Learning by doing: an fMRI study of feedback-related brain activations

Josep Marco-Pallarés, Sandra Verena Müller and Thomas F. Münte

Department of Neuropsychology, University Otto von Guericke, Magdeburg, Germany

Correspondence to Dr Thomas F. Münte, MD, Institut für Psychologie II, Department of Neuropsychology, University Otto von Guericke, Universitätsplatz, Gebäude 24, Postfach 4120, 39106 Magdeburg, Germany
Tel: +49 391 6718469; fax: +49 391 6711947; e-mail: thomas.muente@med.uni-magdeburg.de

Received 14 May 2007; accepted 31 May 2007

Extracting meaningful information from the positive and negative outcomes of actions is a key requirement for learning. To define the neural correlates of feedback processing, rapid event-related functional magnetic resonance imaging was used in an associative learning paradigm in normal human volunteers. Positive (compared with negative) feedback was associated with activations in the ventral striatum, midbrain and anterior and posterior cingulate cortex. No activations were seen for the comparison negative > positive feedback. Blood oxygenation level-dependent responses from the midbrain and the anterior cingulate cortex showed a phasic increase in response to positive feedback, whereas a decrease in response was seen for negative feedback. These results underscore the role of the reward system in feedback learning. *NeuroReport* 18:1423–1426 © 2007 Lippincott Williams & Wilkins.

Keywords: feedback, functional magnetic resonance imaging, learning, reward

Introduction

The adaptation of behaviour based on the results of actions is a prerequisite for the refinement of actions and plans, and for avoiding errors. To determine whether actions have been performed correctly, external information about their results can be used to assess the appropriateness of behaviour. Applying such external information in the form of positive and negative feedback in a systematic fashion, as in conditioning experiments, will lead to the changes in behaviour that is learning. This learning process has been studied in animals, leading to the observation of phasic bursts of dopaminergic activity originating in the midbrain during positive reinforcement [1]. These bursts lead to the learning of rewarding behaviours and thus act as a teaching signal [1,2]. Conversely, negative feedback is associated with dopamine dips that drop below the baseline [1,3]. These neurophysiological findings have been incorporated into computational models of the basal ganglia–dopamine system in reinforcement learning [4].

How feedback is processed in the brain is, hence, a key point in understanding learning processes. A number of studies have addressed this issue using event-related brain potentials (ERPs) and functional magnetic resonance imaging (fMRI) in humans. In ERP studies, typically a frontocentral negativity associated with negative feedback has been observed to appear 250–400 ms after the presentation of the feedback signal, which probably has generators in the anterior and posterior cingulate cortex [5,6]. No specific ERP response has, however, been described for positive feedback. In contrast, most of the fMRI studies on feedback processing have found activations in anterior and posterior cingulate cortex, orbitofrontal cortex and striatum [6–10] in response to positive feedback, but fewer studies have found activations in response to negative feedback [11]. The aim of this study is therefore to reassess the neural correlates of positive and negative feedback during associative learning. Specifically, we were interested in the question of whether or not we could detect an fMRI correlate of the dopaminergic bursts and dips seen in animal experiments following positive and negative feedback, respectively. To reach this goal we performed rapid event-related fMRI during a learning paradigm, in which participants had to learn stimulus–response associations that were based on the information provided by external feedback.

Materials and methods

Participants

Twelve right-handed healthy volunteers (eight women, age range 19–31 years, mean age 23.5 years) participated in the study after giving their written consent. None of them had a history of neurological or psychiatric disorders. The experiment followed the Helsinki protocol and was approved by the ethical committee of the University of Magdeburg.

Experimental procedure

For each experimental run, eight different black-and-white images of animals or objects were presented 12 times each in a random order for a duration of 500 ms. Of these, four pictures were associated with a left-hand response, whereas the remaining four required a right-hand response. The participant was required to make a speeded button press and to determine the correct response for each picture from the response feedback. Following an interval of 1100 ms after the onset of the picture, feedback was given by presenting either a

blue X ('correct', positive feedback) or a red X ('error', negative feedback). In addition, on those trials in which the participant had failed to respond within 800 ms after the onset of the picture a turquoise 0 was shown; this signalled that the response had been too slow. The 'too slow' feedback had priority over the other types of feedback stimuli. This response deadline was imposed on the participants, as time pressure is known to increase the rate of erroneous responses. The total experiment consisted of six blocks containing different pictures. Each block took 4 min to complete.

Magnetic resonance imaging scanning methods

Imaging was performed with a GE Medical Systems 1.5-T Signa Neurovascular MR scanner (Magdeburg, Germany) equipped with a standard quadrature head coil. Visual images were back-projected onto a screen by a light-emitting diode projector and participants viewed the images through a mirror on the head coil. Two magnet-compatible response boxes (one in each hand) were used, containing two response keys each (middle finger and forefinger). Response times and responses were recorded for subsequent analyses. Conventional high-resolution structural images (rf-spoiled GRASS sequence, 60-slice sagittal, 2.8-mm thickness) were followed by functional images sensitive to blood oxygenation level-dependent (BOLD) contrast [echo planar T2*-weighted gradient echo sequence, repetition time (TR)/echo time (TE)/flip angle=2000 ms/40 ms/90°]. Each functional run consisted of 120 sequential whole-brain volumes consisting of 16 axial slices aligned to the plane intersecting the anterior and posterior commissures, of 3.125-mm in-plane resolution and 7-mm thickness, with a 1-mm gap between slices, positioned to cover the entire brain. Volumes were acquired continuously and the two first volumes were discarded owing to T1 equilibration effects. To allow precise coregistration of functional data, a separate T1-weighted two-dimensional spin echo image was acquired in the same slice orientation as the functional scans covering the whole volume.

Data analysis

All analyses were performed using Brain Voyager QX software (Brain Innovation B.V., Maastricht, The Netherlands). Preprocessing steps included three-dimensional motion correction, slice scan time correction and temporal smoothing. The data was then coregistered and normalized to Talairach stereotactic space [12]. We performed random-effects analyses on the *z*-transformed functional data by using two different regressors (positive and negative feedbacks) in the general linear model. Statistical maps were created using a threshold of $P<0.001$ (uncorrected for multiple comparisons) with a cluster threshold of 20 voxels.

Results

Behavioural results

Participants answered correctly/incorrectly in 55.9 ± 8.5 and $40.3\pm7.3\%$ [$t(12)=4.19$, $P<0.005$] of the trials, respectively. Responses exceeding the response deadline were seen in $3.8\pm4.0\%$ of the trials. No significant differences were found between the mean reaction times for correct (598 ± 79 ms) and erroneous responses [598 ± 74 ms; $t(12)=0.11$, $P>0.5$]. Figure 1 illustrates the cumulative positive and negative feedback responses over a run averaged across participants and runs. Clearly, learning of the correct stimulus–response associations occurred.

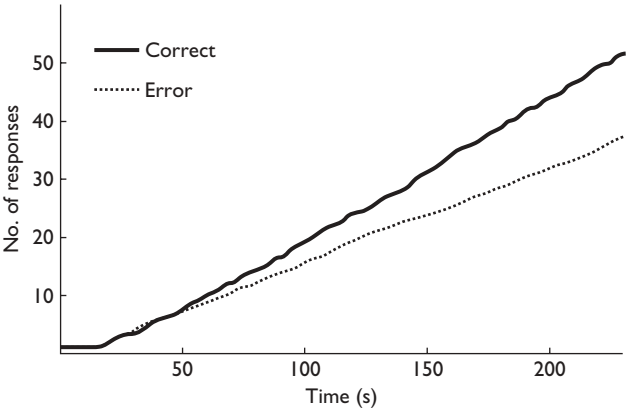

Fig. 1 Mean total number (averaged over blocks and participants) of correct (solid line) and incorrect (dashed line) responses over the course of a block. Although at the beginning of each block the number of correct and incorrect responses is similar, participants learn the association between response and picture over time (see Results).

Table 1 Brain areas presenting significant activity in the contrast positive > negative feedback trials ($P<0.001$, uncorrected for multiple comparisons)

Label	BA	*x*	*y*	*z*	Max *T*
Right cerebellum		33	−72	−30	6.00
Right cerebellum		16	−79	−15	4.18
Right globus pallidus/putamen		19	6	−4	5.12
Posterior cingulate	23	−4	−23	33	5.38
Anterior cingulate	24	−6	29	21	4.05
Left nucleus accumbens		−8	7	−3	4.68
Midbrain		−4	−30	9	4.71
Anterior cingulate cortex	24	−7	4	30	5.34
Left globus pallidus/putamen		−13	4	−2	6.30
Left inferior temporal	37	−46	−55	−14	4.55
Left middle temporal	21	−55	−52	8	6.26

x, *y*, *z*, coordinates given in Tailarach space (approximate). BA, Brodmann area.

Functional magnetic resonance imaging results

Areas that showed a significant increase in the activity of the positive feedback compared with that in negative feedback trials are summarized in Table 1. These included the right and left putamen, left nucleus accumbens (identified using the map published in Ref. [13]), cerebellum, midbrain, anterior and posterior cingulate cortex and left middle and inferior temporal cortex. For the reverse contrast (negative > positive feedback trials) no significant activations were found. Figure 2 illustrates some of the activated brain areas and, in addition, the temporal course of the BOLD response. Interestingly, anterior cingulate cortex (ACC) and midbrain areas showed an increase in the positive and a decrease in the negative feedback trials. To determine if the decrease for negative feedback was significant, a Student's *t*-test between the individual minimum of the BOLD response in the negative feedback condition and baseline was performed, revealing significant results for both areas [midbrain, $t(11)=3.69$, $P<0.01$; ACC, $t(11)=2.86$, $P<0.05$].

To further assess the relationship of the observed brain activations with learning behaviour, we identified in each participant three blocks that showed more efficient learning of the stimulus–response associations, and three that showed less efficient learning. Changes in BOLD response (relative to baseline, average activity at time points 6 and 8 s)

were measured for the midbrain and ACC regions and entered into an analysis of variance with learning (better blocks vs. worse blocks) and feedback (positive vs. negative) as factors. A learning × feedback interaction was identified for the midbrain [$F(1,11)=12.7$, $P<0.005$] suggesting that learning efficacy was related to activity changes in the midbrain (see Fig. 3). While a similar pattern was observed in the ACC, the learning × feedback interaction was not found to be significant in this region [$F(1,11)=0.47$].

Discussion

In this study, we assessed brain activations associated with the processing of positive and negative feedbacks in a learning paradigm. Several areas were found to be active when responses to positive feedback trials were contrasted with those to negative trials. These areas included the anterior and posterior cingulate cortex, the ventral striatum (nucleus accumbens) and putamen, as well as the midbrain. Importantly, these structures have been identified as key players of the dopaminergic reward system and coincide with the brain areas found in response to positive feedback, reward and reward anticipation in previous fMRI studies [6,14–16]. By contrast, and in line with previous fMRI studies, no activations were found in the reverse comparison (negative > positive feedback trials). Inspection of the BOLD time courses, however, revealed a decrease of activation in response to negative feedback in the ACC and the midbrain.

The involvement of dopaminergic brain areas in learning is well described. Animal findings and computational models suggest that positive feedback should facilitate the associated responses by increasing the activity of dopaminergic neurons. This activity is thought to act as the 'teaching signal' [17]. In contrast, negative feedback is associated with a phasic decrease or dip in the midbrain dopaminergic activity, which is further projected to the midfrontal cortex and the ventral striatum. The fMRI is obviously blind to the neurotransmitter changes, however, the brain areas revealed in this study clearly suggest an involvement of the dopaminergic system. Recently, Frank *et al.* [4] described how altered dopamine levels in Parkinson's disease modulate the way people are able to learn from positive or negative reinforcements.

An interesting finding of this study is the absence of specific activations as response to negative feedback. This lack of activation has been described before in fMRI studies [6,18] and is surprising in the face of multiple electroencephalogram studies that have revealed frontocentral negativities in response to negative feedback signals [5,19].

Taking a different approach, Rodriguez *et al.* [20] used a version of the Rescorla–Wagner model to generate prediction errors in a probabilistic classification task with purely cognitive feedback. In their study, activation in the nucleus accumbens increased parametrically with prediction error for negative feedback. By contrast, Yacubian *et al.* [16], employing a guessing task, found that ventral striatum/nucleus accumbens activation scaled with the gain-related part of the computed expected value during an anticipation period. Moreover, this structure also showed activation at the time the outcome of the guessing task was presented; this activation varied as a function of the prediction error (computed as the difference between actual outcome and expected gain). Interestingly, loss-related expected value

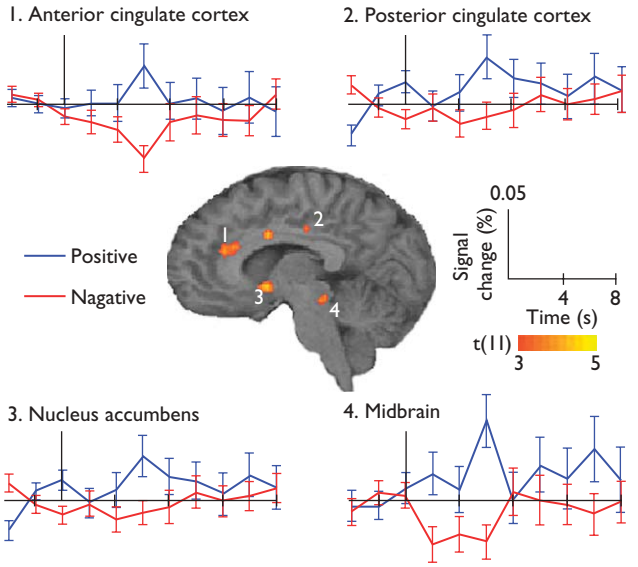

Fig. 2 Sagittal brain slice showing regions of greater activity in positive feedback compared with that in negative feedback trials. The time course of the blood oxygenation level-dependent (BOLD) response in the two conditions is shown for four brain areas. Interestingly, a decrease of the BOLD response is observed for negative feedback trials in the anterior cingulate cortex (ACC) and the midbrain.

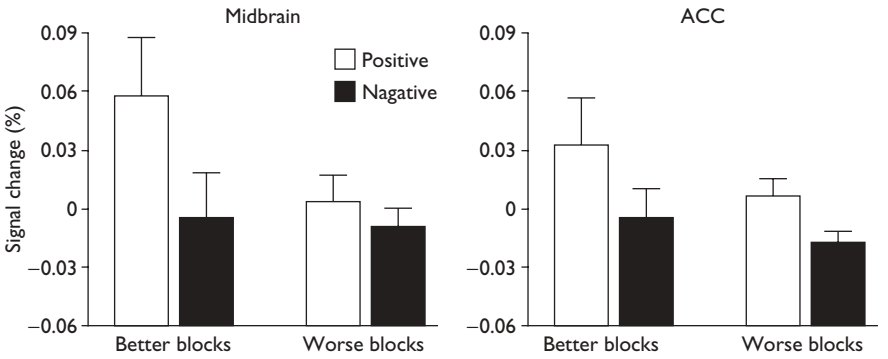

Fig. 3 Blood oxygenation level-dependent (BOLD) changes (average activity for scans 6 and 8 s after the stimulus) relative to baseline in the anterior cingulate cortex (ACC) and the midbrain (coordinates correspond to those in Fig. 2).

and the associated prediction error were represented in the amygdala in the Yacubian *et al.* [16] study. This study used a fast event-related design; therefore, we are unable to resolve activations related to the anticipation of feedback vs. activation as a response to the feedback. Moreover, a purely cognitive feedback was used.

As a novel finding, however, we found two areas (midbrain and ACC) that presented a bidirectional response, that is an increase in activation for positive feedback and a decrease in activation for negative feedback trials. Moreover, the activity difference between negative and positive feedback trials in the midbrain was related to learning efficacy (Fig. 3). Recent studies have shown that some frontal [21] and subcortical areas [16,22,23] increase their activity after reward, whereas they show a decrease with punishment. The current results echo the above findings and extend them to the midbrain. It is intriguing to speculate that this bidirectional response is a correlate of the phasic increase/decrease in dopaminergic activity in the midbrain after positive/negative feedback, respectively; thus it is the brain-imaging manifestation of the teaching signal during the learning process.

Acknowledgements

T.F.M has been financially supported by various grants from the German Research Agency (DFG), the Volkswagen-Stiftung and the BMBF (01GO0202).

References

1. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997; **275**:1593–1599.
2. Pan WX, Schmidt R, Wickens JR, Hyland BI. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward–learning network. *J Neurosci* 2005; **25**:6235–6242.
3. Satoh T, Nakai S, Sato T, Kimura M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci* 2003; **23**: 9913–9923.
4. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 2004; **306**:1940–1943.
5. Müller SV, Möller J, Rodriguez-Fornells A, Münte TF. Brain potentials related to self-generated and external information used for performance monitoring. *Clin Neurophysiol* 2005; **116**:63–74.
6. Nieuwenhuis S, Slagter HA, von Geusau NJA, Heslenfeld DJ, Holroyd CB. Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *Eur J Neurosci* 2005; **21**:3161–3168.
7. Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 2001; **30**:619–639.
8. Elliott R, Newman JL, Longe OA, Deakin JF. Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 2003; **23**:303–307.
9. Knutson B, Westdorp A, Kaiser E, Hommer D. FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 2000; **12**:20–27.
10. Nieuwenhuis S, Heslenfeld DJ, von Geusau NJA, Mars RB, Holroyd CB, Yeung N. Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage* 2005; **25**:1302–1309.
11. Ullsperger M, von Cramon DY. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J Neurosci* 2003; **23**:4308–4314.
12. Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain: an approach to medical cerebral imaging*. New York: Thieme Medical Publishers; 1988.
13. Sturm V, Lenartz D, Koulousakis A, Treuer H, Herholz K, Klein JC, Klosterkötter J. The nucleus accumbens: a target for deep brain stimulation in obsessive-compulsive and anxiety disorders. *J Chem Neuranat* 2003; **26**:293–299.
14. Dreher JC, Kohn P, Berman KF. Neural coding of distinct statistical properties of reward information in humans. *Cereb Cortex* 2006; **16**: 561–573.
15. Knutson B, Taylor J, Kaufman M, Peterson R, Glover G. Distributed neural representation of expected value. *J Neurosci* 2005; **25**:4806–4812.
16. Yacubian J, Glascher J, Schroeder K , Sommer T, Braus DF, Büchel C. Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci* 2006; **26**:9530–9537.
17. Holroyd CB, Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002; **109**:679–709.
18. Van Veen V, Holroyd CB, Cohen JD, Stenger VA, Carter CS. Errors without conflict: implications for performance monitoring theories of anterior cingulate cortex. *Brain Cogn* 2004; **56**:267–276.
19. Miltner WHR, Braun CH, Coles MGH. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a 'generic' neural system for error detection. *J Cogn Neurosci* 1997; **9**:788–798.
20. Rodriguez PF, Aron AR, Poldrack RA. Ventral striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *J Neurosci* 2006; **27**:306–313.
21. Tom SM, Fox CR, Trepel C, Poldrack RA. The neural basis of loss aversion in decision-making under risk. *Science* 2007; **315**:515–518.
22. Jensen J, McIntosh AR, Crawley AP, Mikulis DJ, Remington G, Kapur S. Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 2003; **40**:1251–1257.
23. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. *Neuron* 2003; **38**:329–337.