally modelled subject behavioural data using the hierarchical drift-diffusion modelling package[33] to investigate the influence of value differences on key decision parameters (Fig. 2e). The fit for a drift-diffusion model, in which drift rate and decision threshold parameters are dependent on the difficulty of the trial (hard vs easy), was good (mean square error = 0.0190) and the deviance information criterion[34] for this model was lower than for a drift-diffusion model in which these parameters are not dependent on difficulty (−2,147 versus −1,799; lower values indicate that the

model better explains the data after accounting for complexity). We observed that, for trials in which the mean value of each option was similar (<25-point difference), the estimated decision threshold was lower than for those with clear value differences between options (posterior probability of equal thresholds = 0.0218), while the estimated drift-rate parameter, reflecting differences in integration of information without the need for control, was higher for trials with clear value differences (>25 points) between options (posterior probability of equal drift rates < 0.001). Together these results suggest that, although trials with clear differences between options were easier (as indicated by differences in estimated drift-rate parameters), RT and accuracy data were only partially explained by this and that adjustments in the decision threshold played a significant role in subject behaviour. Adjustments of a decision threshold are a plausible target of proactive control processes[14] and, in our experiment, can be explained as the need for subjects to decrease the decision threshold (that is, invigorate a response) on difficult trials to ensure they respond before the response deadline[35]. Taken together, our behavioural results provide positive evidence that our speeded value-based decision task successfully recruited both proactive and reactive cognitive control processes.

dACC activity during speeded decisions is not explained by choice difficulty. Functional MRI data were modelled using a general linear modelling (GLM) approach, and two GLMs were constructed for our analyses (see Methods). In GLM1, 11 regressors modelled the binned differences in the expected value of the two options (left vs right). In GLM2, eight regressors modelled left and right RT bins with an equal number of trials per bin. For each voxel in an anatomically defined region of interest, including the mPFC (Fig. 3a), dACC and supplementary motor area (SMA), polynomial equations (first to fourth order) were fit to the estimated β values for each voxel across all subjects for each regressor in GLM1 and GLM2.

To assess which polynomial equation better explained the data, Akaike weights[34] were computed for the AIC values obtained for each polynomial to derive a probability that each equation amongst those considered best explained the data (Extended Data Fig. 1). The probability for each equation at each voxel was thresholded at an Akaike weight of >0.999 (equivalent to $P = 0.001$), producing four separate brain probability maps, one for each of the polynomial equations and reflecting the probability that the polynomial equation best accounted for the data relative to the other polynomials. To calculate statistics with volume-based corrections, each of the probability maps was converted to Z-scores and passed to the function spm_uc_clusterFDR used by the Statistical Parametric Mapping (SPM) software package for cluster statistics (see Supplementary Methods for additional information). For both value-binned (GLM1; $P = 0.0087$, voxel extent = 320) and RT-binned (GLM2; $P < 0.001$, voxel extent = 676) values, a cluster of voxels surviving cluster correction (voxel threshold $P = 0.001$, cluster-level family-wise error (FWE) = 0.05) was observed in dACC–mPFC for the quartic polynomial (Fig. 3b). No effects consistent with other polynomial equations (linear, quadratic, cubic) were observed in our volume of interest, even at a lenient cluster threshold ($P = 0.05$, uncorrected). Thus, only the quartic polynomial equation produced significant results in dACC–mPFC, consistent with PRO and EVC2 and inconsistent with the CD or EVC1 models.

dACC activity during speeded decisions is best explained by the PRO account. To distinguish between PRO and EVC2 models, the sign on the quadratic term for the voxels identified in our first step was tested: EVC2 predicts a positive quadratic term, while the PRO model predicts a negative quadratic term. Only voxels with a quadratic term significantly <0 ($P < 0.001$) were identified, consistent with the PRO model and inconsistent with EVC2. In summary, these results favoured PRO over the EVC and CD models in explaining dACC activity during value-based decision-making. The W-pattern predicted by the PRO model is evident in the average β values across the dACC cluster for both our value-binned (Fig. 3c) and RT-binned (Fig. 3d) analyses, while subject behaviour rules out increased error rates (Fig. 3e) or prolonged RTs (Fig. 3f) as an explanation for increased dACC activity in extreme-value-difference trials. Specifically, accuracy for extreme-value bins trials is significantly greater than for the next less-extreme bin ($t(21) = 2.98$, $P = 0.0071$, Cohen's $d = 0.64$, 95% CI = 0.035–0.20), while response times for extreme-value bin trials were significantly lower than for the next less-extreme bin ($t(21) = -2.718$, $P = 0.013$, Cohen's $d = 0.58$, 95% CI = −6.48 to −49.35).

Additional support favouring the PRO model account focuses on the absolute deviation in the overall reward available in a trial versus the long-term average reward available. Simulations of the PRO and EVC models, undertaken in response to reviewer comments on an initial version of this manuscript, suggest a second dissociable prediction made by the two frameworks (Fig. 4a): specifically, the PRO model predicts that large absolute deviations (both positive and negative) from the average reward should produce stronger surprise signals relative to trials with small overall reward deviations. In contrast the EVC model suggests that because, all other factors being equal, dACC activity will track overall reward value, then dACC activity on trials with large absolute deviations will approximately equal dACC activity on trials with small absolute deviations. The rationale for this prediction of the EVC model is that, as available reward increases above the average reward (50% of all trials), the optimal control signal also increases[14], while in the other 50% when reward is lower than the long-term average, the optimal control signal will decrease; on average, both large and small deviations of trial reward, and its consequent influence on the dACC control signal, are centred around the long-term average. Regression of the absolute deviation of the total reward available on each trial from the long-term average yields a significant cluster in the mid-cingulate zone (peak voxel = 12,246, $P < 0.001$, voxel extent = 748), a region that has previously been implicated in signalling differences in reward level (Fig. 4b, green area).

A final piece of evidence for the PRO model in our data comes from analysis of activity at the time of feedback (Fig. 4b, red area). As described above, under the PRO model, dACC activity is related to the 'suprisingness' of a salient event. In our previous analyses, we limited ourselves to the interval starting from the onset of a trial to the generation of a response, primarily because the CD and EVC models do not make clear predictions regarding dACC activity without the requirement for generating behaviour. If our data are consistent with the PRO model account, dACC activity following feedback should correlate with the (unsigned) feedback prediction error. Regression of the per-trial prediction error for each subject on BOLD data yields a significant cluster (Fig. 4; peak voxel = 0, 40, 40, cluster-corrected P(FWE) < 0.001, voxel extent = 3,010, voxel-wise threshold = 0.001) in dACC. The PRO model thus accurately accounts for dACC activity over the course of the entire trial in our speeded decision-making task.

Surprise and reward elsewhere in the brain. Although the main goal of this study was to investigate dACC function, additional regions are known to be active during value-based decision-making and cognitive control. We therefore repeated our polynomial fit analyses over the entire brain. We found that activity in bilateral lateral PFC (lPFC) (right: peak voxel = 46, 32 22, cluster-level $P = 0.004$, voxel extent = 426; left: peak voxel = −52, 32 28, cluster-level $P < 0.001$, voxel extent = 642) and bilateral PFC (right: peak voxel = 44, −42, 42, cluster-level $P < 0.01$, voxel extent = 614; left: peak voxel = −32, −62, 46, cluster-level $P = 0.01$, voxel extent = 348)

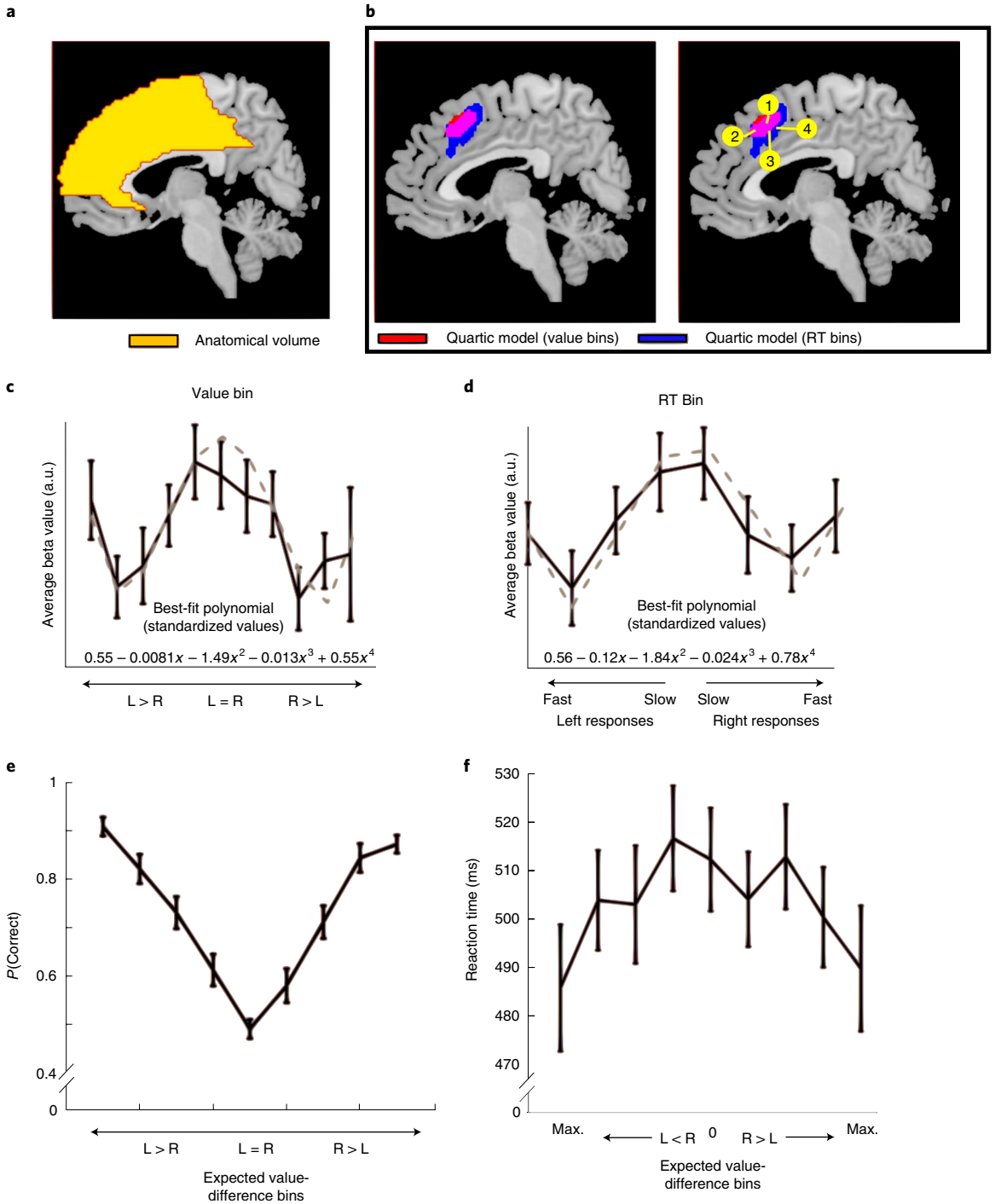


Fig. 3 | fMRI and behavioural performance. a, Anatomically defined region including dACC, mPFC and SMA where model predictions were tested. **b**, Only signals associated with quartic polynomial equations (left, L) were identified in the region, consistent with the PRO and EVC2 models only. The best-fit equation for voxels in this region included a quadratic term that was significantly less than 0 across subjects (right, R), consistent with the PRO model and inconsistent with the EVC2 model. Yellow dots indicate peak voxels reported in previous studies reporting dACC activity related to CD (1, 2)[23,29] and value difference (3, 4)[52]. **c**,**d**, Average β weights for voxels that were best fit by a quartic polynomial (Akaike weight > 0.999) binned by value (**c**) and RT (**d**), controlling for RT in each trial. **e**, Subjects' ability to correctly select the option with the higher expected value improved as a function of the difference in expected value between options. **f**, RTs decreased as a function of value difference. Results of behavioural performance therefore rule out a possible explanation for increased dACC activity in extreme trials as being due either to increased error rate or increased RT. Error bars are s.e.m.

is, as in the dACC, best described by a quartic polynomial function (Fig. 5) while ventromedial PFC (vmPFC; peak voxel = −2, 38, −2, cluster-level $P < 0.001$, voxel extent = 666) is described best by a quadratic polynomial (Fig. 5). Along with dACC, parietal cortex and lPFC are core nodes in the 'salience' network[36,37] that is active during salient or surprising events[38], while vPFC has frequently been identified as a region sensitive to value and reward potential[13,39], consistent with our findings here.

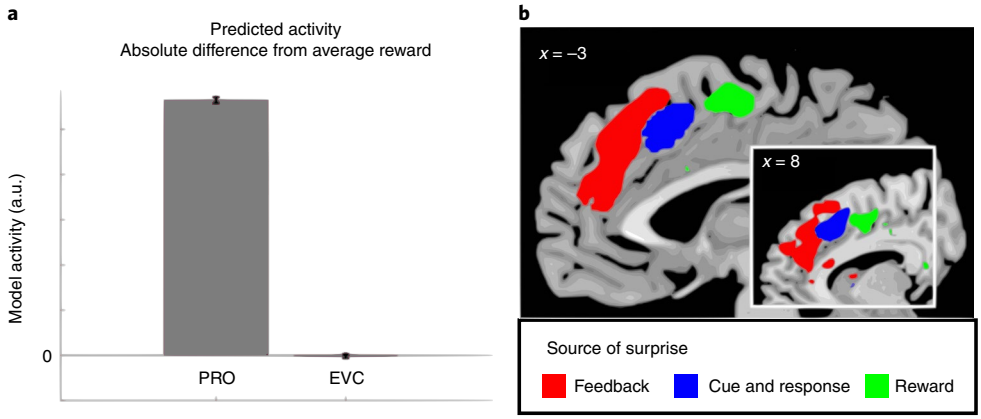
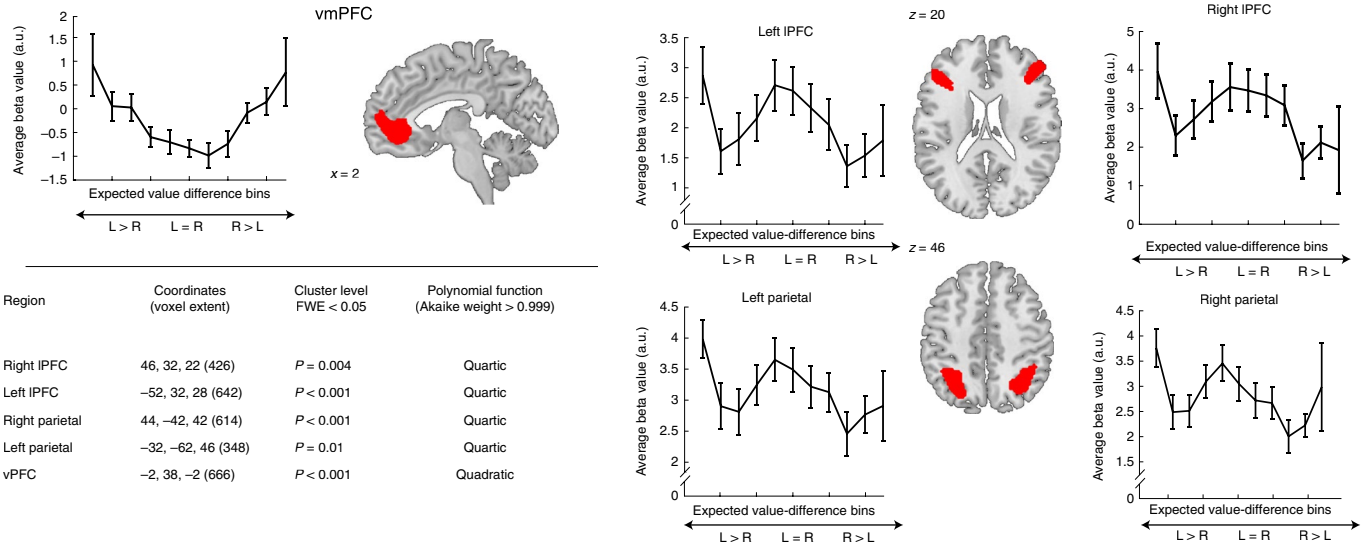


Fig. 4 | Feedback and reward-related surprise. a, The PRO and EVC models make additional competing predictions regarding dACC activity for high- and low-reward-value trials. **b**, Activity in the mid-cingulate cortex–SMA (green) is greater for high-reward trials, consistent with PRO model predictions. dACC activity also correlates with unsigned prediction errors following task feedback (red). The letter *x* denotes the sagittal plane.



Region	Coordinates (voxel extent)	Cluster level FWE < 0.05	Polynomial function (Akaike weight > 0.999)
Right lPFC	46, 32, 22 (426)	$P = 0.004$	Quartic
Left lPFC	–52, 32, 28 (642)	$P < 0.001$	Quartic
Right parietal	44, –42, 42 (614)	$P < 0.001$	Quartic
Left parietal	–32, –62, 46 (348)	$P = 0.01$	Quartic
vPFC	–2, 38, –2 (666)	$P < 0.001$	Quadratic

Fig. 5 | Whole-brain polynomial fits. Main analysis testing for quadratic and quartic trends repeated for the whole brain, to test for the involvement of additional regions. The table shows coordinates, cluster extent and *P* values for all significant clusters. Activity in bilateral PFC and bilateral parietal cortex was best explained by a quartic polynomial; activity in vmPFC) was best explained by a quadratic polynomial. The sagittal and axial planes used to generate images are denoted by *x* and *z*, respectively.

Discussion

Taken together, our fMRI analyses support multiple predictions of the PRO model but fail to find evidence anywhere within the mPFC of an optimal control signal posited by the EVC model. We observed signals in the mPFC during value-based decision-making under time pressure that are best explained by a fourth-order (quartic) polynomial equation, consistent only with predictions of the PRO model of dACC and inconsistent with CD and EVC models. We additionally observed signals consistent with surprise in the mid-cingulate cortex related to the absolute deviation of the total available reward on a trial from the average reward over all trials, consistent with previous work[15] observing activity in the region related to changes in outcome value. Finally, we found activity in the dACC that correlates with an unsigned feedback prediction error. The PRO model accounts for these various effects observed in the cingulate through the single mechanism of negative surprise, derived from standard calculation of prediction errors in reinforcement learning. Negative surprise has previously been used to explain a host of effects observed in dACC under cognitive control and decision-making[15], and signifies the deviation between actual events and their likelihood of occurring.

Our behavioural results replicate previous investigations of value-based decision-making under time pressure[26], which found that RT and accuracy data were influenced by the magnitude of the value difference between options, with more rapid information integration on high-difference trials and lower decision thresholds on low-difference trials. While the rate at which stimulus information is integrated is thought to be independent of control processes, decision threshold adjustment is a plausible target of cognitive control[14], suggesting that proactive control processes were recruited during low-difference trials to ensure a timely response—that is, control was required to invigorate responding (motivated control), and not merely to inhibit a prepotent response. We additionally observed sequential adjustment effects in our data[3] as well as post-feedback

slowing that was dependent on the absolute magnitude of the difference between received points and the expected value of a selected option; both results are consistent with the engagement of reactive control mechanisms[31]. Overall, our behavioural results demonstrate that our speeded decision-making task involved cognitive control, a requirement for testing of predictions made by the EVC model. Moreover, control was required principally on low-value-difference trials: the increased rate of information integration on high-difference trials suggests that these trials could be performed more automatically. Our behavioural data therefore suggest that our speeded decision-making task specifically tapped into motivated control processes, as distinct from inhibitory processes that are typically the target of cognitive control research.

The results of our fMRI analyses provide further evidence for a role of control processes in our study. Within the large anatomical volume (Fig. 3a) in which we tested model predictions of BOLD activity, we observed significant, model-related activity in regions of dACC consistent with regions reported in recent studies of inhibitory control during value-based choice[23,29]. As in the case of behavioural studies of cognitive control, neuroimaging studies have frequently emphasized the role of control regions in response inhibition[3]. Although recent research emphasis has turned to the question of how areas of the brain frequently implicated in inhibitory processes might also support behaviours that rely on motivation and invigoration[2,12,22], especially as related to balancing the costs and benefits of exerting control, a number of salient questions regarding the motivated vs inhibitory control dichotomy remain[2]. One possibility is that, while the processes underlying response invigoration and inhibition occur within the same regions of the brain, they represent two distinct modes of operation that involve different mechanisms—for example, conflict and surprise detection or value computation[14], respectively. Alternately, it may be the case that response invigoration and inhibition are dependent on the same general computations that are broadly applicable across an array of contexts.

In support of this latter possibility, our fMRI analyses found that the pattern of activity we observed in the mPFC was best accounted for by the PRO model relative to two other influential models of dACC. The PRO model casts dACC as an epistemic controller—signalling the deviation between expectations and observations independently of affective valence[15]—and has been used to explain behaviour and brain activity in a wide range of contexts, including the detection of salient events[30], foraging[40], effort-based decision-making[21] and learning task structure[41,42]. In the current study, the W-pattern predicted by the PRO model (and observed in our data) is the product of the same computations applied following the onset of a stimulus, during which dACC activity increases as a function of value differences (Supplementary Fig. 3 and Methods), and following the generation of a response, during which dACC activity increases as a function of value similarity. In line with the epistemic control account, dACC thus tracks deviations from expectations (irrespective of whether the expectation concerns value prediction or response selection). This view is consistent with previous work implicating the dACC in behavioural adaptation[16,17,20,43]. Whether the PRO could account for effort preference (that is, choosing between rewards requiring different levels of effort to be obtained) remains to be empirically confirmed. An additional possibility is that increasing complexity of the decision (for example, by manipulating multiple decision attributes or providing several simultaneous options) may recruit other regions within or outside the mPFC. Interestingly, additional analyses revealed two adjacent clusters in dACC–mPFC coding for feedback-related surprise (anterior mPFC) and reward surprise (mid-cingulate cortex–SMA; see Fig. 4b). Previous fMRI studies identified similar regions using the PRO model as a framework for model-based analysis of BOLD signals. Together, these findings support surprise coding as the core mechanism underlying mPFC function more generally, with the source of the surprise determining the exact neural population implicated. Recent work has proposed hierarchical predictive coding as a working principle for understanding interactions between the mPFC and lPFC[41,42,44], in line with ample evidence of different sub-regions of PFC coding different types of information along a rostro-caudal abstraction gradient[45–48]. In other words, epistemic control unifies cognitive control and motivated control functions previously attributed to the dACC, and provides a possible functional mechanism for the whole PFC. Indeed, our whole-brain analyses identified additional regions, including both bilateral lPFC and parietal cortex, with activity consistent with surprise signalling (Fig. 5). Although these regions show similar patterns of activity, we do not interpret this to mean they are functionally the same. Rather, it may be that our value-based, speeded decision-making task is too simple to meaningfully dissociate activity in the regions, and that more complex tasks involving hierarchical cognitive and attentional control are needed to specify functional differences within this network of regions. Finally, the PRO model, beyond simply predicting the pattern of data observed in this study, made a strong and falsifiable prediction regarding the time course of activity over the entire course of a trial, from the moment of stimulus onset to the generation of a response and the delivery of feedback (Fig. 4) using the exact same computational formalism. This strong prediction may be useful for further efforts to adjudicate amongst computational models of dACC, especially models such as the reduced vector Preisach model (RVPM)[49,50], that similarly propose a central role for dACC in error signalling. Specifically, RVPM and PRO make similar predictions regarding the period following a response but diverge in their predictions following cue onset (see Supplementary Methods). By suggesting a principal role for the dACC in calculating surprise, the PRO model and its extensions provide a perspective on the dACC as a critical node in a network of control regions whose joint, distributed activity ultimately contributes to motivated behaviour[51].

While a broader set of brain regions have been implicated in signalling surprise in the past, dACC activity is critical when such signals are necessary to adapt one's predictive model of the task[38]. In our data, sub-regions of the mPFC revealed specific sensitivity to dissociable types of surprise, all relevant for behavioural adaptation (cue–response related, feedback-related and reward-related surprise; see Fig. 5), indicating surprise as a plausible underlying computational principle of function for the whole mPFC region. While the view of mPFC as being primarily involved in calculating surprise is not necessarily inconsistent with other views of the region as engaging in value-related processes[52], here we explicitly rule out a value-based account of mPFC function[14] and failed to observe any activity in the region unambiguously related to value alone. Because this study was specifically designed to investigate model predictions under speeded value-based decision-making, it is unable to address a number of questions, including explicit testing of the timing of surprise signals in dACC over the course of a trial, how the function of multiple regions showing activity consistent with surprise can be differentiated, and how regions indexing surprise and value information interact during motivated control. Future work should therefore focus on explicit, formal tests to distinguish between value- and surprise-related activity in the mPFC and the rest of the brain, to determine the conditions and neural regions to which these frameworks apply.

Methods

Participants. Twenty-three healthy volunteers participated in this experiment (ten males), with a mean age of 23 ± 2 years ranging between 19 and 28 years. The number of subjects included was determined based on previous studies of dACC activity during value-based choice[23,52], which observed robust value-related effects in the region using sample sizes of 14–20 individuals. The Ethical Committee of Ghent

University Hospital approved the experimental protocol. All participants signed an informed-consent form before participating in the experiment, and filled in a safety checklist to exclude neurological or psychiatric conditions and contraindications for participations. No participating subjects were excluded from analysis.

Experimental procedure. On arrival, participants filled in and signed the informed consent and the pre-scanning checklist. Subsequently, they underwent a training session before entering the scanner. The training consisted of performing a series of decisions. On each trial, participants were presented with two fractal images on the screen, one on the left and one on the right (1,000 ms). Participants could select one image by pressing the left or right response key on the keyboard. After 500 ms, feedback was given showing the selected image and the number of points obtained in that trial (1,000 ms). After showing a blank screen (1,000 ms), the following trial started with eight images in total. Each image was associated with a fixed number of points, ranging (in ten-point increments) between 10 and 80. Participants were instructed to remember the number of points associated with every image. They were informed that the task during the subsequent scanning session would be different, but that the same images would appear and each image would be associated with the same number of points as in the training session. They were therefore specifically encouraged to remember this association. Participants were also informed that all points gained during the scanning session would be converted into money (up to €10, 1 cent per 100 points obtained), which they would be paid at the end of the experiment in addition to the amount paid for participation. At the end of the training session, all participants confirmed that they had learned the number of points associated with each image and that some images would deliver more points than others. Importantly, the number of points associated with each fractal image was changed across participants according to eight possible randomizations, to control for visual features of the images. Data collection and analysis were not performed blind to the conditions of the experiments.

Speeded value-based decision-making task. Participants performed a speeded value-based decision-making task while undergoing fMRI. The task started with a short training session, to acquaint participants with the procedure and the response buttons in the scanner. A total of 20 training trials were performed. In each trial, participants were presented with four fractal images, two on the left side of the screen and two on the right (750 ms; Fig. 1c). They were instructed to select options (left or right side of the screen) to maximize the number of points earned. From the side selected, they had an equal probability of receiving one of the two presented images, and the corresponding amount of points, as feedback. Following a response, a blank screen was presented for a jittered variable interval (randomly selected, range 3,000–5,000 ms, mean 4,000 ms). A feedback display followed, showing the image selected by the computer among the images on the chosen side, with the corresponding number of points obtained by the participant (1,000 ms). Feedback was followed by a randomly jittered interval, ranging between 3,000 and 5,000 ms (mean, 4,000 ms). After completing the training, participants were asked whether the points associated with each image corresponded with what they had learned in the training outside the scanner, and every participant confirmed this. Subsequently the task started, with the same timings and procedure. The task consisted of 160 trials in total in one block, lasting approximately 25 min. No specific criteria were established for identifying and removing outliers, and no subjects were removed from our analyses based on performance.

Behavioural data analysis. Sequential adjustment effects are a behavioural marker of reactive control commonly observed in studies of cognitive control, and are indicated by decreases in RT in trials with control demands similar to those of the previous trial—that is, responses for an easy trial immediately following another easy trial will tend to be more rapid than for easy trials following hard trials. Similarly, responses for hard trials will tend to be more rapid if the trial immediately preceding was also hard than if that trial was easy. To assess the presence of sequential adjustments in our data, we defined hard trials as those in which the expected value difference between left and right options was <25 points (median), while easy trials were defined as those in which the difference was >25 points. Mean subject RTs were calculated for four conditions defined by two current trial conditions (hard vs easy) and two previous trial conditions (hard vs easy) and analysed using two-way analysis of variance. Note that we used an alpha level of 0.05 for all tests, and report two-tailed *P* values. Using the same definitions of hard and easy trials, RT and accuracy data were modelled using hierarchical drift-diffusion modelling[33]. Trials were classified as correct if subjects selected the option with the highest expected value. Trials in which the difference in expected value between options was five or less were discarded, because choosing either option might constitute a correct choice. Drift-diffusion parameters were estimated for drift rate and decision threshold using 15,000 samples (2,000 initial samples discarded). Two drift-diffusion models were estimated: in the first, two drift rate and decision threshold parameters were estimated, one for each trial difficulty condition (hard vs easy). In the second, a single drift rate and decision threshold parameter was estimated for all trials regardless of difficulty condition. Relative model fits (accounting for model complexity) were compared using the deviance information criterion[34], in which lower scores indicate better model fit.

fMRI data acquisition. Data were acquired using a 3T Magnetom Trio MRI scanner (Siemens) with a 32-channel radio-frequency head coil. In an initial scanning sequence, a structural T_1 weighted MPRAGE sequence was collected (176 high-resolution slices, repetition time (t_R) = 1,550 ms, echo time (t_E) = 2.39, slice thickness = 0.9 mm, voxel size = 0.9 × 0.9 × 0.9 mm^3, field of view (FoV) = 220 mm, flip angle = 9°). As the second sequence, functional images were acquired using a T_2*-weighted EPI sequence (33 slices per volume, t_R = 2,000 ms, t_E = 30 ms, no interslice gap, voxel size = 3 × 3 × 3 mm^3, FoV = 192 mm, flip angle = 80°). On average, 760 volumes per participants per task were collected. Each task lasted approximately 25 min.

fMRI data analysis. The first four volumes of each functional run were discarded to allow for steady-state magnetization. The data were preprocessed with SPM 12 (http://www.fil.ion.ucl.ac.uk/spm). Images were realigned to the first image of the run. The structural T_1 image was co-registered to the functional mean image for normalization purposes. Normalization was performed through the unified segmentation and nonlinear warping approach implemented in SPM 12. Functional images were normalized to the Montreal Neurological Institute template. Resulting functional images were smoothed with a Gaussian kernel of 8 mm, full-width, half-maximum.

For each single subject, a GLM approach was applied to identify condition-specific activation. In GLM1, trials were divided into 11 different bins as a function of the value difference between the two sides of the screen (five bins for the left-side average value > right-side average value ranging between 55 and 10 points difference; one bin for trials where both sides were close in value (−5 to +5 average points difference for right–left options); and five bins for right-side average value > left-side average value, ranging between 55 and 10 points difference). Because of the nature of the task, each of the value bins included a different number of trials (with an average of six trials and minimum of four) in the extreme-value bins.

For each regressor, three parametric modulators were added: RT at the current trial, RT at the previous trial and unsigned prediction error at the previous trial. An additional regressor was added to model responses over the time limit (misses). Twelve regressors were added to model feedback in each condition, and in misses. Six further regressors were added to account for motion (X, Y, Z translation, pitch, yaw and roll). In GLM2, trials were divided into eight different bins as a function of RT and side. The purpose of this binning procedure was to equalize the number of trials (20) per bin, as well as to emphasize that quartic effects observed in the dACC were unrelated to RT. For each regressor the RT on current trial was added as a parametric modulator. One regressor was added for responses over the time limit. Nine regressors were added to model feedback in each condition and feedback for misses. Six regressors were added to account for motion, as above.

Because each computational model under consideration predicts effects specifically related to the dACC and medial mPFC, we restricted our analyses to an anatomically defined region including the dACC, SMA and superior mPFC, defined anatomically using the wfupickatlas toolbox (dilation = 2). The predictions of each model considered (see below) can be described in a straightforward fashion by polynomial equations. Therefore, for all voxels within this region, β values estimated for that voxel at the first level, and for all subjects, for each regressor in GLM1 were used to fit polynomial equations (linear, quadratic, cubic and quartic). Following polynomial fits, an AIC value was computed for each polynomial equation and these values were used to compute Akaike weights[53], yielding a value of 0–1 for each model, indicating the probability that that model is the best of all models under consideration. For each voxel, therefore, a value was obtained indicating the probability that the β values at that voxel were best explained by each of the four polynomial equations. These values were thresholded at 0.999 (equivalent to $P < 0.001$) and, for all voxels surviving this threshold, the best-fit polynomial equation was determined averaged over all surviving voxels, and regressed against the BOLD signal to compute whole-brain statistics. These steps were repeated for GLM2 (binned by RT).

To analyse feedback-related activity, a third GLM (GLM3) was created with two regressors: one modelling the choice period and one corresponding to the onset of task feedback was modelled on each trial. A parametric modulator reflecting the (objective) unsigned prediction error (calculated from global task contingencies and based on subjects' trial-by-trial choices) was included for the feedback regressors. A third regressor was used to model non-response trials, and six motion regressors were included as in GLMs 1 and 2.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the corresponding author on request.

Code availability

Detailed modelling methods are included in Supplementary Methods. All scripts used to derive predictions from the various computational models are available online at https://github.com/modelbrains/EVC-Simulations

Received: 10 January 2019; Accepted: 28 November 2019;
Published online: 13 January 2020

References

1. Ebitz, R. B. & Hayden, B. Y. Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nat. Neurosci.* **19**, 1278–1279 (2016).
2. 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).
3. 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).
4. 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).
5. Brown, J. W. & Braver, T. S. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005).
6. Yee, D. M. & Braver, T. S. Interactions of motivation and cognitive control. *Curr. Opin. Behav. Sci.* **19**, 83–90 (2018).
7. Botvinick, M. & Braver, T. Motivation and cognitive control: from behavior to neural mechanism. *Annu. Rev. Psychol.* **66**, 83–113 (2015).
8. Inzlicht, M., Shenhav, A. & Olivola, C. Y. The effort paradox: effort is both costly and valued. *Trends Cogn. Sci.* **22**, 337–349 (2018).
9. Verguts, T., Vassena, E. & Silvetti, M. Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. *Front. Behav. Neurosci.* **9**, 57 (2015).
10. Apps, M. A. J., Grima, L. L., Manohar, S. & Husain, M. The role of cognitive effort in subjective reward devaluation and risky decision-making. *Sci. Rep.* **5**, 16880 (2015).
11. Aarts, E. & Roelofs, A. Attentional control in anterior cingulate cortex based on probabilistic cueing. *J. Cogn. Neurosci.* **23**, 716–727 (2010).
12. Vassena, E. et al. Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS One* **9**, e91008 (2014).
13. Vassena, E., Krebs, R. M., Silvetti, M., Fias, W. & Verguts, T. Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia* **59**, 112–123 (2014).
14. 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).
15. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14**, 1338–1344 (2011).
16. Kolling, N. et al. Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* **19**, 1280–1285 (2016).
17. di Pellegrino, G., Ciaramelli, E. & Làdavas, E. The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *J. Cogn. Neurosci.* **19**, 275–286 (2007).
18. Rushworth, M. F. S., Hadland, K. A., Gaffan, D. & Passingham, R. E. The effect of cingulate cortex lesions on task switching and working memory. *J. Cogn. Neurosci.* **15**, 338–353 (2003).
19. Ridderinkhof, K. R., van den Wildenberg, W. P., Segalowitz, S. J. & Carter, C. S. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* **56**, 129–140 (2004).
20. Sheth, S. A. et al. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioral adaptation. *Nature* **488**, 218–221 (2012).
21. Vassena, E., Deraeve, J. & Alexander, W. H. Predicting motivation: computational models of PFC can explain neural coding of motivation and effort-based decision-making in health and disease. *J. Cogn. Neurosci.* **29**, 1633–1645 (2017).
22. Klein-Flügge, M. C., Kennerley, S. W., Friston, K. & Bestmann, S. Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. *J. Neurosci.* **36**, 10002–10015 (2016).
23. Shenhav, A., Straccia, M. A., Cohen, J. D. & Botvinick, M. M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat. Neurosci.* **17**, 1249–1254 (2014).
24. Lin, H., Saunders, B., Hutcherson, C. A. & Inzlicht, M. Midfrontal theta and pupil dilation parametrically track subjective conflict (but also surprise) during intertemporal choice. *Neuroimage* **172**, 838–852 (2018).
25. Krajbich, I., Lu, D., Camerer, C. & Rangel, A. The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* **3**, 193 (2012).
26. Mormann, M. M., Malmaud, J., Huth, A., Koch, C. & Rangel, A. The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm. Decis. Mak.*, **5**, 437–449 (2010).
27. Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108 (1978).
28. Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414–421 (2014).
29. Shenhav, A., Straccia, M. A., Botvinick, M. M. & Cohen, J. D. Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cogn. Affect. Behav. Neurosci.* **16**, 1127–1139 (2016).
30. Alexander, W. H. & Brown, J. W. A general role for medial prefrontal cortex in event prediction. *Front. Comput. Neurosci.* **8**, 69 (2014).
31. Braver, T. S. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* **16**, 106–113 (2012).
32. Wessel, J. R. & Aron, A. R. On the globality of motor suppression: unexpected events and their influence on behavior and cognition. *Neuron* **93**, 259–280 (2017).
33. Wiecki, T. V., Sofer, I. & Frank, M. J. HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. *Front. Neuroinform.* **7**, 14 (2013).
34. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Linde, A. V. D. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* **64**, 583–639
35. Murphy, P. R., Boonstra, E. & Nieuwenhuis, S. Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nat. Commun.* **7**, 13526 (2016).
36. Downar, J., Crawley, A. P., Mikulis, D. J. & Davis, K. D. A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *J. Neurophysiol.* **87**, 615–620 (2002).
37. Knight, R. T. & Nakada, T. Cortico-limbic circuits and novelty: a review of EEG and blood flow data. *Rev. Neurosci.* **9**, 57–70 (1998).
38. O'Reilly, J. X. et al. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl Acad. Sci. USA* **110**, E3660–E3669 (2013).
39. Grabenhorst, F. & Rolls, E. T. Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cogn. Sci.* **15**, 56–67 (2011).
40. Brown, J. W. & Alexander, W. H. Foraging value, risk avoidance, and multiple control signals: how the ACC controls value-based decision-making. *J. Cogn. Neurosci.* **29**, 1656–1673 (2017).
41. 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).
42. Alexander, W. H. & Brown, J. W. Frontal cortex function as derived from hierarchical predictive coding. *Sci. Rep.* **8**, 3843 (2018).
43. Steenbergen, H., van, Band, G. P. H., Hommel, B., Rombouts, S. A. R. B. & Nieuwenhuis, S. Hedonic hotspots regulate cingulate-driven adaptation to cognitive demands. *Cereb. Cortex* **25**, 1746–1756 (2015).
44. Alexander, W. H., Vassena, E., Deraeve, J. & Langford, Z. D. Integrative modeling of pFC. *J. Cogn. Neurosci.* **29**, 1674–1683 (2017).
45. Badre, D. & Nee, D. E. Frontal cortex and the hierarchical control of behavior. *Trends Cogn. Sci.* **22**, 170–188 (2018).
46. Koechlin, E. & Summerfield, C. An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci.* **11**, 229–235 (2007).
47. Nee, D. E. & Brown, J. W. Rostral–caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *NeuroImage* **63**, 1285–1294 (2012).
48. Nee, D. E. & D'Esposito, M. The hierarchical organization of the lateral prefrontal cortex. *eLife* https://doi.org/10.7554/eLife.12112 (2016).
49. 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).
50. Silvetti, M., Alexander, W., Verguts, T. & Brown, J. W. From conflict management to reward-based decision making: actors and critics in primate medial frontal cortex. *Neurosci. Biobehav. Rev.* **46**, 44–57 (2014).
51. Alexander, W. H., Brown, J. W., Collins, A. G. E., Hayden, B. Y. & Vassena, E. Prefrontal cortex in control: broadening the scope to identify mechanisms. *J. Cogn. Neurosci.* **30**, 1061–1065 (2018).
52. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
53. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196 (2004).

Acknowledgements

We thank C. Holroyd, J. Brown, T. Verguts, Z. Langford, J. Maynard Keenan and M. Pessiglione for useful discussions. W.H.A. was supported by FWO-Flanders Odysseus II Award (no. G.OC44.13N) and by start-up funds provided by Florida Atlantic University. E.V. was supported by the Marie Sklodowska-Curie actions with a standard IF-EF fellowship within the H2020 framework (H2020-MSCA-IF2015,

grant no. 705630). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

E.V. and W.H.A. conceived the study and designed the work. E.V., W.H.A. and J.D. performed acquisition, analysis and interpretation of data and drafted the manuscript. E.V. and W.H.A. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-019-0801-5.

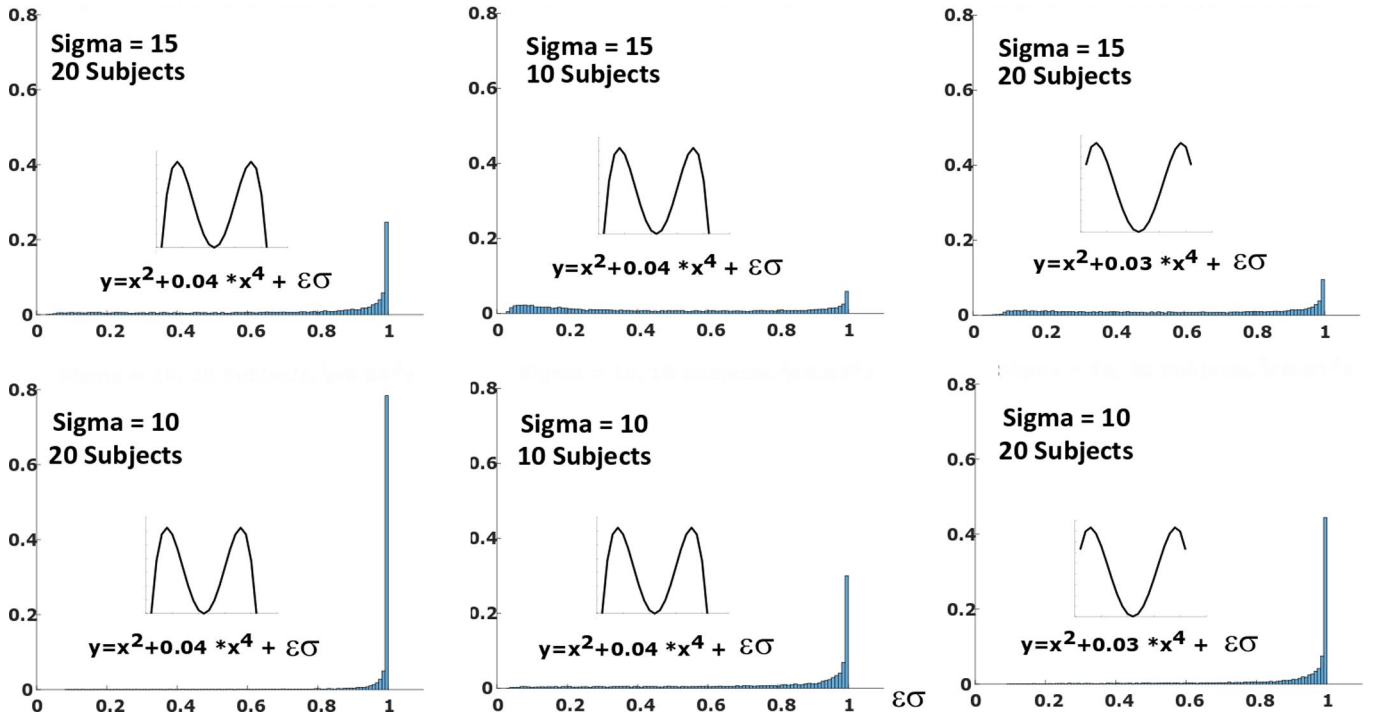
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0801-5.

Correspondence and requests for materials should be addressed to W.H.A.

Peer review information Primary handling Editor: Marike Schiffer

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.



Extended Data Fig. 1 | Factors influencing Akaike Weight estimates. Our analysis of fMRI data using polynomial functions fit to beta estimates from BOLD data is substantially different from classical or model-based fMRI analyses. In order to explore what factors may influence our results, as well as how nested polynomial functions independently account for observed patterns of activity in our data, we carried out additional simulations and analyses. First, we asked what factors affect the ability of our analysis approach to identify quartic effects using Akaike Weights. To answer this, we conducted six simulations of synthetic data generated by a quartic polynomial equation while varying the noise, number of subjects, and shape of the function. 10,000 simulation runs were conducted for each condition, and Akaike Weights calculated to obtain a distribution of the likelihood of Akaike Weights. The results of these simulations suggest that, as in classical univariate analyses, the likelihood of obtaining an Akaike Weight >0.999 for the quartic polynomial function (when the data was in fact generated by a quartic polynomial function) depends both on the quantity and noisiness of the data itself. That is, with less data and more noise, it becomes less likely that our analyses will assign the quartic polynomial function an Akaike Weight of 0.999 or higher. Additionally, changes in the shape of the quartic function that render it more similar to other (quadratic) functions reduce the likelihood of calculating an Akaike Weight >0.999 for the quartic function. These results suggest our analysis approach parallels typical fMRI analyses in terms of the factors that influence the likelihood of observing significant effects.

nature research

Corresponding author(s): William H. Alexander

Last updated by author(s): 10/29/2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collected using the Eprime2 experiment platform
Data analysis	Data were analyzed using SPM12(fMRI) and custom code developed to calculate Akaike Weights implemented in MATLAB and based on Wagenmakers & Farrell, 2004. Behavioral data were analyzed using statistical tests available in MATLAB, as well as the HDDM package for fitting reaction time and choice data (Wiecki, Sofer & Frank, 2013).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study investigates the neural and behavioral correlates of motivated control using theoretical frameworks derived from investigation of inhibitory control based on quantitative measures of behavior (reaction time, accuracy) and brain activity (BOLD signal).
Research sample	Participants in this study were members of the Belgian community associated with Ghent University, and volunteered to participate using an online recruiting portal available to Ghent University students. The mean age of the sample was 23 years (+/- 2 years, minimum 19, maximum 28). The student population at Ghent University provides an accessible pool of research participants.
Sampling strategy	The research sample was selected randomly (self-selected) from the pool of students participating in research studies at Ghent University. A priori sample size calculation was not performed; the number of participants (23 total) was selected to be in line with similar studies investigating the neural bases of decision-making (e.g., Kolling et al., 2011, Science(20 participants); Shenhav et al., 2014, Nature Neuroscience (15 participants in experiment 1, 14 experiments in experiment 2))
Data collection	Data were acquired using a 3T Magnetom Trio MRI scanner (Siemens), with a 32-channel radio-frequency head coil. Experiments were implemented in EPrime2, and participants interacted with the experiment via a 2-button response box and visual display projected into the scanner. No effort was made to standardize visual angle of the display, and participants were only requested to inform the researchers as to whether the entire screen was visible. Only researchers were present during data collection, and all researchers were aware of the experimental hypothesis. As there was only a single condition, research was not blinded, double or otherwise.
Timing	All data were collected between Jan 13 2016 and Feb 11 2016 on a weekly basis.
Data exclusions	No data were excluded
Non-participation	No participants dropped out of the study after providing informed consent.
Randomization	Participants were not allocated into different experimental groups (i.e., the task was the same for each individual). Visual stimuli (total of 8 fractal images) were collected in 8 unique image-point combinations such that each fractal image could be associated with a high, medium or low number of points; beginning with participant 1, each of these stimulus sets was used in order before restarting with the initial stimulus set on subject 9.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above
Recruitment	Participants were recruited from the Ghent University student population through the Experimetrix online recruiting platform. Given that our subject pool is entirely self-selected from a population of university students, it is possible that our sample does

not reflect the general population. However, our sample is comparable to samples used in many neuroimaging studies of cognitive control and decision making that primarily recruit university students, and therefore our results can be directly related to existing results

Ethics oversight: Ethics Committee of the Ghent University Hospital

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type: Task Event-Related Design

Design specifications:
1 Block of ~25 minutes
160 Trials/Block
~10 Seconds/Trial (+/- 2 seconds due to randomly jittered pre- and post- feedback intervals)

Behavioral performance measures: Left/Right button presses and reaction time were recorded. No criteria were established to exclude subjects, and no subjects were excluded for performance issues).

Acquisition

Imaging type(s): Functional and Anatomical

Field strength: 3T

Sequence & imaging parameters:
T1-weighted anatomical images:
TR=1550ms
TE=2.39ms
flip angle = 9 degrees
FoV = 220mm
voxel size = 0.9mmX0.9mmX0.9mm
slice thickness = 0.9mm
T2-weighted function images:
gradient-echo EPI sequence
TR=2000ms
TE=30ms
flip angle = 80 degrees
FoV=192mm
voxel size = 3mmX3mmX3mm
33 interleaved slices/volume

Area of acquisition: whole brain

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software: SPM12

Normalization: Anatomical scans were normalized to MNI space using SPM12 normalization procedures to minimize differences between the anatomical image and template image. A 12-parameter(translation, rotation, zoom, sheear) affine linear transformation was applied, followed by a non-linear warp (deformation fields) transformation. Functional images were co-registered to anatomical scans for each subject.

Normalization template: MNI 152

Noise and artifact removal: 6 Movement Parameters (X,Y,Z translation, pitch, roll, yaw) were entered as additional regressors in 1st level GLMs

Volume censoring: 1st five volumes were discarded to allow for steady state magnetization

Statistical modeling & inference

Model type and settings: Generalized Linear Models were used for 1st level fixed-effects analyses. Average beta estimates for regressors of interest (binned value differences and RTs) were used to fit polynomial equations. 2nd Level (RFX) GLMs were used for additional group analyses

Effect(s) tested: Main Analysis: Polynomial equations (linear, quadratic, cubic, quartic) were fit to beta estimates for each voxel (one beta value per subject per value or RT bin), and the Akaike Information Criterion (AIC) was computed for each polynomial equation in order to determine the best-fit equation for each voxel.

Additional Analyses: 1)Feedback-related surprise - Absolute value of feedback prediction error regressed against BOLD

data. 2) Reward-level surprise - Absolute value of difference of average reward for current trial from long-run average over all trials, regressed against BOLD data.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s): Our specific hypotheses specifically concern dACC/mPFC function, and therefore our primary analyses were restricted to the anatomical region, defined in the wfupickatlas toolbox to include ACC, preSMA, and ventromedial PFC

Statistic type for inference (See Eklund et al. 2016): In order to calculate statistics, AIC values were converted into Akaike Weights (Wagenmakers & Farrell, 2004). The ratio of Akaike Weights was used to calculate the probability that each polynomial equation (of those considered) best explained the data. FWE cluster corrections (FWE $p<0.05$) were applied, and voxels that passed an effective threshold of $p<0.001$ (ratio of Akaike Weights >0.999) were included.

Correction: Family Wise Error correction was used for cluster level statistcs.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis