When the going gets tough, the cingulate gets going

William J Gehring and Stephan F Taylor

Despite extensive study, the exact role of human anterior cingulate cortex in forming behavioral strategies is still controversial. In this issue, Williams and colleagues describe a rare opportunity to test a hypothesis about the function of this brain area in humans, by not only recording from single neurons, but also monitoring behavior immediately after surgical ablation of the area.

Any motivational speaker will agree that the key to success lies not in avoiding mistakes, but in recognizing when things have gone wrong and taking the appropriate steps to recover. Neuroimaging studies of humans performing a wide variety of cognitive and motor tasks point increasingly to a role for the dorsal anterior cingulate cortex (dACC) in detecting when unfavorable circumstances call for a change in behavioral strategy[1]. Activity in dACC increases when subjects make errors[2,3], experience response conflict[4], and lose money in simple decision-making and learning tasks[5,6]. Why this inconspicuous piece of cortex should be so ubiquitous in controlling human behavior is still a mystery, but one common theme that ties together most of these functions is the need for a signal that something unwanted has happened, which tells the subject to switch to a new course of action.

In this issue[7], Williams and colleagues describe a rare opportunity to record from single human dACC neurons during a reward-processing task, and the even rarer opportunity to observe the behavioral effects of dACC ablation in the same individuals. Despite the undoubtedly stressful procedure of surgery, five patients preparing to undergo cingulotomy (ablation of the dACC), a last-resort treatment to alleviate their severe depression or obsessive-compulsive disorder, agreed to participate in the experimental study. Immediately before the ablation, the authors passed recording electrodes through burr holes made in the cranium in preparation for the procedure and recorded from single neurons while the patients performed a task designed to reveal the workings of the dACC (**Fig. 1**).

William J. Gehring is in the Department of Psychology and Stephan F. Taylor is in the Department of Psychiatry, University of Michigan, Ann Arbor, Michigan, USA.
e-mail: wgehring@umich.edu

Non-invasive studies in humans have shown that a loss or reduction of reward can cause ACC activity[5,6,8]. Williams and colleagues hypothesized that the dACC is critical for linking the information that a loss of reward has occurred with the decision to change behavior. To test their hypothesis, they asked the subjects to perform a simple movement task for a small reward under shifting reward conditions. On each trial (**Fig. 2**), subjects had to move a joystick to the left or right according to the instruction displayed on a computer screen. On 80% of the trials, participants were shown five dollar signs ("$$$$$"), which signified a 15-cent reward and told the subjects to move the joystick in the same direction as they had on the preceding trial. In the other 20% of the trials, the stimulus instructed subjects to move the joystick in the direction opposite the previous response. On half of these trials, the stimulus consisted of two arrowheads, which denoted that the subjects would still receive 15 cents. On the remaining trials the stimulus consisted of three dollar signs ("$$$"), signifying a reduction in reward (to 9 cents). Of the 134 neurons recorded across the five subjects, 39% were affected by the instruction to change direction, with the majority increasing their activity when the reduced-reward cue instructed subjects to change response direction.

It is one thing to show that a task can cause activity in some part of the brain; it is quite another to show why that activity is important for producing behavior. To explore the causal role of the dACC in changing behavioral responses, Williams and colleagues took advantage of the small proportion of error trials, finding that dACC activity in response to the reduced-reward cue predicted whether the subject would subsequently respond correctly. This relationship did not hold for activity in response to either of the other instructional cues, suggesting that it was specifically related to using the reduced-reward information to switch to a new response.

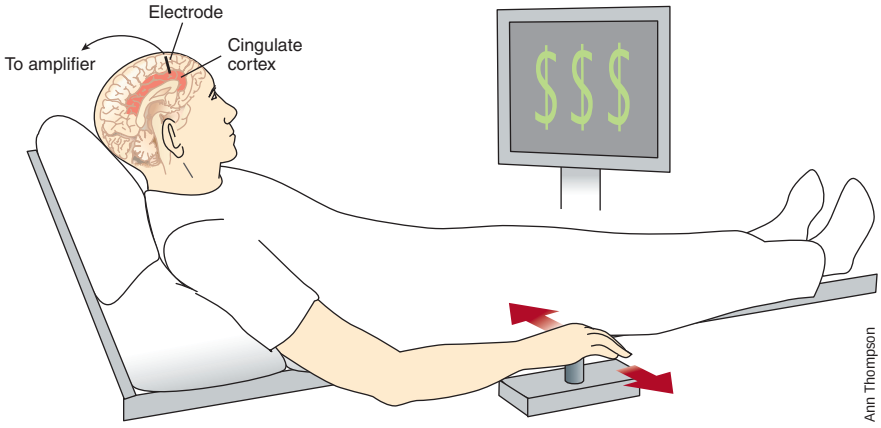

Figure 1 Experimental setup of Williams *et al*. Patients preparing to undergo therapeutic cingulotomy performed a reward processing task in the surgery suite while the authors recorded single-cell activity from dorsal anterior cingulate. The dACC is located within the shaded area, approximately where the electrode is depicted.

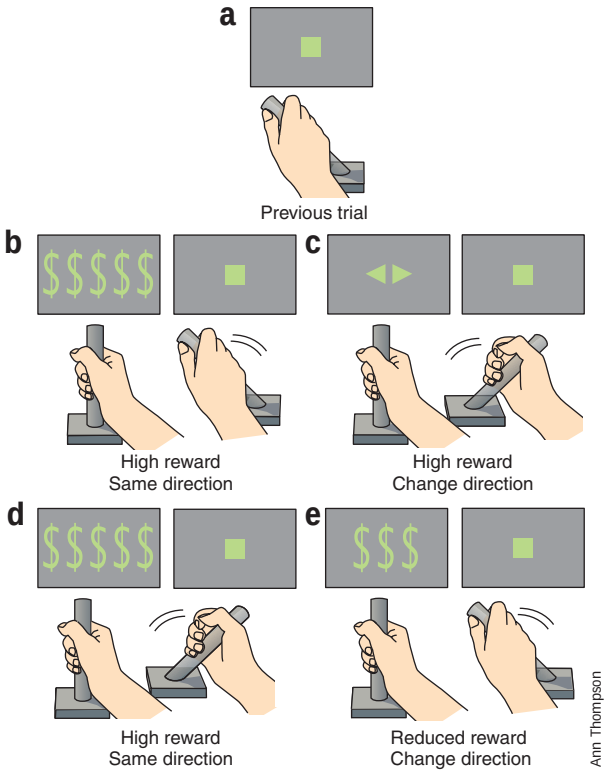

Figure 2 Example trial sequence. Subjects moved a joystick to either the left or right, depending on the instructional stimulus, moving in the same direction as the previous trial (**a**, **b**, **d**) in response to the high reward signal, and changing direction either for the same high reward (**c**) or for reduced reward (**e**).

The most convincing evidence that those neurons were important in this task, however, came when they were no longer available to do their job. The subjects repeated the task within 30 minutes of the end of the surgery, allowing the authors to observe the effects of the ablation. And this is where the study wraps its findings into a tight package: after the dACC was damaged, the subjects were impaired at changing their responses, with the greatest deficits when the cue specified a reduction in reward.

These results support a theory of dACC activity that has been proposed to explain why the dACC is most active in demanding tasks when negative events such as errors, response conflict or losses of reward indicate that something has gone wrong. The theory states that when events occur that are worse than expected, an error signal triggers dACC activity that modifies response strategies to avoid future negative outcomes[5]. The dACC might also modulate autonomic arousal to help behavior meet task demands[9]. Consistent with these ideas, instructional cues in this study that called for change and signaled loss led to maximal dACC activity and produced the strongest behavioral effects following dACC lesions.

In addition to its importance for understanding the dACC, the Williams study provides a remarkable window into human brain function, only rarely available to scientists. Investigators commonly record from single neurons in animals, but extrapolating to the human brain has always been problematic. Studies of humans have depended on indirect techniques, such as functional magnetic resonance imaging (fMRI), which relies on hemodynamic changes, or event-related potentials (ERPs), which rely on electrical scalp potentials. These techniques are limited by poor temporal (fMRI) or spatial (ERP) resolution. Studies of individuals with brain injuries provide valuable information about brain areas necessary for certain functions, but these 'experiments of nature' are hampered by a lack of control over the location of damage and by an inability to compare the individuals' behavior before and after damage. Williams and colleagues have not only surmounted these limitations in a single study, but they have also gone a step further and shown how the physiological activity of the dACC relates to behavior.

Although an experimental *tour de force*, the study has several limitations. It is doubtful that a patient undergoing surgery will perform with the same the same level of concentration and with the same motivations as typical participants in most psychology experiments—usually healthy college students in quiet laboratories. In the surgical setting, one must take special measures to confirm that the subjects interpreted the task and performed it in the way intended by the experimenters. In this case, a key assumption is that subjects find the 15-cent reward desirable and the 6-cent reduction undesirable. This assumption is not implausible, but one wonders how much the difference between a 15-and 9-cent reward concerns someone lying in an operating room in the midst of brain surgery. Certainly the subjects' behavior shows that they cared about the task, but there is no behavioral evidence that they cared about the rewards.

An important consideration in interpreting this study is that the brain area of interest is the same area that must be destroyed to alleviate the patients' psychopathological symptoms. One must be concerned that these findings might be misleading because of possible abnormalities in the dACC before surgery. Williams and colleagues consider this possibility, however, pointing out that consistencies between their experiment and other work (fMRI studies in healthy humans, as well as single-unit recordings in monkeys) suggest that their findings were not significantly affected by psychopathology.

The effects of the ablation also raise some questions: it is puzzling that despite the apparent centrality of dACC function to behavior, cingulotomy patients exhibit very little impairment of cognitive and motor function following the surgery[10]. Williams and colleagues measured behavioral effects only within minutes of dACC ablation, leaving open the possibility that compensatory changes could enable other brain regions eventually to take over tasks carried out by the dACC. These considerations suggest that theories of dACC function will need to explain both the behavioral impairments seen in the reward processing task and the clinical improvements and relatively normal behavior that follow cingulotomy.

Beginning with the groundbreaking work of Wilder Penfield[11] in the 1950s, there is a long history of studying the brain in awake humans undergoing neurosurgical procedures. Building on this tradition, the Williams study is an excellent illustration of the advances that are possible when direct intraoperative recordings are combined with an elegant experimental procedure for analyzing how physiology and brain lesions influence behavior. Obviously, ethical considerations limit the use of intraoperative recordings in humans. Patients with a psychiatric illness severe enough to warrant this extraordinary treatment may be particularly vulnerable to exploitation in the research setting. Careful

protocols, such as the one used in this study, need to ensure that patients who participate in the research do so without any feeling of coercion or expectation of benefit. Conducted with appropriate care and precaution, studies like this one will be critical in supplementing the more widely available non-invasive tools of cognitive neuroscience.

1. Ridderinkhof, K.R., Ullsperger, M., Crone, E.A. & Nieuwenhuis, S. *Science* **306**, 443–447 (2004).
2. Miltner, W.H.R., Braun, C.H. & Coles, M.G.H. *J. Cogn. Neurosci.* **9,** 788–798 (1997).
3. Ullsperger, M. & von Cramon, D.Y. *Cortex* **40**, 593–604 (2004).
4. Carter, C.S. *et al. Science* **280**, 747–749 (1998).
5. Holroyd, C. & Coles, M.G.H. *Psychol. Rev.* **109**, 679–709 (2002).
6. Gehring, W.J. & Willoughby, A.R. *Science* **295,** 2279–2282 (2002).
7. Williams, Z.M., Bush, G., Rauch, S.L., Cosgrove, G.R. & Eskandar, E.N. *Nat. Neurosci.* **7,** 1370–1375 (2004).
8. Bush, G. *et al. Proc. Natl. Acad. Sci. USA* **99**, 523–528 (2002).
9. Critchley, H.D. *et al. Brain* **126,** 2139–2152 (2003).
10. Kim C.H. *et al. Acta Psychiatry Scand.* **107**, 283–290 (2003).
11. Penfield, W. & Jasper, H. *Epilepsy and the Functional Anatomy of the Human Brain* (Little and Brown, Boston, 1954).

Multiple routes to similar network output

Scott L Hooper

Neuronal networks are built from neurons with different properties and from synapses of different strengths. Modeling suggests that networks can tune these parameters to many different combinations that nonetheless produce very similar network outputs.

What would happen if we measured the ability of humans to write with their left and right hands and averaged the results without accounting for the population's bimodality? Rather than identifying the true case that 10% of people write exclusively with their left hand and 90% exclusively with their right, we would report that people write 90% of the time with their right hand and 10% with their left.

Characterizing every conductance and synaptic strength within a neuronal network in a single experiment is generally impossible. Instead, experimentalists often make a few measurements in the 'same' neuron or synapse from multiple animals, then repeat the process for another set of measurements in a different group of animals and eventually obtain multiple measurements of all the network's conductances and synaptic strengths. These measurements are then typically reported as an average ± standard deviation or error. This procedure is perfectly acceptable if the neurons and networks that are said to be the 'same' have similar conductances and synaptic strengths in different individuals, and if variations in one parameter are uncorrelated with variations in other parameters. However, it does not account for the possibility that neurons or networks with different conductances or synaptic composition might be able to produce the same activity if changes in one conductance or synapse were compensated by changes in other conductances or synapses. Eve Marder and her co-workers have investigated this hypothesis extensively in single neurons, both experimentally[1] and computationally[2–4]. In this issue, Prinz *et al.*[5] now show that model neural networks with different combinations of intrinsic neuronal properties and synaptic strengths can produce extremely similar outputs (**Fig. 1**). The neurons or networks that are compared over different individuals might therefore be the same only with respect to their output, not their underlying makeup.

Prinz *et al.* simulated 20,250,000 model neural networks consisting of three neurons, each of which existed in five or six different intrinsic-property versions and could be interconnected with five or six different synaptic strengths. They then identified which of these models produced outputs resembling the triphasic bursting activity of the well-known decapod crustacean pyloric network. Remarkably, even with fairly stringent selection criteria based on 15 measures of experimental spiking activity, over 2% (452,516) of the models produced correct ('pyloric') outputs. Even more remarkably, models producing pyloric output could be built with all six model-neuron types and with synapses spanning the entire range of conductance values (with the exception of one synapse that needed to be weak). This suggests that compensatory, function-maintaining changes in cellular and synaptic properties can occur in a graded fashion.

Such graded correlations would be much less obvious experimentally than simple cases of bimodality—where, for example, a low value of current Y is associated with a strong synapse X and a high value of Y with a weak synapse X.=w The Prinz *et al.* work thus indicates that experimentalists must re-examine their data to test whether their standard deviations actually represent graded, correlated changes of neuronal and network properties. Wide current-density variations (up to threefold)[1] and synaptic-strength variations (with standard deviations as large as ±100% of the mean)[6,7] have been experimentally observed in the pyloric network. Cross-correlations among the measured parameters were not calculated, and it is therefore not

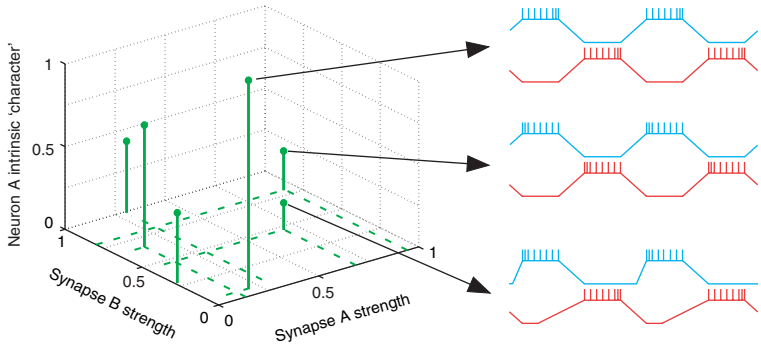

Figure 1 Example illustrating how different combinations of synaptic strength and intrinsic neuron properties could produce outputs that have the same cycle period, spike number and phase relationships (albeit with fine differences in action potential timing and slow-wave trajectories). The system modeled by Prinz *et al.* had nine dimensions (three neurons whose intrinsic properties could vary, and six synapses).

The author is in the Department of Biological Sciences, Ohio University, Athens, Ohio 45701, USA.
e-mail: hooper@ohio.edu