Supplementary Note

Although our study focused on the activation of dorsal anterior cingulate cortex, it is worth noting that, in addition to dorsal anterior cingulate cortex, the pre-supplementary motor area and the supplementary motor area, activity associated with error responses was greater than activity associated with correct responses (fixed trials contrast) in left rostral anterior cingulate cortex (area 32; x = -8, y = 36, z = 26), right insular cortex (area 45; x = 37, x = 18, y = 8), and right parietal cortex (area 40; x = 58, y = -36, z = 32). Furthermore, as determined by the conservative statistical threshold, activity associated with correct feedback was greater than activity associated with error feedback (random trials contrast) in the right caudate nucleus (x = 10, y = 9, z = 2) and in the right medial orbitofrontal cortex (area 10; x = 4, y = 48, z = 5). In contrast, at that threshold, no brain area was identified in which activity associated with correct responses was greater than activity associated with error responses (fixed trials contrast), nor in which activity associated with error feedback was greater than activity associated with correct feedback (random trials contrast).

Supplementary Methods

Participants

Eleven young adults (4 women), ranging in age from 19 to 30 years (M = 22.9) participated in the experiment. All participants had normal or corrected-to-normal visual acuity, all but one were right-handed, and all were screened for neurological illness (via self-report). They were paid $30, plus a performance-related bonus, as described below. All participants provided written informed consent. This experiment was approved by the institutional review panel at Princeton University.

Experimental Task

Participants engaged in a probabilistic learning task adapted from a previous ERP experiment to an event-related fMRI design [1]. On each trial of the task, participants were presented with an imperative stimulus, were required to respond within 600 ms following the presentation of the stimulus by pressing one of two buttons on a response pad (IFIS 5-button response glove, MRI Devices, Waukesha, WI), and were presented with a feedback stimulus indicating to them that their response was either correct, incorrect, or too late. Participants were informed that each correct response would be rewarded 10 cents, that each incorrect response would be penalized 10 cents, and that each late response would be penalized 20 cents. They were not informed of the appropriate stimulus-response mappings, and were instructed to respond by trial-and-error in such a way that would maximize the total amount of money earned by the end of the experiment. Participants began the task with $1 in bonus money and were awarded the bonus money at the end of the task (about $10).

Each trial began with the presentation of a central fixation cross for 1 s, followed by the

presentation of an imperative stimulus for 0.5 s and a blank screen for 3.5 s. Then, the fixation cross was presented for 1 s, followed by the presentation of the feedback stimulus for 0.5 s and a blank screen for 3.5 s. Thus, the interval between consecutive task-relevant stimuli was 5 s. Imperative stimuli consisted of six randomly chosen images of buildings, animals, musical instruments, and so on. Images of a head of lettuce and of a carrot served as feedback stimuli indicating to the participant that they were rewarded or penalized on that trial; the mappings between reward/punishment and the feedback stimuli were counterbalanced across participants and kept fixed across the experiment. A different feedback stimulus, an image of a cherry, was presented in case the response time exceeded the 600 ms deadline. These feedback images were used in the original ERP experiments from which this task was derived [1,2]. The deadline was introduced to ensure that participants made some errors due to premature responding even after the stimulus-response mappings had been learned. All stimuli were part of a public Corel image library, were scaled to a uniform size so that they subtended approximately 5^{o} x 5^{o}, and were presented in color against a black visual display projected into the scanner.

Throughout the task, two of the six imperative stimuli were mapped to the left button, so that participants were rewarded if they pressed the left button and penalized if they pressed the right button, and two other stimuli were mapped to the right button in a similar fashion (fixed mappings). For the two remaining stimuli, feedback was delivered randomly, irrespective of the given response. As a result, participants were rewarded on 50% of the trials and penalized on 50% of the trials (random mappings). Before the data acquisition phase of the experiment, participants received written instructions and performed two practice blocks of 30 trials outside of the scanner. During the data acquisition phase of the experiment, participants performed 10

blocks of 30 trials in the scanner; each of the six imperative stimuli was presented 50 times in a random order. A new set of six images was chosen for each participant, and they saw only that set of six images throughout the practice and experimental phases of the experiment. At the end of each block, participants were informed about the total amount of bonus money they had earned throughout the experiment.

FMRI image acquisition

Images were collected with a 3.0 Tesla Siemens Allegro head-dedicated MRI scanner. Anatomical images were collected using a T1-weighted MP-RAGE protocol (256 x 256 matrix, FOR = 256 mm, 176 1-mm sagittal slices). Functional images were reconstructed from twenty-five axial slices parallel to the AC-PC plane acquired using a T2*-weighted echoplanar pulse sequence (TR = 1500 ms; TE = 30 ms; flip angle 75°; 192 x 192 cm FOV; 64 x 64 matrix; 3.0-mm isotropic voxels; 1-mm interslice spacing). Image acquisition varied systematically across trials with respect to stimulus onset, yielding an effectively higher temporal sampling rate [3]. Ten functional runs (207 scans each) were collected. The first three scans of each run, recorded before the longitudinal magnetization reached a steady state recovery value, were discarded.

fMRI image analysis

Data were preprocessed and analysed with BrainVoyager software (Maastricht, The Netherlands). Image preprocessing consisted of: rigid-body 3D motion correction using trilinear interpolation; slice scan time correction using sinc interpolation; spatial smoothing with a 4 mm fullwidth at half maximum Gaussian kernel; voxel-wise linear detrending; and high-pass filtering of frequencies below 3 cycles per time course. Spatial normalization was performed using the standard 9-parameter landmark method of Talairach and Tournoux [4]. For each participant, the

blood oxygen-level dependent response across the scanning run was modeled with a general linear model [5] that included ten regressors. Four regressors accounted for imperative stimuli with correct responses and error responses on trials with random mappings and on trials with fixed mappings (correct responses on random mapping trials, error responses on random mapping trials, correct responses on fixed mapping trials, and error responses on fixed mapping trials). Four additional regressors accounted for correct and error feedback on trials with random mappings and fixed mappings (correct feedback on random mapping trials, error feedback on random mapping trials, correct feedback on fixed mapping trials, and error feedback on fixed mapping trials). Finally, two regressors accounted for the fixation period before each stimulus and the rest period following each stimulus. Responses exceeding the deadline and corresponding feedback were rare and were not modeled. The hemodynamic response to each event was estimated by convolving each regressor with a standard gamma function [6]. For each voxel and each event type, a parameter estimate was generated that indicated the strength of covariance between the data and the hemodynamic response function; these estimates were corrected for temporal autocorrelation using a first-order autoregressive model. Pair-wise contrasts between parameter estimates for different events were calculated for each participant, and the results submitted to a group analysis that treated inter-subject variability as a random effect [5]. Statistical parametric maps were derived from the resulting t-values associated with each voxel and were thresholded at an uncorrected p-value (see main text). The location of the peak activity associated with each cluster of activation was reported in Talairach coordinates [4].

We predicted that a caudal and dorsal region of anterior cingulate cortex implicated in the cognitive control of motor behavior [7,8] would be more active on error trials than on correct

trials, regardless of whether the source of the error information was the response or the feedback [1]. However, the anatomy of human medial frontal cortex is highly variable across individuals [9,10], and hence the boundaries of this region are somewhat poorly defined. Moreover, although previous empirical studies have indicated that a broad area of medial frontal cortex is strongly activated by error responses relative to correct responses, with peaks of activity in rostral anterior cingulate cortex, caudal anterior cingulate cortex, and the supplementary motor area (e.g., [11-13]), medial frontal cortex is only weakly activated by error feedback relative to correct feedback (Van Veen et. al. Program No. 16.1 2002 Abstract Viewer/Itinery Planner. Washington, D. C.: Society for Neuroscience 2002. Online), with an apparent peak of activity in dorsal anterior cingulate cortex [14,15]. For these reasons, we identified feedback-related error activity in caudal and dorsal anterior cingulate cortex by defining a region of interest as described in the main text.

References

[1] Holroyd, C. B. & Coles, M. G. H. *Psychol. Rev.* **109**, 679-709 (2002).

[2] Nieuwenhuis, S. *et al. Cogn. Affect. Behav. Neurosci.* **2**, 19-36 (2002).

[3] Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E. & Buckner, R. L. *Neuroimage* **11**, 735-759 (2000).

[4] Talairach, J. & Tournoux, P. *Co-Planar Stereotaxic Atlas of the Human Brain: An Approach to Medical Cerebral Imaging* (Thieme, Stuttgart, Germany, 1988).

[5] Friston, K. J. *et al. Neuroimage* **7**, 30-40 (1998).

[6] Boynton, G. M., Engel, S. A., Glover, G. H. & Heeger, D. J. *J. Neurosci.* **16**, 4207-4221 (1996).

[7] Bush, G., Luu, P. & Posner, M. I. *Trends. Cogn. Sci.* **4**, 215-222 (2000).

[8] Picard, N. & Strick, P. L. *Cereb. Cortex* **6**, 342-353 (1996).

[9] Paus, T. *et al. Cereb. Cortex* **6**, 207-214 (1996).

[10] Vogt, B. A., Nimchinsky, E. A., Vogt, L. J. & Hof, P. R. *J. Comp. Neuro.* **359**, 490-506 (1995).

[11] Kiehl, K. A., Liddle, P. F. & Hopfinger, J. B. *Psychophysiology* **37**, 216-223 (2000).

[12] Menon, V., Adleman, N. E., White, C. D., Glover, G. H. & Reiss, A. L. *Hum. Brain Mapp.* **12**, 131-143 (2001).

[13] Ullsperger, M. & Cramon, D. Y. *NeuroImage* **14**, 1387-1401 (2001).

[14] Ullsperger, M. & von Cramon, D. Y. *J. Neurosci.* **23,** 4308-4314 (2003).

[15] Bush, G. *et al. Proc. Natl. Acad. Sci. U. S. A.* **99**, 523-528 (2002).