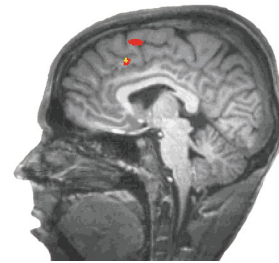


# Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals

Clay B Holroyd<sup>1,6</sup>, Sander Nieuwenhuis<sup>1,2,6</sup>, Nick Yeung<sup>1</sup>, Leigh Nystrom<sup>1,3</sup>, Rogier B Mars<sup>4,5</sup>, Michael G H Coles<sup>4</sup> & Jonathan D Cohen<sup>1,3</sup>



**Figure 1** Sensitivity of dorsal anterior cingulate cortex (dACC) to both internal and external sources of error information. Red: error-related clusters of activation in dACC (ventral cluster) and in the pre-supplementary motor area (dorsal cluster), associated with a contrast between error responses and correct responses on trials with fixed mappings. Yellow: error-related cluster of activation in dACC associated with a contrast between error feedback and correct feedback on trials with random mappings.

**In our event-related functional magnetic resonance imaging (fMRI) experiment, participants learned to select between two response options by trial-and-error, using feedback stimuli that indicated monetary gains and losses. The results of the experiment indicate that error responses and error feedback activate the same region of dorsal anterior cingulate cortex, suggesting that this region is sensitive to both internal and external sources of error information.**

Theories of anterior cingulate cortex (ACC) emphasize a role for this brain area in performance monitoring<sup>1</sup> and action selection<sup>2</sup>. One recent theory proposes that an area in caudal and dorsal ACC (dACC) implicated in the cognitive control of action selection<sup>3,4</sup> uses reward prediction error signals conveyed there by the mesencephalic dopamine system to reinforce adaptive behaviors<sup>2,5</sup>. The theory holds that this neural region is activated by unexpected error information, regardless of whether the source of the information is 'internal' (from an 'efference copy' of the response command) or 'external' (from feedback in the outside environment). Although the results of event-related brain potential (ERP)<sup>2,5,6</sup> and hemodynamic neuroimaging<sup>7–9</sup> studies are consistent with this position, it remains to be determined whether or not a single motor-related area in ACC is sensitive to both types of error information.

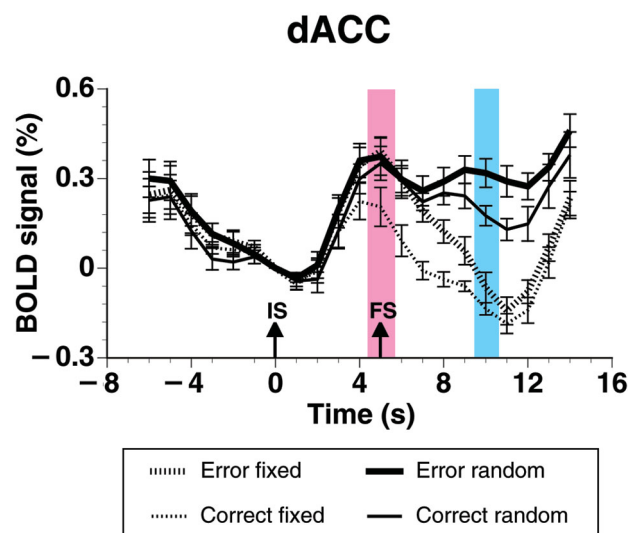
Here we used fMRI to investigate this question. We adapted to an event-related design<sup>10</sup> a probabilistic learning task previously used in ERP experiments to study error activity associated with ACC<sup>2,6</sup>. In this task, participants used feedback stimuli to learn by trial-and-error which of two buttons to press after presentation of imperative stimuli. The feedback stimuli indicated whether the participants received a financial reward (+10 cents) or incurred a financial penalty (–10 cents) on each trial. Some of the imperative stimuli were associated with fixed stimulus-response mappings that could be learned by the participants, and other imperative stimuli were associated with random stimulus-response mappings that could not be learned by the participants. We predicted that dACC would be more active for error trials than for correct trials, whether or not the source of the information was related to self-detected errors following the response or to externally provided error feedback. Furthermore, the theory holds that the error activity should be elicited only by unexpected error events—namely, by the first event on each trial indicating that an error has occurred. On trials with fixed stimulus-response mappings, this event is the error response, because the mappings are known and the system can detect the error at the time of the response. However, on trials with random stimulus-response mappings, this event is the

error feedback, because the mappings are essentially unknown and the system must wait for the error feedback to detect the error. Thus, we predicted that response-related error activity would occur on trials with fixed mappings but not on trials with random mappings, and that feedback-related error activity would occur on trials with random mappings but not on trials with fixed mappings.

Written consent was obtained from all participants. Participants were briefly given some practice at the task and an opportunity to become familiar with the stimulus-response mappings before entering the scanner. Data acquisition and analysis followed commonly adopted procedures (see **Supplementary Methods** online). Participants responded correctly on about 77% ( $\pm 3\%$ ) of the trials associated with fixed mappings. On trials associated with random mappings, correct feedback and error feedback were delivered at random, so the accuracy measure is not meaningful for this condition. About 6% ( $\pm 1\%$ ) of all trials were classified as 'too late' (response time >600 ms).

A statistical parametric map associated with a contrast between error responses and correct responses on trials with fixed mappings was thresholded at a conservative value ( $P < 0.00025$ , uncorrected). This procedure yielded clusters of activity associated with the peak difference in activation between error responses and correct responses (Fig. 1, red areas), including right dACC (area 32:  $x = 1$ ,  $y = 18$ ,  $z = 44$ ), right pre-supplementary motor area (area 6:  $x = 4$ ,  $y = 12$ ,  $z = 59$ ) and left supplementary motor area (area 6:  $x = -5$ ,  $y = -6$ ,  $z = 65$  and  $x = -11$ ,  $y = 0$ ,  $z = 62$ ). In contrast, at this threshold, a comparison between error feedback and correct feedback on trials with random mappings did not reveal any neural areas in which feedback-related error activity was greater than feedback-related correct activity. To identify feedback-related error activity, the dACC cluster was defined as a region of interest (ROI). We predicted that this area would be sensitive to error feedback, that is, that feedback-related error activity would be associated with the same region as the peak response-related error activity. A neural area within this ROI (area 32:  $x = 4$ ,  $y = 18$ ,  $z = 44$ ) was indeed more activated by error feedback than by correct feedback on trials with random mappings ( $P < 0.05$ , uncor-

<sup>1</sup>Department of Psychology, Green Hall, Princeton University, Princeton, New Jersey 08544, USA. <sup>2</sup>Department of Cognitive Psychology, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands. <sup>3</sup>Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, New Jersey 08544, USA. <sup>4</sup>F. C. Donders Centre for Cognitive Neuroimaging, Adelbertusplein 1, 6525 EK Nijmegen, The Netherlands. <sup>5</sup>Nijmegen Institute for Cognition and Information, University of Nijmegen, Box 9104, 6500 HE, Nijmegen, The Netherlands. <sup>6</sup>These authors contributed equally to this work. Correspondence should be addressed to C.B.H. (cholroyd@princeton.edu).



**Figure 2** Event-related averages associated with dorsal anterior cingulate cortex (dACC). Data correspond to the cluster of activation produced by the feedback contrast on trials with random mappings (yellow area in Fig. 1). Arrows indicate time of presentation of the imperative stimulus (IS; 0 s) and of the feedback stimulus (FS; 5 s). Pink and blue areas correspond to the time of the maximum hemodynamic response associated with the response and feedback, respectively. Error bars show standard error of the mean.

rected; Fig. 1, yellow area). Furthermore, the locations of the peak activities in dACC for the response comparison and for the feedback comparison were identical (within the resolution of the experimental design) along the rostral-caudal and ventral-dorsal dimensions of Talairach space ( $y = 18$ ,  $z = 44$ ), and extremely close along the lateral dimension ( $x = 1$  versus  $x = 4$ ).

To confirm these observations, we created event-related averages<sup>10</sup> from the data associated with the cluster of activation corresponding to the feedback contrast, for correct trials and error trials associated with fixed mappings and with random mappings (Fig. 2). A three-factor repeated-measures ANOVA on trial type (fixed, random), event type (response, feedback) and outcome (error, correct) on the event-related averaged data associated with the periods of maximum hemodynamic activity (5 s after each event<sup>11</sup>, colored areas in Fig. 2) revealed main effects of trial type ( $F_{1,10} = 14.5$ ,  $P < 0.005$ ), event type ( $F_{1,10} = 31.5$ ,  $P < 0.001$ ) and outcome ( $F_{1,10} = 10.7$ ,  $P < 0.01$ ); an interaction between trial type and event type ( $F_{1,10} = 49.5$ ,  $P < 0.001$ ); and a three-way interaction between trial type, event type and outcome ( $F_{1,10} = 14.4$ ,  $P < 0.005$ ). The other interactions were not significant ( $P > 0.05$ ). Paired  $t$ -tests on the event-related average data associated with fixed trials indicated that this region was significantly more activated by error responses compared to correct responses,  $t_{10} = 4.0$ ,  $P < 0.005$  (Fig. 2, pink area), but not by error feedback compared to correct feedback,  $t_{10} = 1.6$ ,  $P = 0.13$  (Fig. 2, blue area). Conversely, paired  $t$ -tests on the event-related average data associated with random trials indicated that this region was significantly more activated by error feedback compared to correct feedback,  $t_{10} = 3.9$ ,  $P < 0.005$  (Fig. 2, blue area), but not by error responses compared to correct responses,  $t_{10} = 0.4$ ,  $P = 0.73$  (Fig. 2, pink area). Although not the focus of this study, we note that this area was also activated by both error responses and correct responses on random trials relative to correct fixed trials, a finding that is predicted by the response conflict theory of ACC

(because the random trials are associated with relatively equal response biases, and therefore with a greater probability of response coactivation and conflict)<sup>1</sup>. Activations associated with other contrasts are given in the **Supplementary Note** online.

These results demonstrate that a single area within dACC is more activated both by error responses than by correct responses and by error feedback than by correct feedback, suggesting that this area is involved in processing both sorts of error information. Along the rostral-caudal and ventral-dorsal dimensions of Talairach space, the peak areas associated with the activity were in fact identical. This focus is located in the posterior part of the rostral cingulate zone<sup>4</sup>, exactly on the border of Brodmann areas 32 and 8. It is located very near to regions identified in previous studies as being sensitive to error feedback<sup>8</sup> and to unexpected decreases in reward<sup>9</sup>, and to a region of ACC that is activated by pain<sup>12</sup>—a primal form of error feedback. Furthermore, consistent with the error-related ERP activity observed in this task<sup>2,6</sup>, these error activations were greatest following unexpected error events, that is, after the response on fixed trials and after the feedback on random trials.

dACC is believed to contribute to the cognitive control of motor behavior<sup>3,4</sup>. The activity of this region is modulated by midbrain dopamine neurons<sup>13</sup>, which convey reinforcement signals that indicate errors in reward prediction<sup>14</sup>. A recent theory proposes that dACC uses these performance-related dopamine signals to select and reinforce adaptive behaviors<sup>2,5</sup>. The theory further posits that these dopamine signals modulate dACC activity such that unexpected unfavorable events are associated with more activity than unexpected favorable events. Until now, this position has been supported primarily by ERP data that, because of the inverse problem, have provided relatively poor information about the location of the underlying neural sources. The results of the present study provide converging evidence from a complementary experimental technique with superior spatial resolution.

*Note: Supplementary information is available on the Nature Neuroscience website.*

#### ACKNOWLEDGMENTS

The authors thank P. Hu, J. Aronson, P. Hutter, M. Richter and A. Engell for technical assistance. This research was supported by National Institute of Mental Health grant MH064445 and postdoctoral fellowship MH63550. The research of S. Nieuwenhuis was supported by the Netherlands Organization for Scientific Research.

#### COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 20 November 2003; accepted 3 March 2004

Published online at <http://www.nature.com/natureneuroscience/>

1. Botvinick, M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. *Psychol. Rev.* **108**, 624–652 (2001).
2. Holroyd, C.B. & Coles, M.G.H. *Psychol. Rev.* **109**, 679–709 (2002).
3. Paus, T. *Nat. Rev. Neurosci.* **2**, 417–424 (2001).
4. Picard, N. & Strick, P.L. *Cereb. Cortex* **6**, 342–353 (1996).
5. Holroyd, C.B., Nieuwenhuis, S., Mars, R.B. & Coles, M.G.H. in *Cognitive Neuroscience of Attention* (ed. Posner, M.) (Guilford Publications, New York, in press).
6. Nieuwenhuis, S. *et al. Cogn. Affect. Behav. Neurosci.* **2**, 19–36 (2002).
7. Carter, C.S. *et al. Science* **280**, 747–749 (1998).
8. Ullsperger, M. & von Cramon, D.Y. *J. Neurosci.* **23**, 4308–4314 (2003).
9. Bush, G. *et al. Proc. Natl. Acad. Sci. USA* **99**, 523–528 (2002).
10. Dale, A.M. & Buckner, R.L. *Hum. Brain Mapp.* **5**, 329–340 (1997).
11. Boynton, G.M., Engel, S.A., Glover, G.H. & Heeger, D.J. *J. Neurosci.* **16**, 4207–4221 (1996).
12. Ploghaus, A. *et al. Science* **284**, 1979–1981 (1999).
13. Allman, J.M., Hakeem, A., Erwin, J.M., Nimchinsky, E. & Hof, P. *Ann. NY Acad. Sci.* **935**, 107–117 (2001).
14. Schultz, W. *Neuron* **36**, 241–263 (2002).