NeuroImage 64 (2013) 1–9

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Reinforcement learning signals in the anterior cingulate cortex code for others' false beliefs

M.A.J. Apps [a,*], R. Green [a,b], N. Ramnani [a]

[a] Dept. of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK
[b] School of Psychology, The University of Birmingham, Edgbaston Birmingham, B15 2TT, UK

ARTICLE INFO

Article history:
Accepted 4 September 2012
Available online 13 September 2012

Keywords:
Cingulate
Social
Reinforcement
False belief
Prediction error

ABSTRACT

The ability to recognise that another's belief is false is a hallmark of our capacity to understand others' mental states. It has been suggested that the computational and neural mechanisms that underpin learning about others' mental states may be similar to those that underpin first-person Reinforcement Learning (RL). In RL, unexpected decision-making outcomes constitute prediction errors (PE), which are coded for by neurons in the Anterior Cingulate Cortex (ACC). Does the ACC signal the PEs (false beliefs) of others about the outcomes of their decisions? We scanned subjects using fMRI while they monitored a third-person's decisions and similar responses made by a computer. The outcomes of the trials were manipulated, such that the actual outcome was unexpectedly different from the predicted outcome on 1/3 of trials. We examined activity time-locked to privileged information which indicated the actual outcomes only to subjects. Activity in the gyral ACC was found when the outcomes of the third-person's decisions were unexpectedly positive. Activity in the sulcal ACC was found when the third-person's or computer's outcomes were unexpectedly positive. We suggest that a property of the ACC is that it codes PEs, with a portion of the gyral ACC specialised for processing the PEs of others.

Introduction

Successfully interacting in social environments depends upon the ability to understand that others can have mental states which are distinct from one's own (Frith and Frith, 2003). Understanding that another's belief is false is widely regarded as one of the most sensitive and reliable measures of whether an individual is able to understand another's mental state (Hughes et al., 2000; Perner and Lang, 1999; Wellman et al., 2001) and is a fundamental component of cooperative and competitive behaviours. The ability to detect that another's belief is false can be tested for using the 'Sally–Anne' test devised by Wimmer and Perner (1983). During this task a subject reads a story, receiving privileged information that one of the protagonists does not. This scenario enables the subject to determine whether the belief of that protagonist is true or false. If the subject can process others' mental states, they can identify that the protagonist's belief is both different from their own and also false.

There has been a wealth of research examining the neural basis of false belief processing (Frith and Frith, 2003). However, as far as we are aware no study has examined how others' false beliefs about the contingencies between decisions and rewarding outcomes are processed. In addition, no study has investigated how false beliefs are processed when they pertain to the mental states of another with whom the subject is engaged in a 'real-time' social interaction. In this study, we used fMRI to investigate brain activity during a false belief task, where the subjects monitored the decisions of another (a confederate) in real-time. We were specifically interested in examining activity time-locked to events that signalled to a subject that another's belief about the outcome of their decision was false.

Recently it has been suggested that the Anterior Cingulate Cortex (ACC) may play an important role in monitoring the outcomes of both one's own and others' decisions (Apps et al., 2012; Behrens et al., 2009). This area has been implicated in processing information in a manner that conforms to the principles of Reinforcement Learning Theory (RL). In RL prediction error signals occur when new information reveals that the actual outcome of an action is discrepant from the predicted outcome. In accordance with the predictions of this theory, single-unit recording studies have found neurons in the ACC which are sensitive to both unexpectedly positive and unexpectedly negative outcomes (Amiez et al., 2005; Matsumoto et al., 2007; Seo and Lee, 2009). In addition, fMRI studies in humans have reported that the ACC is activated when one's own predictions about the outcome of a decision are erroneous (Ribas-Fernandes et al., 2011). However, no study has investigated whether such prediction error signals in the ACC occur when a subject is informed that another's prediction of an outcome is false.

False belief processing and RL processes share some similarities. In RL, it is hypothesised that error signals occur when the actual

* Corresponding author at: Department of Psychology, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK.
E-mail address: m.apps@rhul.ac.uk (M.A.J. Apps).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.09.010

outcome of a decision is discrepant from the predicted outcome (Schultz, 2006). In a false belief task the subject must identify that the predictions of the protagonist are discrepant from the actual outcomes they will receive. Thus, both false belief processing and RL require a subject to identify a discrepancy between a prediction and an outcome. Is ACC activity sensitive to both one's own erroneous predictions about the outcome of a decision and others' erroneous predictions about the outcomes of their decisions? A number of studies have shown that activity in different portions of the ACC is sensitive to social and non-social information. Specifically, there may be functional distinctions between the Sulcus (ACCs) and Gyrus (ACCg) (Beckmann et al., 2009; Behrens et al., 2009). Lesions specifically to the ACCs, which leave the ACCg intact, disrupt monkeys' abilities at using the outcomes of decisions to guide future choices (Rudebeck et al., 2008). In contrast, lesions to the ACCg disrupt monkeys' processing of social stimuli (Rudebeck et al., 2006). Behrens and colleagues (Behrens et al., 2008) tested the hypothesis that the ACCg and ACCs may perform similar computations on information from social and non-social sources respectively. Subjects performed a reward-based decision-making task with a confederate who gave them trial-by- trial advice and activity in the ACC was examined at the time that feedback about the outcome of a decision was given to a subject. They found that the ACCs processed the extent to which participants should change their behaviour in the future as a function of their own decisions. In contrast, the ACCg processed the extent to which they should change their behaviour as a function of the advice from the confederate. This suggests that the ACCg is engaged when new information is revealed about the outcomes of others' decisions, in the same manner that the ACCs is activated when new information is revealed about one's own decisions. However, the design of that study did not enable them to test whether the ACCg is engaged when a subject receives privileged information that tells them in real-time that another's belief about the outcome of a decision is false, i.e. signalling the discrepancy between another's prediction and the actual outcome of their decision.

In this study subjects performed a false belief task where the nature of the false belief related to another's predictions about the outcomes of their decisions. We used a design in which subjects monitored the predictions and outcomes of another's (a confederate) trials (see Fig. 1). Similar to the Sally–Anne task, the subjects received privileged information that was not received by the confederate. This took the form of cues that the subjects believed were relayed only to them, not the confederate, while they were inside the scanner. The cues informed the subject of the outcome of the trial before the confederate would have access to this information, enabling them to infer whether the predictions of the confederate were true or false. The subject's task was to decide whether this outcome was the same or different from the predicted outcome. On each trial one of two Predictions could be made (positive or negative) by the confederate and there could be one of two actual Outcomes (positive or negative). However, the actual Outcomes of trials were experimentally manipulated, such that on 1/3 of all trials the actual Outcome was different from the Prediction. Thus, at the time of the privileged cue the subjects were able to infer whether the predictions of the confederate were true or false.

In addition to the trials of the confederate, the participants also monitored trials performed by a computer. The computer performed trials in an identical manner to the confederate. The subjects were required to perform the same task on these trials as they were on the trials of the confederate This design (see Table 1) enabled us to test

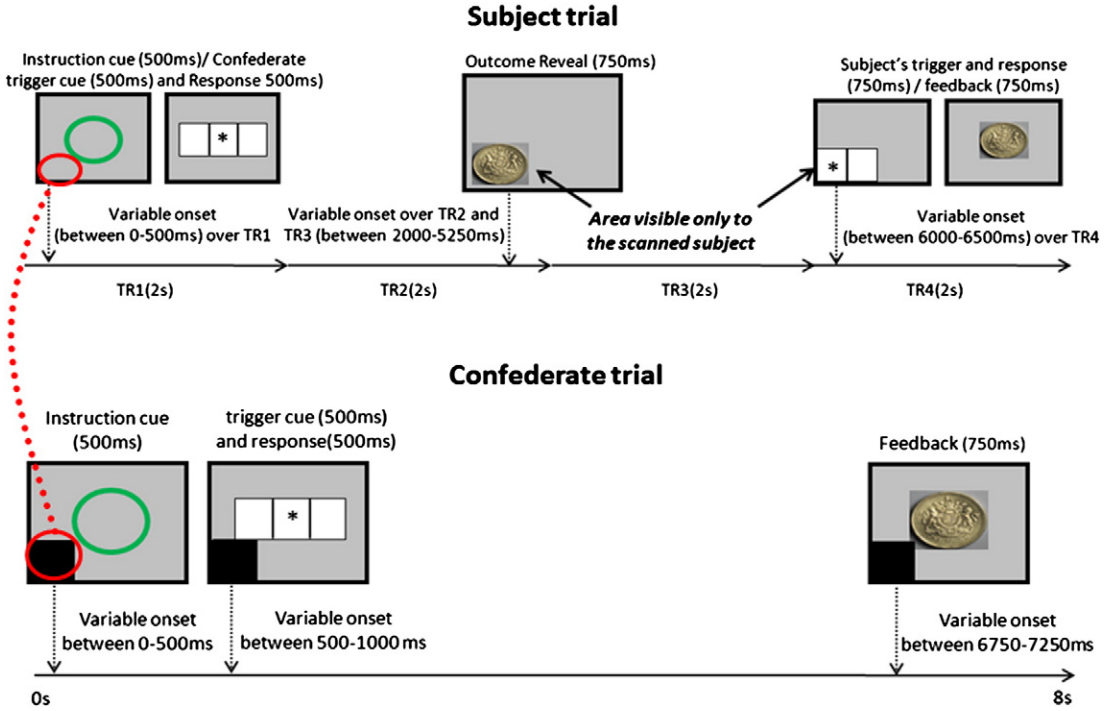

Fig. 1. Trial structure for the subject (top) and for the confederate (bottom). The subject saw all of the same trial events as the confederate, but also observed additional cues that the confederate could not see. For both the subject and the confederate each trial contained an instruction cue (a colour coded shape; green—the confederate would respond, black—a computer would respond), a trigger cue (instructing the confederate to respond) and the response (an asterisk indicating which of the three buttons was pressed); and finally feedback (either a one pound coin or a crossed out one pound coin). The confederate (seated in the control room next to the scanner) or computer responded to one of four shapes on each trial. The correct arbitrary associations had been learnt during a training session for three of these shapes ('learnt' cues). The correct response randomly changed on each trial for the fourth shape ('random' cue). The subject received privileged information in the corner of the screen that the confederate could not see (area in black in the corner indicates the area of the screen that could only be seen by the scanned subject).They received an additional outcome reveal cue (the actual outcome of the trial); the scanned subject trigger (a cue triggering a response from the subject) and they made a response (true or false, for the contingency between the prediction and outcome). We time-locked activity to the events that signalled the privileged information to the subject.

two hypotheses: Firstly, that the ACCs will be activated when new information reveals that Outcomes are unexpected on both the computer and the confederate trials and secondly, that the ACCg will be activated exclusively when new information reveals that an Outcome is unexpected on the confederate's trials.

Materials and methods

Subjects

Subjects were sixteen healthy right-handed participants (aged between 18 and 30; 9 female) screened for Neurological and Psychological disorders. One subject (1 female) failed to complete the whole scanning session and was excluded from the analyses. Subjects gave written informed consent and the study was approved by the Royal Holloway University of London Psychology Department Ethics Committee and the study conformed to regulations set out in the CUBIC MRI Rules of Operations (http://www.pc.rhul.ac.uk/sites/cubic/). Subjects were paired with one of two confederate participants, who they believed were a naïve participant. The subjects were not paid for their participation but were offered a picture of their brain as an incentive. The subjects were informed that the other participant performing the task with them (confederate) were being paid £5 for their participation as they were not being scanned.

Apparatus

Subjects lay supine in an MRI scanner with the fingers of the right hand positioned on a on MRI-compatible response box. Stimuli were projected onto a screen behind the subject and viewed in a mirror positioned above the subjects face. Presentation software (Neurobehavioral Systems, Inc., USA) was used for experimental control (stimulus presentation and response collection). A custom-built parallel port interface connected to the Presentation PC received transistor–transistor logic (TTL) pulse inputs from the response keypad. It also received TTL pulses from the MRI scanner at the onset of each volume acquisition, allowing events in the experiment to become precisely synchronised with the onset of each scan. The timings of all events in the experiment were sampled accurately, continuously and simultaneously (independently of Presentation) at a frequency of 1 kHz using an A/D 1401 unit (Cambridge Electronic Design, UK). Spike2 software was used to create a temporal record of these events. Reaction times were calculated off-line, and event timings were prepared for subsequent general linear model (GLM) analysis of fMRI data (see below).

Training

Subjects were pre-trained in pairs with a confederate one day prior to the scanning session. Training was conducted in two phases. In the first phase, the subject and the confederate were seated in front of the same monitor, each with their own keypad. They each performed a series of delayed-outcome conditional motor learning trials. During this phase, each trial consisted of an instruction cue (a coloured shape), a trigger cue (three white lines which indicated that a response should be made), the response cue (three white lines with an asterisk over one line indicating which response was made. "Missed" was presented at this point if a response was not made within the response window) and finally a feedback cue was presented (a one pound coin indicating a correct response or a one pound coin with a cross through it, indicating an incorrect response, "missed" if a response had not been made at the time of the trigger cue). They were both required to learn the arbitrary stimulus–response associations between three cues and three motor responses by trial and error. They also observed computer 'learning' associations (a non-biological control, as used in previous studies; Apps et al., 2012; Ramnani and Miall, 2004; Sanfey et al., 2003). Subjects had a 750 ms window in which to make a response following the onset of the trigger cue. The instruction cues were colour-coded, such that the subject responded to red shapes, the confederate responded to green shapes and the computer responded to black shapes. However, the form of the shapes was identical and all three agents learnt the same associations. As the confederates were paired with multiple different subjects throughout the piloting and experimental phases, they were highly over-trained on the associations. However, they were told to make deliberate errors (both responses that were too slow and also incorrect responses) to mimic the learning of a real participant. This first training phase ensured that the subject understood each of the stimulus–response associations and enabled them to observe and monitor the responses of the confederate. Once the subject, the confederate and the computer had made three contiguous, correct responses for each instruction cue, the task was completed.

In the second phase, the subject and the confederate practised the task that would be performed in the scanner on the following day (see task design below). During this phase, the subject was played the sound of the scanners EPI sequence through headphones, inside a mock scanner. The subject observed the confederate being seated in front of a monitor with a response keypad, before they entered the bore of the mock scanner. This practise session lasted 12 min and consisted of 90 trials.

Table 1
Experimental design. A 2×2×2 factorial design was used. The first factor manipulated the Agency of the respondent on the trial (confederate or computer,) the second factor manipulated the predicted outcome (positive or negative) and the third factor manipulated the actual outcome (positive or negative).

Outcome	Prediction			
	Confederate		Computer	
	Positive	Negative	Positive	Negative
Positive	*+ve true belief*	*+ve false belief*	*+ve true belief*	*+ve false belief*
Negative	*−ve false belief*	*−ve true belief*	*−ve false belief*	*−ve true belief*

Scanning session

Task design

During the second training phase and the scanning session the confederate and the computer continued to make responses on conditional motor learning trials in the same manner as they had during the earlier training session. However, the subject no longer performed trials in the same manner, instead they performed a false belief task on the conditional motor learning trials of the other agents (see below). The subjects were informed that they would see the responses of the computer and confederate in real-time from inside the scanner. However, the responses they observed were actually a series of computer controlled responses which occurred in a pre-programmed sequence. To ensure that the subject maintained the belief that the confederate was a naïve participant responding to the visual cues, four missed trials and three incorrect responses were programmed to occur on the confederate's trials. None were included on the computer's trials, to ensure that subject's maintained a sense that the confederate was a biological agent and the computer was not. The behaviour of the confederate (the number of errors and missed responses) was based on the responses of subjects in a pilot experiment. To ensure that all subjects maintained the belief in the deception, subjects were interviewed during a debriefing session. All participants that were included in the analysis maintained the belief that they were observing the decisions of the confederate in real-time and that only they were able to see the privileged information in the corner of the screen.

The trials of the confederate and the computer consisted of an instruction cue, a trigger cue and feedback, presented in the centre of the scanned subjects' visual display (see Fig. 1). All of these trial

elements were presented in real-time to the scanned subject inside the bore of the scanner. The same colour-coding of instruction cues was used for the computer and confederate as during training (green for confederate, black for computer). There were 360 trials in total, 180 of which consisted of one of the three instruction cues for which the correct association had been learned during the training session (90 confederate trials, 90 computer trials; 30 trials for each cue for each agent). The subject and confederate were also reminded of the stimulus–response associations they had learned before entering the scanner. However, feedback was pre-determined on these learnt cue trials, such that rewarding outcomes were only delivered on 2/3 of the trials (20 for each shape, for each Agent), even if a correct response was made. Thus, on 1/3 of the learnt cue trials (10 for each shape) a negative outcome was delivered.

In addition to the 'learnt' cues, an extra instruction cue was now presented on 50% of the trials (180 trials; 90 computer, 90 confederate trials). For this 'random' cue, the correct response changed randomly between each of the three buttons across trials. As such, it was not possible to learn the correct response for this instruction cue. Thus, the confederate would predict receiving a negative outcome on 2/3 of trials, given that there was a 1 in 3 chance of guessing the correct response on each trial. Unbeknown to the subjects, rewarding outcomes were fixed to be received on only 1/3 of all 'random' cue trials (30 computer trials, 30 confederate trials). The subjects were informed that these positive outcomes were delivered when the confederate or computer had made the correct response. Therefore, the confederate and computer would receive rewarding outcomes on 2/3 of the learned cue trials and negative outcomes on 2/3 of the random cue trials. Based on probabilities, a positive outcome would be predicted on the learnt cue trials and a negative outcome would be predicted on the random cue trials (see Table 1).

It is important to note that whilst the outcomes of the trials were experimentally manipulated, there was still an apparent contingency between the responses of the confederate and their resulting outcomes. On the random cue trials, subjects were told that a positive outcome would only be delivered if a correct response had been made, which would "probably only occur on 1/3 of the trials". On the learnt cue trials the subjects were informed that a positive outcome would only be delivered if a correct response was made. That is, if the confederate made an incorrect response, they would only ever receive a negative outcome. As such, the outcomes of the confederate's responses were still contingent on their choices. This contingency was apparent to both the subjects and the confederates as they were jointly instructed on the task prior to scanning. As such, subjects were aware that confederate would have the same expectations about all trials as themselves.

During the scanning session, the subject performed a false-belief task inside the bore of the scanner (see Fig. 1). In this session the subject received privileged information that was not given to the confederate. This took the form of a privileged cue that was presented after the trigger cue on each trial. To provide the subject with privileged information, we dedicated the corner of the screen to the information that was only to be displayed to them. The same information was displayed on the monitor used by the confederate and that used by the subject. However, we covered the corner of the confederate's screen, such that information displayed in that corner was only visible to the subject. Before the subject entered the scanner they were shown the monitor that would be used by the confederate and its covered area. Thus, the subjects were under the belief that they were receiving privileged information that the confederate could not see. On each trial, the subjects received an additional cue (the 'privileged cue') in the corner of the screen, which informed them what the actual outcome of the trial was. At this point in time, the subject knew both the confederate's prediction of the outcome and the actual outcome of the trial. Thus, when there was a discrepancy between the confederate's prediction and the actual outcome of the trial, the subject knew that the confederate was a holding a false belief. However, a potential confound was present in this design. On each trial the subject would have an expectation about the likely outcome of the trial. When the privileged information revealed that an outcome was unexpected it was therefore a violation of the subject's own expectations, as well as a violation of the predictions of the confederate. For this reason we introduced the computer conditions. As the computer trials were identical to the confederates, the subject would have matching expectations for the comparable conditions on the trials of each agent. Unexpected outcomes on the computer's trials would therefore signal a violation of the subject's own predictions, in the same manner as they would on the confederate's trials. Any area that showed a different profile of activity between the computer and confederate conditions could therefore not be a result of the subjects own predictions being erroneous, as they were matched on every aspect apart from the agency of the respondent.

Following the presentation of this privileged information, a trigger cue appeared in the corner of the screen. At this point the subject was required to indicate whether the confederate held a false belief. Subjects were instructed to "determine if the outcome is what would be predicted" on the computer trials. Subjects had 750 ms to indicate whether the belief was true or false by pressing the first button on the keypad for true and the second button on the keypad for false.

Trial structure

Trials for the subject (see Fig. 1) consisted of an instruction cue (colour-coded shapes for training partner or computer), immediately followed by a trigger cue (instructing a computer or training partner response), followed by a response cue displaying the response of the partner or computer. After a variable delay period a privileged cue was presented (informing the scanned subject what the actual outcome of the trial would be, in the corner of the screen). After a further variable delay, the scanned subject trigger appeared (a cue displayed in the corner of the screen, instructing the response from the scanned subject) followed by the scanned subject response (displaying the response of the scanned subject in the corner of the screen). Finally feedback was presented after another variable delay (displaying the outcome of the confederate or computer decision in the centre of the screen).

Conditions

To investigate activity that occurred at the specific moment in time when new information revealed that the confederate had a false belief, activity time-locked to the 'privileged' cues was examined. A $2\times2\times2$ Factorial design was used (see Table 1). The first factor was the Predicted outcome or 'Prediction' which could be positive or negative. The second factor was Outcome, which could be positive or negative and the third factor was Agency which could be either confederate or computer. This created eight different conditions that occurred time-locked to the cues which signalled the privileged information, which were as follows:

1. Confederate positive false belief (the confederate is predicting a negative outcome on the random cue trials, but the actual outcome is positive; 30 trials).
2. Confederate negative false belief (the confederate is predicting a positive outcome on the learnt cue trials, but the actual outcome is negative; 30 trials).
3. Confederate positive true belief (the confederate is predicting a positive outcome on the learned cue trials, and the actual outcome is positive; 60 trials).
4. Confederate negative true belief (the confederate is predicting a negative outcome on the random cue trials and the actual outcome is negative; 60 trials).

5. Computer positive false belief (a negative outcome is expected on the computer random cue trials, but the actual outcome is positive; 30 trials).
6. Computer negative false belief (a positive outcome is expected on the computer learnt cue trials, but the actual outcome will be negative; 30 trials).
7. Computer positive true belief (a positive outcome is expected on the learned cue trials and the actual outcome is positive; 60 trials).
8. Computer negative true belief (a negative outcome is expected on the random cue trials and the actual outcome is negative; 60 trials).

The aim of this investigation was to examine activity occurring when new information revealed another had a false belief and also activity that occurred whenever an outcome was unexpected. To examine these two occurrences, two main contrasts were conducted. The first looked for an interaction between Prediction (positive x negative) and Outcome (positive x negative), independent of the level of Agency. This would identify voxels that showed a difference in the BOLD response between true and false belief trials, irrespective of the Agent who made the initial response. The second looked for a three-way interactions between Agency (computer × confederate), Prediction (positive × negative) and Outcome (positive × negative). This would identify voxels that showed a differential response between prediction and outcome on the computer trials and the confederate trials.

It could be argued that an effect on the false belief trials could be attributed to the different frequency at which stimuli that signal these conditions to the subject are presented. In our design, there are different numbers of repetitions of the true belief conditions (60 trials) compared to the false belief (30 trials) conditions on the confederate trials. However, we control for any effects of differences in the number of trials across conditions, by comparing activity on the confederate trials with activity on the computer trials. Each computer condition was matched with one of the confederate conditions in terms of the predicted outcome, the actual outcome and the number of repetitions. As a result, we only report activity that shows a differential response between conditions that are presented the same number of times. We also note that the statistics employed in a general linear model based fMRI analysis are not biased by the number of repetitions of each condition. Indeed, it is not uncommon for neuroimaging studies to use different numbers of trials across conditions, particularly when uncommon events are essential for the aims of the experiment. Thus, we argue that none of the results reported in this study for the confederate false belief conditions can be attributed to the frequency of these conditions.

Finally, an additional analysis was performed examining activity time-locked to the instruction cues. At the time of the instruction cues, the subjects would be able to code the Predicted outcome of the trials. To examine activity time-locked to these instruction cues a 2 × 2 factorial design was used. The first factor was the Prediction (positive or negative) and the second factor was the Agency (confederate or computer). No significant results were identified in this analysis at the whole-brain level, or with small volume corrections in the ACC. This absence of a response is not surprising, given that these events were only jittered over a short time-frame. This was a necessity in order to maximise the sampling of activity time-locked to the privileged cue event.

Experimental timing

An important feature of the study was that activity was time-locked specifically to the point in time when privileged information was revealed to the subject. In order to do this, a variable delay was introduced between the instruction cue and the privileged cue. An additional delay was also introduced between the privileged cue and the scanned subject trigger cue. This allowed BOLD activity time-locked to the privileged cue to be isolated, without contaminating effects of either prior or subsequent trial events. Events in each trial took place across four TRs (0–8 s; TR = 2 s). The interval between scan onset and instruction cue onset was varied over the first TR from trial-to-trial. To optimally sample the cue of interest, the privileged cue, a randomly varying interval between the scan onset and these cues was introduced over the second and third TRs. This achieved an effective temporal sampling resolution much finer than one TR for the conditions of interest. These intervals were uniformly distributed for each condition, ensuring that Evoked Haemodynamic responses (EHRs) time-locked to the privileged cues were sampled evenly across the time period following each outcome reveal. The scanned subject trigger and feedback cues were randomly jittered over the fourth TR.

Functional imaging and analysis

Data acquisition

1470 EPI scans were acquired from each participant using a 3 T Siemens Trio scanner (Royal Holloway University of London). 27 slices were acquired in an interleaved manner, at an oblique angle (≈30°) to the AC-PC line to decrease the impact of susceptibility artefact in the ACC (Deichmann et al., 2003). A voxel size of 3 × 3 × 4 mm (25% slice gap, 0.8 mm) was used; TR = 2 s, TE = 32, flip angle = 80°. The functional sequence lasted 49 min. High resolution T1-weighted structural images were also acquired at a resolution of 1 × 1 × 1 mm using an MPRAGE sequence.

Image preprocessing

Scans were pre-processed using SPM5 (www.fil.ion.ucl.ac.uk/spm) by spatial realignment to the first scan, normalisation to the ICBM EPI template using both linear affine transformations and non-linear transformations (Friston et al., 1995a). Lastly, a Gaussian kernel of 8 mm was applied to spatially smooth the images in order to conform o the Gaussian assumptions of the GLM implemented in SPM5.

Statistical analysis

Event definition and modelling. Nine separate event types were modelled in the analysis. Each of the eight kinds of 'outcome reveal' cues were modelled as a separate event type. The instruction cues, trigger cues, feedback cues, and the outcome reveal cues from trials which were either missed or incorrect responses, were modelled as one regressor. Trials were classed as missed if the response was too early (before the scanned subject's trigger cue) or too late (a reaction time > 1000 ms). Each event type was used to construct a series of regressors by convolving the event timings with a Fourier set of five harmonic functions (two sine, two cosine, one envelope function with a Hanning window of 32 s). This strategy was employed, as it has been in previous research (Apps et al., 2012; Balsters and Ramnani, 2008; Lau et al., 2006), because it allowed us to investigate potentially complex haemodynamic activity without making stringent prior assumptions about its timecourse. The residual effects of head motion were modelled in the analysis by including the six parameters of head motion acquired from the realignment stage of the preprocessing as covariates of no interest. Prior to the study, a set of planned experimental timings were carefully checked so that they resulted in an estimable GLM in which the statistical independence of the eight event types was preserved.

First-level analysis. The GLMs were estimated in SPM5 (Friston et al., 1995b) on a Dual Core AMD Pentium 32 MHz PC with 2 GB of RAM, running Ubuntu and Matlab 6.5 (MathWorks Inc). SPM{t} contrast images were computed at the first-level, one image per basis

function. Thus, 45 SPM{t} images were created at the first-level to be used in the second level.

Random effects group analysis. A random effects analysis (Full-Factorial ANOVA) was applied to determine voxels significantly different at the group level. SPM{t} images from all subjects at the first-level were grouped into two factors, basis function and condition. We conducted F-contrasts across the Fourier basis functions to look for significant interaction effects (see 'conditions' above). To apply correction for multiple comparisons, we used 80% probability anatomical masks of the ACCg and ACCs. To create each mask, subject-specific masks of the ACCg and ACCs were constructed in FSL (http://www.fmrib.ox.ac.uk/fsl/) . Although the cytoarchitectonic boundaries of the ACC have no corresponding gross anatomical landmarks, we defined the anatomical boundaries based on the location of these boundaries in previous literature investigating Cingulate cytoarchitecture(Vogt et al., 1995). We used a posterior vertical extent to each mask extending 22 mm posterior to the Anterior Commisure, i.e. the posterior border of the midcingulate cortex. We included all voxels that lay within the ACCs or the ACCg extending anterior to this border, including subgenual Cingulate cortex. The final ACCs and ACCg masks included only voxels which were within each region in 80% of our subjects.

Results

Behavioural results

Subjects performed a false belief task on the contingency between predicted and actual outcomes of responses made by a computer or a confederate. Subjects were required to indicate whether the Predicted outcome on a trial was the same as the actual Outcome. Subjects performed the task at a high level of accuracy (mean of 92.9% of 355 trials performed correctly; mean 25.2 trials incorrect or missed SD $\pm$ 13.54). Thus, they were able to correctly understand both the predicted and actual outcomes of trials. To examine the subjects' performance of the task in both the confederate and computer conditions the number of correct trials in each condition was converted into an overall percentage for the confederate and computer trials. To test for any significant difference in task performance between the confederate and computer trials, a repeated measures *t*-test was conducted. No significant difference in task accuracy was found between the computer and confederate conditions ($t(14) = 0.174$, $p = 0.865$). Therefore, subjects were able to perform the task at a high level of accuracy regardless of whether it was a computer or confederate trial.

fMRI results

The main aim of this experiment was to investigate activity at the point in time when new information revealed that another's belief was false. A $2\times2\times2$ factorial design was used to examine activity time-locked to the cue which signalled privileged information only to the scanned subject. The first factor was the predicted outcome of the trial (Prediction), which could be positive or negative; the second factor was the actual outcome of the trial (Outcome), which could be positive or negative and the third factor was the respondent on the trial (Agent) which could be either the confederate or the computer.

The experimental design contained four conditions in which the actual outcome was discrepant from the predicted outcome. There were two conditions in which the confederate's prediction of the outcome would be false and two conditions in which an unexpected outcome occurred on the computer's trials (see Table 1). Two hypotheses were tested: (i) the ACCs responds on trials where the actual Outcome is different from the Prediction, regardless of whether it was a computer or confederate trial; (ii) the ACCg responds only when the confederate's Prediction is false. To test the anatomical specificity of these hypotheses, masks of the ACCg and ACCs were used as a small volume correction for multiple comparisons. These masks ensured that any activated voxel at the group level would be within the ACCg or ACCs of 80% of the subjects. The results of whole-brain analyses are included in supplementary material.

Prediction error

To test the first hypothesis, we looked for a two-way interaction between Prediction and Outcome, independently of whether the trial was that of the computer or the confederate (see supplementary Table 2 for a full list of whole-brain uncorrected results). This contrast identified voxels which showed a significant effect of any unexpected Outcome, regardless of the Agent performing the trial. A significant effect (MNI coordinates: 4,14,32, $Z = 3.95$, $p<0.005$svc) was found in the ACCs (putatively within the Rostral Cingulate Zone (RCZ) in midcingulate area 24c'). Examination of peristimulus time histograms (PSTH), of data from the peak voxel, revealed that this effect was being driven by a significant response to both the confederate and computer positive false belief trials (see Fig. 2). To examine whether there were any differences in the response between the confederate and computer trials in the ACCs we performed a three-way interaction (Agency $\times$ Prediction $\times$ Outcome). No voxels showed a three-way interaction effect in the ACCs, even at a lower threshold ($p<0.01$ uncorrected).

To test whether the effect identified in the two-way interaction was being driven by the response to these two positive false belief conditions, additional contrasts were conducted (see Table 2). These contrasts revealed significant differences between the computer and confederate positive false belief conditions, where the outcome was unexpectedly positive, and all other conditions ($p<0.005$ uncorrected). A cluster of 31 voxels showed this effect. Thus, the interaction was driven by responses to both the confederate and computer positive false belief conditions i.e., the effect of positive false beliefs. Importantly, there was no significant difference between these two conditions, even at a much lower threshold ($p>0.01$ uncorrected). This indicates that the ACCs made a response on all trials where there was an unexpectedly positive outcome.

Confederate false belief

To test the second hypothesis, we looked for an interaction between Prediction, Outcome and Agency in the ACCg (see supplementary Table 1 for a list whole-brain uncorrected results). A significant effect was found in the ACCg (0, 8, 28; $Z = 3.41$; $p<0.05$ svc; see Fig. 2) putatively in midcingulate area 24a'/24b' (Vogt et al., 1995). This cluster also showed an interaction between Prediction and Outcome on the confederate trials. To confirm that no portion of the ACCg was responsive to the interaction between Prediction and Outcome on the computer trials, we performed a two-way interaction between these factors on the computer trials and also a two-way interaction between these factors regardless of Agency. No voxels within the ACCg showed a significant effect of either of these two interactions ($p>0.01$ uncorrected). Examination of the PSTH revealed that the identified effect was driven by a response only to the confederate positive false belief condition. To test statistically whether this interaction effect was being driven by the simple effect of confederate positive false beliefs, contrasts were run between this condition and every other condition. A cluster of 17 voxels that showed the three way interaction, also showed a significant effect ($p<0.005$ uncorrected) in each of these seven contrasts.

In summary, these findings show that parts of the ACC have the common property of sensitivity to unexpectedly positive outcomes of decisions. However, responses of the ACCs and ACCg differ in

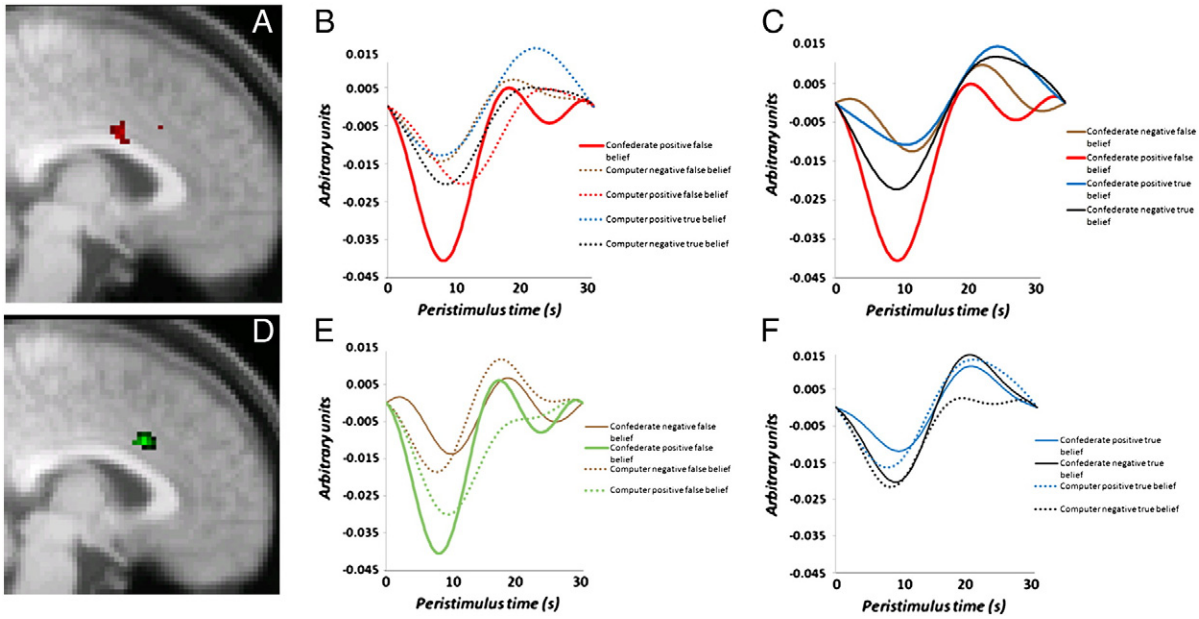

Fig. 2. fMRI results. *False belief:* three-way interaction between Prediction, Outcome and Agency (A) Activity shown in the ACCg on a saggital plane (B) Plots of the Fourier Basis sets in peri-stimulus time, weighted by their parameter estimates, for the four confederate conditions in the peak voxel and (C) Plots of the four computer conditions and the confederate positive false belief condition. Prediction Error: Two-way interaction between Prediction and Outcome on both the confederate and computer trials. (D) Activity shown in the ACCs on a saggital plane. (E) Plots of the Fourier Basis sets in peri-stimulus time, for the four false belief conditions in the peak voxel and (F) plots of the four true belief conditions.

terms of whose errors are being processed. Whereas the ACCs responds to any unexpectedly positive outcome, the ACCg responds exclusively when the outcomes of another's decision is unexpectedly positive.

Discussion

This study tested two hypotheses about the processing of prediction errors and false beliefs in the ACC. Firstly, that the ACCg would be engaged only at the point in time that privileged information revealed that another's prediction of an outcome was false. Secondly, that the ACCs would be activated at the point in time that new information revealed that an outcome of a decision was unexpected, regardless of the agent who made the response. Our results show that both the ACCg and the ACCs respond to unexpectedly positive outcomes of decisions. In line with our hypotheses, the ACCs signal a prediction error for any decision and the ACCg signals others' erroneous predictions. These findings therefore support the claim that the ACCs and the ACCg compute similar information about the outcomes of decisions, but, the ACCg processes this information about the outcomes of other's decisions.

Typically, studies that examine the processing of false beliefs report activation in three interconnected areas, namely the paracingulate cortex, the posterior Superior Temporal Sulcus (STS) and the Temporal Poles (TPs) (Aichhorn et al., 2009; Fletcher et al., 1995; Frith and Frith, 2003; Gallagher et al., 2000; Grezes et al., 2004; Perner et al., 2006; Saxe and Kanwisher, 2003; Sommer et al., 2007). To date, only one study has shown activation in the ACC when processing others' false beliefs (Sommer et al., 2007). The discrepancy between the findings of our study and those of previous studies can be attributed to two features of our design. Firstly, none of these studies have investigated brain activity that occurs at the point in time when new information reveals that another's belief is false. Instead they have investigated activity aggregated across entire trials, blocks of trials, or examined activity time-locked to retrospective false belief judgements. One previous study has reported activity at the time that privileged information was presented to subjects, that of Sommer et al. (2007). However, in their study the absence of temporal jitter between the privileged information and the other events in the trials means that activity evoked by the cue of interest may have been confounded by other events in the trials. In our study, activity was time-locked to the particular events that allowed a subject to infer the true or false nature of the mental state of another. These events are the first opportunity in our trials that subjects have to judge that another's beliefs are false. They were sampled independently of the other trial events. Thus, unlike previous studies, activity was examined at the exact moment in time when new information reveals that another's belief is false, making our study sensitive for identifying activity in the ACC time-locked to this event. Secondly, in most false belief tasks the subject does not reason about the mental states of another individual whilst interacting with them in real-time (Frith and Frith, 2003). Instead, the subject is reasoning about the hypothetical mental state of a protagonist. In this study, the subject was making inferences about specific predictions made by another person whilst monitoring their behaviour in real-time. Thus, when privileged information revealed that the outcome of another's decision was to be unexpected, it indicated that their current mental state, their prediction, was false. Although many studies have investigated the processing of decisions during real-time social interactions (Rilling et al., 2008) none have investigated the contribution of the ACCg to processing other's false beliefs about the outcomes of reward-related decisions. These important differences have allowed us to highlight how an area,

Table 2
Additional contrasts conducted on the cluster identified in the ACCs.

Contrast	*Z*	*P*
(Confederate and computer positive true belief) <> (confederate and computer positive false belief)	3.48	$P<0.005$
(Confederate and computer negative true belief) <> (confederate and computer positive false belief)	3.48	$P<0.005$
(Confederate and computer negative false belief) <> (confederate and computer positive false belief)	3.22	$P<0.005$
Confederate positive false belief <> computer positive false belief	1.73	$P>0.01$

outside of the classical core-circuit engaged when processing others' mental states, is also engaged when processing information about others' mental states. We therefore argue that the ACCg may constitute part of this network that is specialised for processing social information.

There is a considerable body of evidence which shows that the ACC is involved in processing the outcomes of decisions and particularly in signalling when they are different from expectations (Behrens et al., 2007; Frank et al., 2005; Kennerley et al., 2011; Ribas-Fernandes et al., 2011; Walton and Mars, 2007). Single-unit recording studies have shown that neurons in the ACCs increase their firing rate as the probability of reward increases (Hayden and Platt, 2010; Kennerley and Wallis, 2009; Kennerley et al., 2009) and also do so in line with the magnitude of prediction error responses (Amiez et al., 2005; Matsumoto et al., 2007; Sallet et al., 2007). This would suggest that ACC neurons code reward predictions and errors in the manner proposed by RL theory. That is, there are neurons which code predictions about the outcomes of choices during choice selection and also neurons in the same region which signal when the outcome of the decision reveals that predictions were erroneous.

One might argue that finding prediction error signals in a small sample of neurons in these studies is not informative of the overarching functional properties of the ACC. However, there is additional, converging evidence from studies investigating the behavioural effects of lesions in non-human primates, as well as electrophysiological and neuroimaging studies in humans. These studies suggest that the error signals in the ACC may be involved in learning and guiding behaviour. For instance, lesions to the ACCs in monkeys result in an inability to adapt behaviour appropriately in a task requiring reinforcement learning processes (Kennerley et al., 2006). In humans, the dipole of the error-related negativity (ERN), an event-related potential detected using electroencephalography (EEG), has been localised to the ACC. This ERN occurs whenever errors are detected including when the outcome of a decision is unexpectedly positive (Frank et al., 2005; Holroyd and Coles, 2008). fMRI studies in humans also show analogous BOLD response that occurs when an outcome of a decision reveals that a prediction was erroneous (Carter et al., 1998; Jessup et al., 2010; Ribas-Fernandes et al., 2011; Vickery et al., 2011). As such, the signal identified in the ACCs in this study is consistent with many of the findings in the literature. However, in this study activity in the ACCs occurred whenever there was an unexpectedly positive outcome on a computer or third-persons trials. This is distinct from the findings of previous studies, which have shown such signals to occur when the unexpected outcome is that of the subjects' own decisions. This therefore suggests that the ACCs may code for any discrepancies between a subject's predictions and actual outcomes of decisions, regardless of the source of the decision. The ACCg also responded when prediction were erroneous. However, the ACCg was only sensitive to the discrepancy between another's prediction and the actual outcome of their decision. Thus, this study supports the notion that neurons the ACCs and the ACCg perform similar computations on the contingencies between decisions and outcomes. However, the ACCg processes the contingency between another's prediction about the outcome of their decision and its actual outcome.

An interesting aspect of the results was that no significant effects were identified within the ACC for either the confederate or the computer negative false belief trials i.e., no part of the ACC responded to unexpectedly negative outcomes. Whilst this is somewhat surprising, as it is well established that there are neurons in the ACCs which signal both positive and negative prediction errors (Matsumoto et al., 2007; Sallet et al., 2007), a recent neurophysiology study identified that such a finding would be expected, if you record the population response from the ACCs (Kennerley et al., 2011). Kennerley et al. (2011) found neurons that responded to positive, negative, and both types of prediction error in the ACCs at the time of the outcomes. In addition, they found neurons that signalled the value of outcomes regardless of the prediction in the same region. However, the overarching response at the population level occurred when an outcome was unexpectedly positive. Thus, our finding that the ACCs responds differentially to unexpectedly positive outcomes compared with the response to any other expected or unexpected outcomes, is supported by research investigating the neurophysiological properties of the ACC. In addition, we found that the ACCg also responds to unexpectedly positive outcomes when they occur following the decision of a third-person. This raises the possibility that the neurophysiological properties of the ACCg may be similar to those of the ACCs, with the ACCg responding to outcomes when the decision has been made by another. Future research should therefore focus on investigating whether the ACCg and the ACCs share similar neurophysiological properties when observing another's or making one's own decisions.

The view that the ACCs and the ACCg perform similar reward-related computations is supported by anatomical evidence. Both areas receive direct projections from dopamine neurons within the Ventral Tegmental Area (VTA; (Williams and Goldman-Rakic, 1998)) and also from the striatum (Yeterian and Pandya, 1991). Importantly, the VTA and the ventral striatum have been shown to signal reward prediction errors (D'Ardenne et al., 2008; McClure et al., 2003; O'Doherty et al., 2003; Seymour et al., 2007). Such a 'connectional fingerprint' implicates the ACC in processing information related to the prediction error signals in the ventral striatum and VTA. In addition, both the ACCs and ACCg share connections to areas implicated in decision-making and reward processing, including portions of the intraparietal sulcus (Vogt and Pandya, 1987), the hippocampal formation (Vogt and Pandya, 1987) and parts of the orbitofrontal cortex (Carmichael and Price, 1995; Pandya et al., 1981). However, despite many common connections the ACCg and the ACCs also exhibit several that are distinct. These include connections between the ACCg, medial parts of the superior frontal gyrus (bordering areas 8, 9 and 32'), portions of the pSTS and the TPs (i.e. the core-circuit engaged in previous false belief studies). Connections to these areas are not found in the ACCs (Pandya et al., 1981; Petrides and Pandya, 2006; Seltzer and Pandya, 1989; Vogt and Pandya, 1987). This would suggest that both the ACCg and the ACCs have access to reward-related information, but the ACCg has additional access to information in the core-circuit which processes others' mental states. Thus, anatomical evidence supports the notion that the ACCs and the ACCg signal when there is discrepancy between a prediction about a rewarding outcome and an actual outcome. However, they process this information for one's own predictions or the predictions of another, respectively.

Recently it has been suggested that the computations that are hypothesised in RL theory may accurately account for many different forms of information processing, beyond the learning of decision-outcome contingencies. Furthermore, some have suggested that RL processes may reflect a general mechanism of how the brain processes decisions (Rushworth et al., 2009). That is, prediction errors are a common mechanism for how new information is used to update prior expectations, with many brain areas processing discrepancies between predicted and actual outcomes (Vickery et al., 2011). A few recent studies have shown that signals predicted by RL also occur when subjects are making and monitoring decisions during social interactions. These studies have shown that the ACCg and the striatum are engaged when there are discrepancies between the predicted and actual outcomes of another's decisions (Apps et al., 2012; Behrens et al., 2008; Burke et al., 2010). In each of these studies, therefore, signals predicted by RL theory are evoked when monitoring the decision-making of others and not when processing the outcomes of one's own decisions. Here, we show that the function of a portion of the ACCg may be to compute the discrepancy between another's prediction and the actual outcome of their choice. Thus, prediction error processing in the ACCg may play an important role

in guiding behaviour during social interactions, by signalling the erroneous nature of others' predictions.

Funding

ESRC 1 + 3 PhD studentship awarded to MAJA and NR.

Acknowledgments

The authors would like to thank Ari Lingeswaran and Vivien Ainley for their help with the data collection.

Appendix A. Supplementary

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.09.010.

References

Aichhorn, M., Perner, J., Weiss, B., Kronbichler, M., Staffen, W., Ladurner, G., 2009. Temporo-parietal junction activity in theory-of-mind tasks: falseness, beliefs, or attention. J. Cogn. Neurosci. 21, 1179–1192.
Amiez, C., Joseph, J.P., Procyk, E., 2005. Anterior cingulate error-related activity is modulated by predicted reward. Eur. J. Neurosci. 21, 3447–3452.
Apps, M.A.J., Balsters, J.H., Ramnani, N., 2012. The anterior cingulate cortex: monitoring the outcomes of others' decisions. Soc. Neurosci. 7, 424–435.
Balsters, J.H., Ramnani, N., 2008. Symbolic representations of action in the human cerebellum. NeuroImage 43, 388–398.
Beckmann, M., Johansen-Berg, H., Rushworth, M.F.S., 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. J. Neurosci. 29, 1175–1190.
Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221.
Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., Rushworth, M.F.S., 2008. Associative learning of social value. Nature 456, 245–249.
Behrens, T.E.J., Hunt, L.T., Rushworth, M.F.S., 2009. The computation of social behavior. Science 324, 1160–1164.
Burke, C.J., Tobler, P.N., Baddeley, M., Schultz, W., 2010. Neural mechanisms of observational learning. Proc. Natl. Acad. Sci. U. S. A. 107, 14431–14436.
Carmichael, S.T., Price, J.L., 1995. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J. Comp. Neurol. 363, 615–641.
Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280, 747–749.
D'Ardenne, K., McClure, S.M., Nystrom, L.E., Cohen, J.D., 2008. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319, 1264–1267.
Deichmann, R., Gottfried, J.A., Hutton, C., Turner, R., 2003. Optimized EPI for fMRI studies of the orbitofrontal cortex. NeuroImage 19, 430–441.
Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S.J., Frith, C.D., 1995. Other minds in the brain—a functional imaging study of theory of mind in story comprehension. Cognition 57, 109–128.
Frank, M.J., Woroch, B.S., Curran, T., 2005. Error-related negativity predicts reinforcement learning and conflict biases. Neuron 47, 495–501.
Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S.J., 1995a. Spatial registration and normalization of images. Hum. Brain Mapp. 3, 165–189.
Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C.R., Frackowiak, R.S.J., Turner, R., 1995b. Analysis of fMRI time-series revisited. NeuroImage 2, 45–53.
Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 459–473.
Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 38, 11–21.
Grezes, J., Frith, C.D., Passingham, R.E., 2004. Inferring false beliefs from the actions of oneself and others: an fMRI study. NeuroImage 21, 744–750.
Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. J. Neurosci. 30, 3339–3346.
Holroyd, C.B., Coles, M.G.H., 2008. Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behaviour. Cortex 44, 548–559.
Hughes, C., Adlam, A., Happe, F., Jackson, J., Taylor, A., Caspi, A., 2000. Good test–retest reliability for standard and advanced false-belief tasks across a wide range of abilities. J. Child Psychol. Psychiatry 41, 483–490.
Jessup, R.K., Busemeyer, J.R., Brown, J.W., 2010. Error effects in anterior cingulate cortex reverse when error likelihood is high. J. Neurosci. 30, 3467–3472.
Kennerley, S.W., Wallis, J.D., 2009. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur. J. Neurosci. 29, 2061–2073.
Kennerley, S.W., Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rushworth, M.F.S., 2006. Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. 9, 940–947.
Kennerley, S.W., Dahmubed, A.F., Lara, A.H., Wallis, J.D., 2009. Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci. 21, 1162–1178.
Kennerley, S.W., Behrens, T.E.J., Wallis, J.D., 2011. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat. Neurosci. 14, 1581–1589.
Lau, H.C., Rogers, R.D., Passingham, R.E., 2006. On measuring the perceived onsets of spontaneous actions. J. Neurosci. 26, 7265–7271.
Matsumoto, M., Matsumoto, K., Abe, H., Tanaka, K., 2007. Medial prefrontal cell activity signaling prediction errors of action values. Nat. Neurosci. 10, 647–656.
McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.
O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. Neuron 38, 329–337.
Pandya, D.N., Vanhoesen, G.W., Mesulam, M.M., 1981. Efferent connections of the cingulate gyrus in the rhesus-monkey. Exp. Brain Res. 42, 319–330.
Perner, J., Lang, B., 1999. Development of theory of mind and executive control. Trends Cogn. Sci. 3, 337–344.
Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Thinking of mental and other representations: the roles of left and right temporo-parietal junction. Soc. Neurosci. 1, 245–258.
Petrides, M., Pandya, D.N., 2006. Efferent association pathways originating in the caudal prefrontal cortex in the macaque monkey. J. Comp. Neurol. 498, 227–251.
Ramnani, N., Miall, R.C., 2004. A system in the human brain for predicting the actions of others. Nat. Neurosci. 7, 85–90.
Ribas-Fernandes, J.J.F., Solway, A., Diuk, C., McGuire, J.T., Barto, A.G., Niv, Y., Botvinick, M.M., 2011. A neural signature of hierarchical reinforcement learning. Neuron 71, 370–379.
Rilling, J.K., King-Casas, B., Sanfey, A.G., 2008. The neurobiology of social decision-making. Curr. Opin. Neurobiol. 18, 159–165.
Rudebeck, P.H., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2006. A role for the macaque anterior cingulate gyrus in social valuation. Science 313, 1310–1312.
Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci. 28, 13775–13785.
Rushworth, M.F.S., Mars, R.B., Summerfield, C., 2009. General mechanisms for making decisions? Curr. Opin. Neurobiol. 19, 75–83.
Sallet, J., Quilodran, R., Rothe, M., Vezoli, J., Joseph, J.P., Procyk, E., 2007. Expectations, gains, and losses in the anterior cingulate cortex. Cogn. Affect. Behav. Neurosci. 7, 327–336.
Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. Science 300, 1755–1758.
Saxe, R., Kanwisher, N., 2003. People thinking about thinking people—the role of the temporo-parietal junction in "theory of mind". NeuroImage 19, 1835–1842.
Schultz, W., 2006. Behavioral theories and the neurophysiology of reward. Annu. Rev. Psychol. 57, 87–115.
Seltzer, B., Pandya, D.N., 1989. Frontal-lobe connections of the superior temporal sulcus in the rhesus-monkey. J. Comp. Neurol. 281, 97–113.
Seo, H., Lee, D., 2009. Behavioral and neural changes after gains and losses of conditioned reinforcers. J. Neurosci. 29, 3627–3641.
Seymour, B., Daw, N., Dayan, P., Singer, T., Dolan, R., 2007. Differential encoding of losses and gains in the human striatum. J. Neurosci. 27, 4826–4831.
Sommer, M., Dohnel, K., Sodian, B., Meinhardt, J., Thoermer, C., Hajak, G., 2007. Neural correlates of true and false belief reasoning. NeuroImage 35, 1378–1384.
Vickery, T.J., Chun, M.M., Lee, D., 2011. Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron 72, 166–177.
Vogt, b.a, pandya, d.n, 1987. Cingulate cortex of the rhesus-monkey.2. Cortical afferents. J. Comp. Neurol. 262, 271–289.
Vogt, B.A., Nimchinsky, E.A., Vogt, L.J., Hof, P.R., 1995. Human cingulate cortex—surface-features, flat maps, and cytoarchitecture. J. Comp. Neurol. 359, 490–506.
Walton, M.E., Mars, R.B., 2007. Probing human and monkey anterior cingulate cortex in variable environments. Cogn. Affect. Behav. Neurosci. 7, 413–422.
Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development: the truth about false belief. Child Dev. 72, 655–684.
Williams, S.M., Goldman-Rakic, P.S., 1998. Widespread origin of the primate mesofrontal dopamine system. Cereb. Cortex 8, 321–345.
Wimmer, H., Perner, J., 1983. Beliefs about beliefs—representation and constraining function of wrong beliefs in young childrens understanding of deception. Cognition 13, 103–128.
Yeterian, E.H., Pandya, D.N., 1991. Prefrontostriatal connections in relation to cortical architectonic organization in rhesus-monkeys. J. Comp. Neurol. 312, 43–67.