

When your error becomes my error: anterior insula activation in response to observed errors is modulated by agency

Emiel Cracco,^{1,2,*} Charlotte Desmet,^{1,2} and Marcel Brass^{1,2}

¹Department of Experimental Psychology, Ghent University, B9000 Ghent, Belgium and ²Ghent Institute for Functional and Metabolic Imaging of the brain, Ghent University, B9000 Ghent, Belgium

*Correspondence should be addressed to Emiel Cracco, Department of Experimental Psychology, Ghent University, B9000 Ghent, Belgium. E-mail: emiel.cracco@ugent.be

Abstract

Research on error observation has focused predominantly on situations in which individuals are passive observers of errors. In daily life, however, we are often jointly responsible for the mistakes of others. In the current study, we examined how information on agency is integrated in the error observation network. It was found that activation in the anterior insula but not in the posterior medial frontal cortex or lateral prefrontal cortex differentiates between observed errors for which we are partly responsible or not. Interestingly, the activation pattern of the AI was mirrored by feelings of guilt and shame. These results suggest that the anterior insula is crucially involved in evaluating the consequences of our actions for other persons. Consequently, this region may be thought of as critical in guiding social behavior.

Key words: error observation, agency, fMRI

Introduction

Errors often have negative consequences for the agent of the error. For example, a rider in a cycle race is likely to crash if he makes a steering mistake during a descent. In order to avoid such negative outcomes, it is important that we are able to efficiently monitor our actions for errors. A long tradition of neurophysiological research has identified that this function is supported primarily by a network including the posterior medial frontal cortex (pmFC), the lateral prefrontal cortex (LPFC) and the anterior insula (AI) (for reviews, see Ridderinkhof *et al.*, 2004; Taylor *et al.*, 2007; Ullsperger *et al.*, 2014). More precisely, it is assumed that activation in the pmFC reflects the detection and prediction of errors (Falkenstein *et al.*, 1990; Botvinick *et al.*, 2001; Holroyd and Coles, 2002; Brown and Braver, 2005; Rushworth and Behrens, 2008), whereas the LPFC is responsible for subsequent adaptations in cognitive control (Kerns *et al.*, 2004; Marco-Pallarés *et al.*, 2008), and the AI is involved in the conscious awareness (Ullsperger *et al.*, 2010) and affective evaluation (Koban *et al.*, 2010, 2013; Koban and Pourtois, 2014) of errors.

Interestingly, the same network is also found to be activated when another person is observed making an error (Shane *et al.*, 2008; De Bruijn *et al.*, 2009; Koban *et al.*, 2013; Yu *et al.*, 2014), suggesting that monitoring self-generated and observed behavior relies on the same neural mechanisms (e.g. Miltner *et al.*, 2004; van Schie *et al.*, 2004; Bates *et al.*, 2005; De Bruijn *et al.*, 2009).

However, research on error observation has primarily focused on situations in which participants are passive observers of errors (Shane *et al.*, 2008; Newman-Norlund *et al.*, 2009; Howard-Jones *et al.*, 2010; Monfardini *et al.*, 2013; Desmet *et al.*, 2014). In daily life, on the other hand, errors are often embedded in an interactive context. As a result, we may sometimes share responsibility for the errors other people make. Imagine, for instance, that the abovementioned rider made a steering mistake because he was pressured by the team leader to speed up. In this situation the team leader is partly responsible for the error of the rider. Preventing the situation from reoccurring therefore requires that the team leader acknowledges his share

Received: 22 May 2015; Revised: 7 September 2015; Accepted: 17 September 2015

© The Author (2015). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

in the rider's mistake. In more general terms, the above example illustrates that prosocial and moral behavior crucially rests on the ability to correctly assign agency in interactive situations (e.g. Koban et al., 2013; Yu et al., 2014; Cui et al., 2015; Lepron et al., 2015). In spite of its important social function, however, it is still unknown how the evaluation of agency is integrated in the error observation network. Therefore, the aim of the current study is to test how activation in this network varies as a function of agency.

Considering the functions that have been ascribed to the different regions of the error network, a likely candidate to evaluate the degree of agency over an observed error is the AI. As outlined above, this region is consistently found in response to both self-generated (e.g. Ullsperger and von Cramon, 2001, 2004; Marco-Pallarés et al., 2008) and observed (e.g. De Bruijn et al., 2009; Koban et al., 2013; Yu et al., 2014) errors. In line with its role in interoceptive awareness (Critchley et al., 2004; Craig, 2009; Ullsperger et al., 2010), it has been suggested that the role of the AI is to signal the affective consequences of intentional actions (Brass and Haggard, 2010). Interestingly, evidence linking the AI to the simulation of vicarious emotions (Wicker et al., 2003) and vicarious pain (Singer et al., 2004; Lamm et al., 2011) suggests that it may not only evaluate the consequences of an action for oneself, but also for others (Koban et al., 2013; Koban and Pourtois, 2014). In support of this idea, a number of studies have now shown that the AI responds more strongly to vicarious pain when it is the result of an error made by oneself compared with an error made by someone else (Koban et al., 2013; Cui et al., 2015) or by both oneself and someone else (Yu et al., 2014; Cui et al., 2015). Considering the earlier, the AI thus seems well suited to perform an affective evaluation of the degree of agency one has over an observed error. Specifically, observed errors for which an individual shares responsibility may generate an emotional signal in the AI, perhaps in the form of guilt or shame, that encourages the individual to adapt his/her behavior in accordance with the social norm (Koban et al., 2013; Koban and Pourtois, 2014; Yu et al., 2014).

To test these assumptions we designed an interactive task in which participants would sometimes accidentally create a difficult situation for their co-player. Because the co-player was more likely to make an error in the difficult situation, this task allows us to study the influence of agency on the error observation network. Specifically, in the current task, agency over an observed error can be studied both before and after the actual error is made. First, the influence of agency can be investigated before an error is observed by measuring error network activation when the actions of the participant create a difficult situation for the co-player. In line with studies that have shown error activation both in response to actions that have negative consequences for oneself and to actions that have negative consequences for others (Koban et al., 2013; Yu et al., 2014; Cui et al., 2015), we expect the action monitoring network to be activated when an action is likely to cause an error by the co-player. When the difficult situation does indeed result in an error by the co-player, we expect the observation of this error to primarily activate the AI. This is based on the idea that the AI appraises the degree of agency over an observed error. Finally, if the AI signals shared responsibility over an error by generating an emotional signal, we can expect that feelings of guilt and shame mirror the activation pattern in the AI.

Materials and methods

Participants

Twenty-five subjects took part in the experiment (17 females, $M_{\text{age}} = 23.16$, $SD_{\text{age}} = 2.43$) and were paid 28 euro for their participation. However, as explained below, six participants were excluded from analysis. This resulted in a final sample of 19 subjects (14 females, $M_{\text{age}} = 23.00$, $SD_{\text{age}} = 2.58$). All of them were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971), had no history of neurological disorders, and gave written informed consent. The study was approved by the Medical Ethic Review Board of the Ghent University Hospital.

Task, design and stimuli

To test how error observation processes are modulated by agency, we used an adaptation of the social speeded flanker task. In this task, two individuals alternately execute and observe a flanker trial. In the execution trials, participants see an array of arrows and are asked to respond to the middle arrow while ignoring the surrounding arrows. In the observation trials, participants see only the middle arrow and the response given by the other person. Note that in the current experiment one of these persons was a confederate because only one MRI scanner was available. Over the different experimental sessions, four confederates (one male, three female) were used. All confederates first received a short training in which they were familiarized with the task and the protocol. Participants and confederates were not matched on gender. However, because most participants and most confederates were female, most couples were female–female couples.

To introduce an element of agency, the participant was assigned the role of 'agent', while the confederate was assigned the role of 'subject'. For the agent, each execution trial was followed by an agency trial in which the performance of the agent determined the difficulty of the flanker trial for the subject. During these trials, the agent was required to stop a square from switching its color between green and red. Crucially, (s)he was told that stopping the square on green would result in a 'normal' trial for the subject, whereas stopping the square on red would result in a 'difficult' trial for the subject (Figure 1). This was important because the agent was informed that the subject would start with 15 € and win or lose 10 eurocent on each flanker trial depending on whether a correct or an erroneous response was given.

However, a problem with this approach is that it introduces two confounds. First, in the agency task it is not possible to distinguish brain activation related to agency from brain activation related to empathy because the agent is always responsible when the subject receives a difficult trial. Second, in the flanker observation task difficult trials are accompanied by a higher percentage of errors than normal trials. Therefore, observed errors on difficult trials do not only differ from observed errors on normal trials with respect to agency but also with respect to error likelihood. To address these issues, we added a second block (PC block) to the experiment that differed only from the first block (ME block) with respect to agency. That is, in the PC block the agent saw a stationary red or green square on the screen. He or she was told that (s)he had to press a button to activate the computer. The computer would then alternate the square's color a random amount of times between green and red. As a result, it was now the computer that was responsible

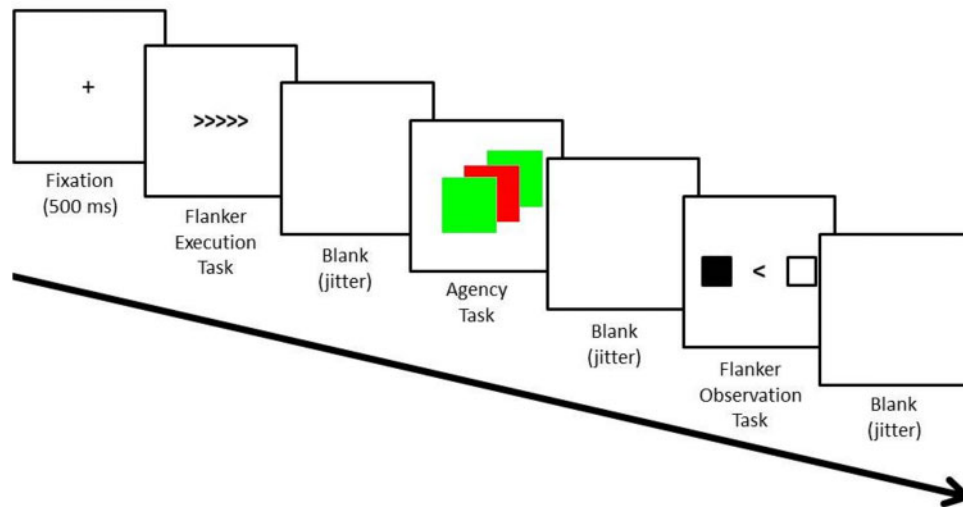


Fig. 1. Trial procedure of the experiment. The only difference between the ME and PC block related to the agency task. In the ME block, the squares started switching color automatically and participants could stop the alternations by pressing a button. In the PC block, the agency task started with a stationary green or red square. When participants pressed a button, the squares started alternating between green and red for a random number of times. The intervals between the different tasks were jittered according to a pseudologarithmic distribution with 50% short intervals (200–2000 ms), 33.3% intermediate intervals (2600 and 4400 ms) and 16.7% long intervals (5000–6800 ms).

for the difficult trials of the subject. Note that this means that the ME and PC block were equal in all aspects except for agency. Comparing brain activation in these two blocks hence allows us to examine the brain correlates of being responsible for another person's errors.

Procedure

The experiment was programmed in Tscope (Stevens et al., 2006). Participants entered the building together with the confederate and the task was explained to them. In addition, it was told that one person (the participant) would perform the task inside the MRI scanner because (s)he had signed up for an fMRI experiment, whereas the other person (the confederate) would perform the task at a computer outside the scanner room because (s)he had signed up for a behavioral experiment. In reality, however, the performance of the confederate was pre-programmed. As outlined earlier, the participant was always assigned the agent role, whereas the confederate was always assigned the subject role.

In order to make it plausible for the participant that (s)he would perform the task together with the confederate, they first practiced an interactive version of the experiment outside of the scanner room. In the practice phase, the participant and confederate sat across each other with a monitor placed in front of each person. It was shown that both monitors were connected to a single computer and it was explained that this made it possible to observe the trials of the other person. Because we wanted participants to get acquainted with the role of agent as well as the role of subject, both roles were practiced for 20 trials each. After the practice phase, the participant and confederate were shown the monitor at which the confederate would sit during the actual experiment. It was shown that this monitor was connected to the same computer as the monitor in the scanner room. Participants were told that this setup would allow them to observe the trials of the confederate even though they would no longer be in the same room.

As explained above, the experiment consisted of two blocks (ME and PC) of 100 trials each. The blocks were counterbalanced and each trial within the blocks could be subdivided into three consecutive tasks, namely (i) flanker execution task, (ii) agency task and (iii) flanker observation task (Figure 1). The task trials were separated by a blank screen that was presented for a jittered duration. The jitter duration followed a pseudo-logarithmic distribution in which 50% of the blank screens was presented for a short period (200–2000 ms in steps of 600 ms), 33.3% for an intermediate period (2600 and 4400 ms in steps of 600 ms) and 16.7% for a long period (5000–6800 ms in steps of 600 ms).

In what follows, we will outline the trial procedure of each task. The flanker execution task was included to define regions of interest (ROI) for the action monitoring network. In this task, participants first saw the flankers for 80 ms (e.g. <<<<) before the relevant stimulus (50% congruent, 50% incongruent) was added (e.g. <<<><<). The presentation time of the relevant stimulus varied between 30 and 80 ms. Specifically, the presentation time was decreased when the error rate was below 30% and increased when the error rate was above 30%. This was done to ensure that the error rate in the flanker execution task was similar to the error rate of the equivalent trials (i.e. normal) of the flanker observation task (i.e. 30%). After the presentation of the whole stimulus array, a mask (#####) appeared on the screen for 120 ms. Starting from the flanker presentation, participants had 700 ms to respond. When subjects did not respond in time, the message 'TOO SLOW!!!' appeared on the screen in red letters for 500 ms.

In the agency task, it was determined on the basis of a green/red alternating square whether the co-player would receive a normal (green) or difficult (red) flanker trial. The trial procedure of the agency task differed slightly between the two blocks. In the ME block, the square automatically alternated between green and red and participants had to stop the square from switching color by pressing a button. Because we expected participants to aim for the green square, the switch time of the square was adjusted online to ensure a sufficient amount of red squares. More precisely,

if >60% of the squares were stopped on green, the switch rate was increased with 25 ms up until a maximum of 100 ms per switch. If <60% of the squares were stopped on green, the switch rate was decreased with 25 ms. In the PC block, participants first saw a stationary green or red square. Participants then had to press a button to activate the computer. Finally, the computer would switch the color of the square a random number of times between two and seven with a random switch rate between 100 and 250 ms before stopping it on green (60%) or red (40%). Subjects were given 3000 ms to respond in both blocks. Once the square was stopped, it remained on the screen in its end color for 1000 ms. When no response was given within the response interval, the color of the square turned red and stayed on the screen for 1000 ms. This was followed by the message "TOO SLOW!!!" in red letters for 500 ms.

In the flanker observation task, participants observed the flanker trial of the other person. The stimulus display in this task was equal for the normal and difficult trials. Note, however, that participants experienced both trial types themselves in the practice phase. Participants saw the middle arrowhead of the co-player's stimulus together with two squares that represented his/her response buttons. When a response was given, the corresponding square turned black for 500 ms. Because there was no actual co-player doing the task, responses were determined by the computer. In order to simulate the response behavior of a real person, we drew the reaction time from a normal distribution with a mean of 500 ms and a standard deviation of 100 ms (for a similar procedure see Castellar et al., 2011). Furthermore, limits were imposed on the reaction time so it was never faster than 300 ms or slower than 700 ms. Finally, to make the difference in difficulty between normal and hard trials believable, the computer was programmed to make an error on 30% of the normal trials and on 50% of the difficult trials.

On 5% of the trials within each block, two questions were asked after the flanker observation task. This was done to check whether participants were attentive to the task. The first question asked if the other person received a normal or a difficult flanker trial. The second question asked if the other person gave a correct or an erroneous response. Subjects were required to respond to these questions by pressing a button.

Finally, after the experiment, participants were debriefed and filled out a number of questions that measured feelings of guilt and shame during the experiment (see Supplementary Materials). Participants were explicitly asked whether they believed that they played with another person during the experiment. All participants indicated that they believed this.

Image acquisition

MRI images were acquired with a 3 T scanner (Siemens Trio) combined with a 32-channel radiofrequency head coil. Participants were entered head first and supine into the scanner. The scanning procedure started with an anatomical scan in which 176 high-resolution anatomical images were acquired using a T1-weighted 3D MPRAGE sequence [repetition time (TR)=2530 ms, echo time (TE)=2.58 ms, image matrix=256 × 256, field of view (FOV)=220 mm, flip angle=78, slice thickness=0.90 mm, voxel size=0.9 × 0.86 × 0.86 mm (resized to 1 × 1 × 1 mm)]. Next, two runs (i.e. ME block and PC block) were conducted in which whole-brain functional images were obtained. These functional images were acquired using a T2*-weighted echo planar imaging (EPI) sequence, sensitive to BOLD contrast (TR=2000 ms, TE=28 ms, image matrix=64 × 64, FOV=224 mm, flip angle=80°, slice thickness=3.0 mm,

distance factor=17%, voxel size 3.5 × 3.5 × 3 mm, 34 axial slices). The amount of EPI images in the two runs depended on the speed at which participants performed the task.

Statistical analyses

All data was analyzed with SPM8 (Wellcome Department of Imaging Neuroscience, UCL, London, UK; www.fil.ion.ucl.ac.uk/spm). In order to account for T1 relaxation effects, the first four scans of both runs were dummy scans. With regard to the preprocessing, all functional images were first spatially realigned using a rigid body transformation. Second, the functional images were slice-time corrected with respect to the middle acquired slice. Next, the structural image of each subject was co-registered with their mean functional image. T1 images were then segmented according to the tissue probability maps available in SPM and the parameters estimated during segmentation were used to normalize the functional images to standard MNI space. Finally, the images were resampled into 3 mm³ voxels and spatially smoothed with a Gaussian kernel of 8 mm (full-width at half maximum).

A high-pass filter of 128 Hz was applied during fMRI data analysis. First level analyses were performed using the general linear model as applied by SPM8. A model was created with separate regressors for the three tasks (i.e. flanker execution task, agency task and flanker observation task). In addition, twelve regressors (i.e. six for each run) were added to the model to control for head movement artifacts. The hemodynamic response function (HRF) in the flanker execution task was modeled for error and correct trials at the time of the response. The HRF in the agency task was modeled when the square's end color was presented with block (ME vs PC) and outcome (normal vs difficult) as factors. The HRF in the flanker observation task was modeled at the time of the observed response with block (ME vs PC), difficulty (normal vs difficult) and accuracy (error vs correct) as factors. Contrast images of interest were created at the first level and were then entered into a second level analysis using one-sample *t* tests. We will focus on brain activation that was significant at the cluster-level ($P < 0.05$, FWE-corrected) after applying a whole brain uncorrected threshold of $P < 0.001$ at the peak-level. The ROI analyses were conducted using the MARSBAR package for SPM8 (Brett et al., 2002).

Six participants were excluded from the analyses. One participant was excluded for favoring the red square over the green square in the agency task (60% red). Another participant was excluded because (s)he did not manage to distinguish between normal and difficult trials in the PC block (probe accuracy: 40%). Finally, four participants were excluded due to excessive movement (>3 mm).

For the remaining participants, the following trials were removed from the data analysis. For the flanker execution task analyses, we excluded the trials in which the flanker task RT exceeded the response deadline together with the trials in which the flanker task RT was faster than 200 ms (8.62%). For the agency task analyses, we excluded trials in which the agency task RT exceeded the response deadline (1.03%). Finally, for the flanker observation analyses, we excluded trials in which the agency task RT exceeded the response deadline and trials that were followed by a probe (5.90%).

Results

Behavioral data

In the flanker execution task, the mean response time was 490 ms (SD=50 ms) and the mean error rate was 24.21%

(SD = 10.27%). In the agency task, the mean response time of the participants was 1152 ms (SD = 363 ms) in the ME block and 786 ms (SD = 156 ms) in the PC block. In the ME block, the square was stopped on red in 40.39 % (SD = 0.92%) of the trials.

The question on the difficulty of the co-player's trial was answered correctly in 97% (SD = 8%) of the probe trials in the ME block and in 94% (SD = 9%) of the probe trials in the PC block. A paired samples *t* test revealed that the percentage of correctly answered probe trials did not differ between the two blocks, $t = 1$. The question on the accuracy of the co-player's trial was answered correctly in 97% (SD = 8%) of the probe trials in the ME block and in 97% (SD = 7%) of the probe trials in the PC block. Again, a paired samples *t* test revealed no difference between the two blocks, $t < 1$.

Questionnaire data

Feelings of guilt and shame after observing an error are summarized in Table 1 for the different experimental conditions. In order to test whether participants felt guiltier and more ashamed after observing an error for which they were responsible, we subjected the questionnaire scores to a block (ME vs PC) \times difficulty (normal vs difficult) repeated measures ANOVA. This revealed a significant block \times difficulty interaction both for feelings of guilt, $F(1, 18) = 28.25$, $P < 0.001$, and for feelings of shame, $F(1, 18) = 11.03$, $P = 0.004$. Follow-up analyses showed that participants felt guiltier for observed errors on difficult compared with normal trials in both the ME block, $t(18) = 8.49$, $P < 0.001$, and the PC block, $t(18) = 2.69$, $P = 0.015$, but the difference was considerably larger in the ME block. With regard to feelings of shame, follow-up analyses revealed that participants felt more ashamed for observed errors on difficult compared with normal trials in the ME block, $t(18) = 4.32$, $P < 0.001$, but not in the PC block, $t < 1$.

fMRI data

Flanker execution task. As expected, error related activation (Error > Correct) was found in the rostral cingulate zone (RCZ), the pre-supplementary motor area (pre-SMA), the left AI, and the right middle frontal gyrus (MFG) (Figure 2 and Table 2). These activation clusters were subsequently used as ROIs for analyses in the agency task and flanker observation task. Because activation in the pMFC extended over two anatomically distinct brain regions and because these regions have been related to different aspects of error processing (e.g. Ullsperger and von Cramon, 2001, 2004; Desmet et al., 2011; Yoshida et al., 2012), we created a separate ROI for the ventral pMFC (vpMFC) and the dorsal pMFC (dpMFC) by splitting the pMFC activation cluster in two at the border of the pre-SMA and the RCZ.

Agency task. To examine the brain response to actions that were likely to elicit an error by the co-player, we subjected activation in the error monitoring ROIs to a block (ME vs PC) \times outcome (normal vs difficult) repeated measures ANOVA (Figure 3). This revealed a significant block \times outcome interaction in the vpMFC, $F(1, 18) = 12.54$, $P = 0.002$, the dpMFC, $F(1, 18) = 28.92$, $P < 0.001$, the AI, $F(1, 18) = 13.61$, $P = 0.002$, and the LPFC, $F(1, 18) = 17.26$, $P = 0.001$. Follow-up two-tailed *t* tests showed that brain activation in all ROIs was increased in the ME block, all $t(18) \geq 3.52$, all $P \leq 0.002$, but not in the PC block, all $t(18) \geq 3.52$, all $P \leq 0.002$, when the outcome for the co-player was a difficult compared with a normal trial.

Table 1. Guilt and shame ratings for observed errors in the different conditions

	Guilt	Shame
ME normal	2.63 (1.46)	2.42 (1.31)
ME difficult	5.53 (1.12)	3.95 (1.75)
PC normal	2.05 (1.39)	1.90 (0.99)
PC difficult	2.74 (1.79)	2.00 (1.33)

Note. Means are displayed with standard deviations between parentheses. Ratings were given on a 7-point Likert scale.

Flanker observation task. For the flanker observation task, we first established whether the error execution defined ROIs were also involved in error observation. To this end, we compared observed errors with observed correct responses in the normal trials separately for each ROI. Note that we focused on these trials because they resembled the flanker execution trials most closely in terms of error likelihood (i.e. 30%). Two-sided paired *t* tests revealed no difference in the vpMFC, $t < 1$, but a stronger response to observed errors in the dpMFC, $t(18) = 3.07$, $P = 0.007$, AI, $t(18) = 2.65$, $P = 0.016$, and the LPFC, $t(18) = 2.63$, $P = 0.017$.

Next, we investigated whether participants processed observed errors differently depending on whether or not they shared responsibility for these errors (Figure 4). For this purpose, we first removed confounding activation attributable to error likelihood by subtracting brain activation in the PC block from brain activation in the ME block separately for each condition (see 'Materials and methods' section). The obtained difference scores were then subjected to an accuracy (correct vs error) \times difficulty (normal vs difficult) repeated measures ANOVA. We first computed the main effect of accuracy to test if the brain response to observed errors differed between the ME and PC block. This revealed a near-significant accuracy main effect in the dpMFC, $F(1, 18) = 4.37$, $P = 0.051$, indicating stronger error observation activation in the ME block compared with the PC block. No main effect of accuracy was observed in the other ROIs, all $F < 1$. We then computed the difficulty \times accuracy interaction to test the influence of agency on error observation while controlling for differences in error likelihood. This revealed a significant interaction in the AI, $F(1, 18) = 7.54$, $P = 0.013$. Follow-up two-tailed *t* tests confirmed that this was due to increased brain activation for observed errors compared with observed correct responses on the difficult trials, $t(18) = 2.87$, $P = 0.010$, but not on the normal trials, $t < 1$. The difficulty \times accuracy interaction did not reach significance in the three remaining ROIs, all $F(1, 18) \leq 1.27$, all $P \geq 0.275$.

Discussion

Despite the fact that our actions often have a strong influence on the behavior of others, research on error observation has hitherto focused primarily on the passive observation of errors (Shane et al., 2008; Newman-Norlund et al., 2009; Howard-Jones et al., 2010; Monfardini et al., 2013; Desmet et al., 2014). To extend error observation research to more realistic interactive contexts, we therefore investigated how agency over an observed error can affect how these errors are processed. More specifically, we designed a task in which the actions of the participants determined whether another person would get a normal or a difficult trial. This task allowed us to study agency over observed errors both before and after the actual error was made.

The study yielded three important results. First, it was found that the error network was activated when participants created

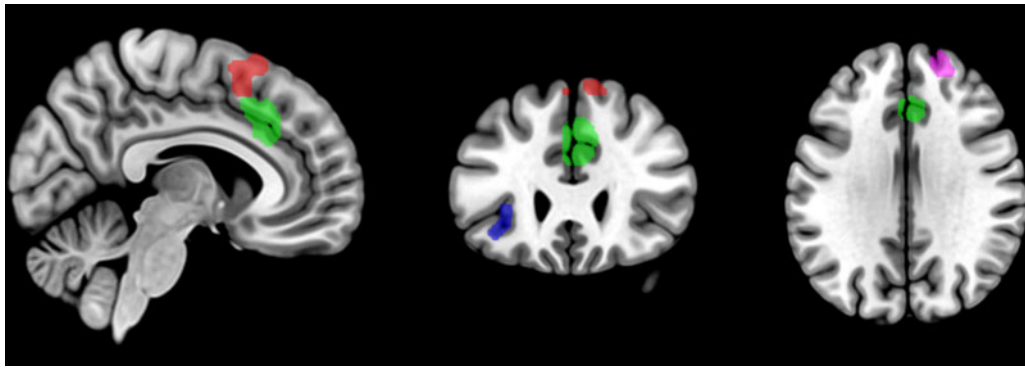


Fig. 2. Sagittal ($x=6$), coronal ($y=27$) and axial ($z=31$) planes depicting brain activation resulting from the flanker execution task contrast Error > Correct (see fig. S1 for t-score maps). Based on this activation, ROIs were created for the analysis of the agency task and the flanker observation task. Activation clusters in the figure are color coded to represent the different ROIs. The red area represents the dpMFC ROI, the green area represents the vpMFC ROI, the blue area represents the AI ROI, and the purple area represents the LPFC ROI. Note that the vpMFC and dpMFC ROI were created by splitting the pMFC activation cluster in two at the border of the pre-SMA and the RCZ.

a difficult situation for someone else. Crucially, the same activation was not found when the difficult situation was created by the computer, excluding the possibility that it reflected a general empathic response. Instead, the results suggest that an action that will likely lead to an error by someone else is evaluated in the same way as an error. This is in line with previous studies that revealed error network activation when an action caused pain to someone else (Koban *et al.*, 2013; Yu *et al.*, 2014; Cui *et al.*, 2015) and confirms that the action monitoring system is not only sensitive to actions with negative consequences for oneself, but also to actions with negative consequences for other persons.

Second, we confirmed previous studies in which activation in the dpMFC, LPFC and AI was found in response to both self-generated (e.g. Ullsperger and von Cramon, 2001, 2004; Marco-Pallarés *et al.*, 2008) and observed (e.g. Shane *et al.*, 2008; De Bruijn *et al.*, 2009; Desmet *et al.*, 2014) errors. Activation in the vpMFC, on the other hand, was found to be restricted to self-generated errors. Interestingly, a similar dissociation was recently reported in a single-neuron study in macaques (Yoshida *et al.*, 2012). This study revealed that the ventral and dorsal regions of the pMFC have different functions within the error observation process. More specifically, it was found that the dpMFC is involved in the general detection of vicarious errors, whereas the vpMFC is involved in adapting personal behavior to these errors. Applied to the current study, the absence of error observation related activation in the vpMFC could be explained by the fact that observed errors were uninformative to improve personal performance. That is, the stimulus display in the flanker observation task provided information about the outcome of the co-player's trial, but it did not provide information on how this outcome was achieved. Without the latter, however, observed errors do not inform participants how they should adapt their behavior to increase personal performance.

Finally, sense of agency over an observed error differentially modulated brain activation in the previously defined error observation network. In particular, it was found that activation in the AI but not in the pMFC or LPFC was sensitive to whether or not the error was caused by the observer. Interestingly, the AI activation pattern was accompanied by higher feelings of guilt and shame when seeing errors for which participants bore responsibility (see also Koban *et al.*, 2013; Yu *et al.*, 2014). In accordance with the idea that the AI is associated with the affective evaluation of intentional behavior

(Brass and Haggard, 2010; Koban and Pourtois, 2014), these findings suggest that the AI performs an emotional evaluation of observed errors based on the degree of agency over the error. Interestingly, this interpretation fits well with a recent theoretical framework concerning the influence of social factors in error monitoring (Koban and Pourtois, 2014). In this framework, it was proposed that the different regions within the error monitoring network perform different yet complementary functions. Specifically, it was argued that the role of the pMFC is to determine whether an action is erroneous or correct, whereas the role of the AI is to evaluate the resulting information in a broader context of agency and social factors. Although this framework was primarily constructed to account for findings in the domain of error execution, the findings of the current study suggest that similar functions can be ascribed to these regions during error observation. Taken together with previous studies on error execution (Koban *et al.*, 2013; Yu *et al.*, 2014; Cui *et al.*, 2015), the results of the current study thus support the idea that the AI generates an affective signal that indicates the need to adjust our behavior when our actions cause another person harm. As such, the AI can be thought of as an important brain region in regulating social behavior.

It should be noted however, that it appears from Figure 4 that agency primarily influenced AI responses to observed correct responses rather than to observed erroneous responses. However, directly comparing activation between the two blocks and between the two situations is not trivial in the flanker observation task. The fact that participants created a difficult situation for the co-player in the difficult ME trials makes it likely that an emotional response was generated in this situation even before the performance of the co-player was observed. Such emotional responses are known to be the target of emotion regulation processes (Ochsner and Gross, 2005, 2008) that inhibit activation in emotion areas such as the AI (e.g. Ochsner and Gross, 2005, 2008; Goldin *et al.*, 2008; Grecucci *et al.*, 2013). It is therefore probable that overall AI activation was reduced in the flanker observation task when participants had created a difficult situation for the co-player. Such baseline differences strongly limit the conclusions that one can draw from a direct comparison between the two blocks or between the two situations. Therefore, we instead chose to focus on relative comparisons based on the difference in activation between observed errors and observed correct responses. These relative comparisons clearly indicated that agency modulated the AI response to observed errors.

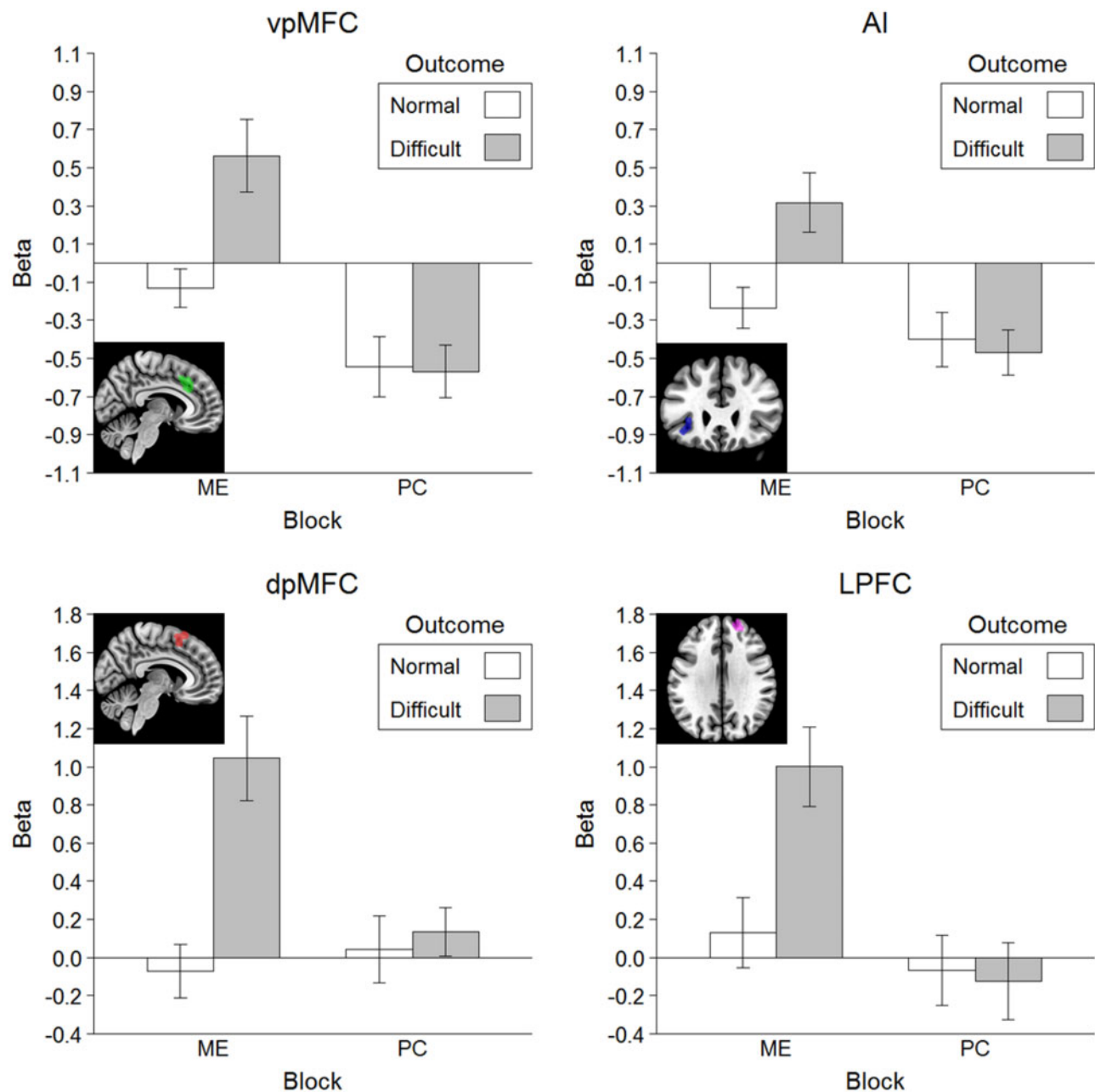


Fig. 3. Results of the agency task ROI analysis. Error bars represent standard errors of the mean corrected for within subject designs according to Morey (2008).

Table 2. MNI Coordinates of the Error > Correct contrast in the flanker execution task

	Peak coordinates	Z-score	Cluster size
RCZ, pre-SMA	3, 26, 37	4.91	366
	9, 23, 61	4.27	
	-3, 11, 64	4.18	
Right MFG	21, 53, 31	4.24	78
Left AI, IFG	-42, 20, -11	4.15	89
	-30, 29, 1	3.82	

RCZ, rostral cingulate zone, pre-SMA, pre-supplementary motor area; MFG, middle frontal gyrus; AI, anterior insula; IFG, inferior frontal gyrus.

In contrast to the activation pattern of the AI, activation in the LPFC and dpMFC suggested that these regions perform a more general error detection function during error observation (e.g. Yoshida et al., 2012). Interestingly, however, the observation that dpMFC activation to vicarious errors was still seen after activation in the PC block was subtracted from activation in the ME block suggests that error detection in this region is modulated by context. Specifically, it indicates that the dpMFC is more involved in monitoring the actions of others in an active compared with a passive situation. At least two explanations could be postulated for this finding. First, it is possible that individuals simply attended more closely to the task of the other person in the ME block. However, this seems unlikely given that probe questions were answered equally well in both blocks. Second, it is possible that vicarious errors in the ME block triggered the dpMFC more strongly because the active context caused participants to

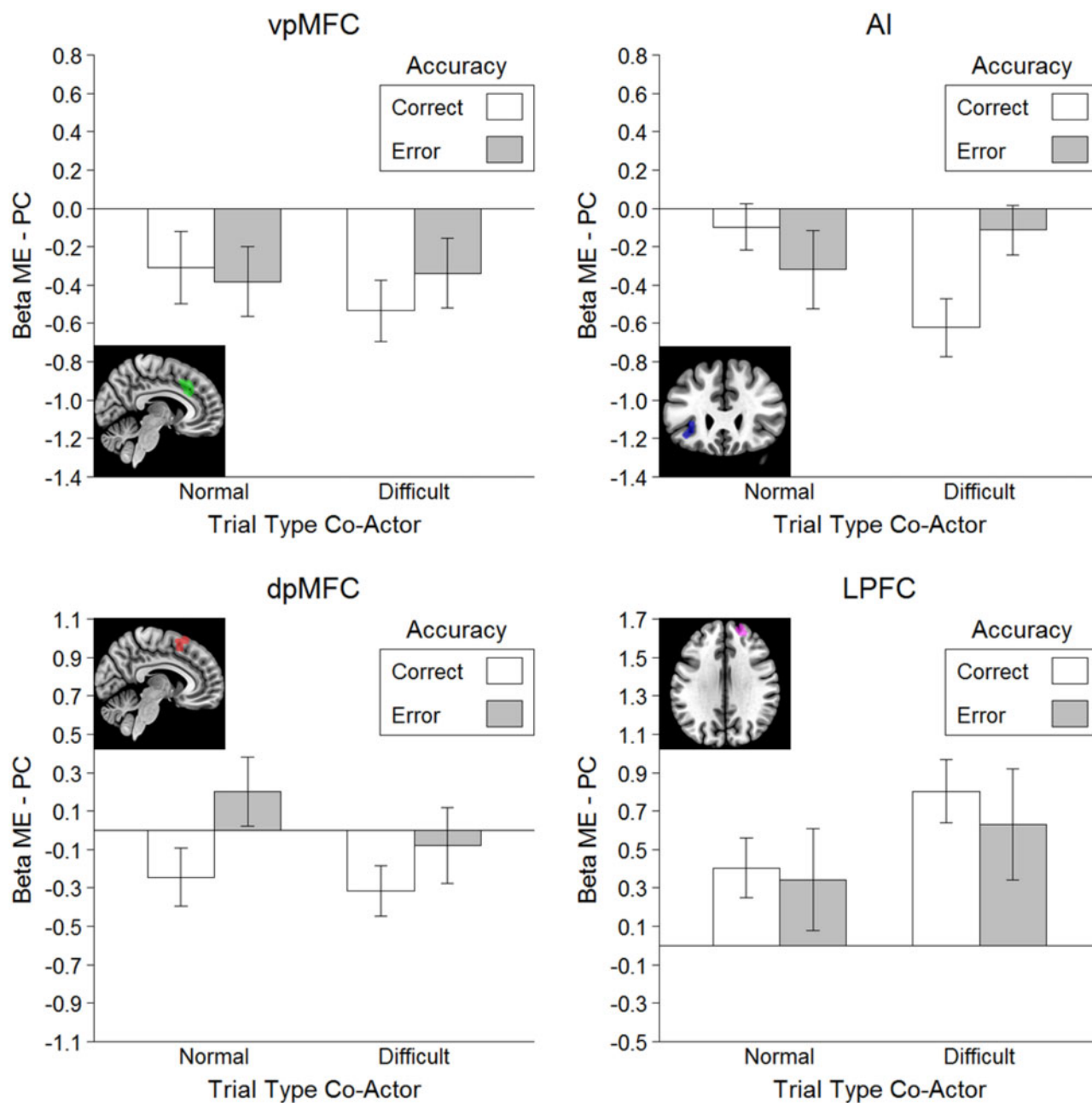


Fig. 4. Results of the flanker observation task ROI analysis. Each bar represents the difference score between the ME and the PC block (ME-PC) to control for the confounding influence of error likelihood. Error bars represent standard errors of the mean corrected for within subject designs according to Morey (2008). Because the gender distribution was unbalanced, we additionally checked whether the hypothesized pattern in the AI was present for both males and females. A visual analysis confirmed that the same pattern was obtained for both genders (see Supplementary Figure S2).

empathize more with the other person. This is in line with previous studies on error observation that have linked activation in the error network to individual differences in empathy (e.g. Newman-Norlund *et al.*, 2009; Shane *et al.*, 2009).

In summary, the current study examined the influence of agency on activation in the error observation network. In accordance with the idea that the outcome of an action is first evaluated in the pMFC and then forwarded to the AI where it is integrated with information on agency and social context (Koban *et al.*, 2013; Koban and Pourtois, 2014), it was found that activation in the AI but not the pMFC was modulated by agency over the observed error. Interestingly, the activation pattern in the AI was mirrored by feelings of guilt and shame. This

suggests that the AI is crucially involved in evaluating the social consequences of an action.

Acknowledgements

We thank Alko van der Wiel, Ena Coenen, Stephanie Cobbaert and Amanda Clauwaert for acting as confederates in this study.

Funding

This work was supported by a grant from the Research Council of Ghent University (B/13089).

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Bates, A.T., Patel, T.P., Liddle, P.F. (2005). External behavior monitoring mirrors internal behavior monitoring. Error-related negativity for observed errors. *Journal of Psychophysiology*, **19**(4), 281–8.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**(3), 624–52.
- Brass, M., Haggard, P. (2010). The hidden side of intentional action: the role of the anterior insular cortex. *Brain Structure and Function*, **214**, 603–10.
- Brett, M., Anton, J.L., Valabregue, R., Poline, J.B. (2002). Region of interest analysis using an SPM toolbox. *Abstract Presented at the 8th International Conference on Functional Mapping of the Human Brain*. Sendai, Japan.
- Brown, J.W., Braver, T.S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, **307**, 1118–21.
- Castellar, E.N., Notebaert, W., Van Den Bossche, L., Fias, W. (2011). How monitoring other's actions influences one's own performance: post-error adjustments are influenced by the nature of the social interaction. *Experimental Psychology*, **58**(6), 499–508.
- Craig, A.D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, **10**, 59–70.
- Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., Dolan, R.J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, **7**(2), 189–95.
- Cui, F., Abdelgabar, A.-R., Keysers, C., Gazzola, V. (2015). Responsibility modulates pain-matrix activation elicited by the expressions of others in pain. *NeuroImage*, **114**, 371–78.
- De Bruijn, E.R.A., de Lange, F.P., Von Cramon, D.Y., Ullsperger, M. (2009). When errors are rewarding. *The Journal of Neuroscience*, **29**(39), 12183–6.
- Desmet, C., Deschrijver, E., Brass, M. (2014). How social is error observation? The neural mechanisms underlying the observation of human and machine errors. *Social Cognitive and Affective Neuroscience*, **9**, 427–35.
- Desmet, C., Fias, W., Hartstra, E., Brass, M. (2011). Errors and conflict at the task level and the response level. *The Journal of Neuroscience*, **31**(4), 1366–74.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., Blanke, L. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Brunia, C., Gaillard, A., Kok, A., editors, pp. 192–5. *Psychophysiological Brain Research*. Tilburg: Tilburg University Press.
- Goldin, P.R., McRae, K., Ramel, W., Gross, J.J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, **63**(6), 577–86.
- Grecucci, A., Giorgetta, C., Bonini, N., Sanfey, A.G. (2013). Reappraising social emotions: the role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. *Frontiers in Human Neuroscience*, **7**, 523.
- Holroyd, C.B., Coles, M.G.H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, **109**(4), 679–709.
- Howard-Jones, P.A., Bogacz, R., Yoo, J.H., Leonards, U., Demetriou, S. (2010). The neural mechanisms of learning from competitors. *NeuroImage*, **53**(2), 790–9.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, **303**, 1023–6.
- Koban, L., Corradi-Dell'Acqua, C., Vuilleumier, P. (2013). Integration of error agency and representation of others' pain in the anterior insula. *Journal of Cognitive Neuroscience*, **25**, 258–72.
- Koban, L., Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neuroscience and Biobehavioral Reviews*, **46**, 1–14.
- Koban, L., Pourtois, G., Vocat, R., Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Social Neuroscience*, **5**, 360–74.
- Lamm, C., Decety, J., Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, **54**(3), 2492–502.
- Lepron, E., Causse, M., Farrer, C. (2015). Responsibility and the sense of agency enhance empathy for pain. *Proceedings of the Royal Society B-Biological Sciences*, **282**(1799), 20142288.
- Marco-Pallarés, J., Camara, E., Münte, T.F., Rodríguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *Journal of Cognitive Neuroscience*, **20**, 1595–610.
- Miltner, W.H.R., Brauer, J., Hecht, H., Trippe, R., Coles, M.G.H. (2004). Parallel brain activity for self-generated and observed errors. In: Ullsperger, M., Falkenstein, M., editors, pp. 124–9. *Errors, Conflicts and the Brain. Current Opinions on Performance Monitoring*. Leipzig: Max Plank Institute of Cognitive Neuroscience.
- Monfardini, E., Gazzola, V., Boussaoud, D., Brovelli, A., Keysers, C., Wicker, B. (2013). Vicarious neural processing of outcomes during observational learning. *PLoS One*, **8**(9), e73879.
- Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, **4**(2), 61–4.
- Newman-Norlund, R.D., Ganesh, S., van Schie, H.T., De Bruijn, E.R.A., Bekkering, H. (2009). Self-identification and empathy modulate error-related brain activity during the observation of penalty shots between friend and foe. *Social Cognitive and Affective Neuroscience*, **4**, 10–22.
- Ochsner, K.N., Gross, J.J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, **9**(5), 242–9.
- Ochsner, K.N., Gross, J.J. (2008). Insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science*, **17**(2), 153–8.
- Oldfield, R.C. (1971). Assessment and analysis of handedness – Edinburgh Inventory. *Neuropsychologia*, **9**, 97–113.
- Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, **306**(2004), 443–7.
- Rushworth, M.F.S., Behrens, T.E.J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, **11**(4), 389–97.
- Shane, M.S., Stevens, M., Harenski, C.L., Kiehl, K.A. (2008). Neural correlates of the processing of another's mistakes: a possible underpinning for social and observational learning. *NeuroImage*, **42**, 450–9.
- Shane, M.S., Stevens, M., Harenski, C.L., Kiehl, K.A. (2009). Double dissociation between perspective-taking and empathic concern as predictors of hemodynamic response to another's mistakes. *Social Cognitive and Affective Neuroscience*, **4**, 111–8.

- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, **303**, 1157–62.
- Stevens, M., Lammertyn, J., Verbruggen, F., Vandierendonck, A. (2006). Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*, **38**(2), 280–6.
- Taylor, S.F., Stern, E.R., Gehring, W.J. (2007). Neural systems for error monitoring: recent findings and theoretical perspectives. *The Neuroscientist*, **13**, 160–72.
- Ullsperger, M., Danielmeier, C., Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, **94**, 35–79.
- Ullsperger, M., Harsay, H.A., Wessel, J.R., Ridderinkhof, K.R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure and Function*, **214**(5–6), 629–43.
- Ullsperger, M., von Cramon, D.Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, **14**(6), 1387–401.
- Ullsperger, M., von Cramon, D.Y. (2004). Neuroimaging of performance monitoring: error detection and beyond. *Cortex*, **40**, 593–604.
- van Schie, H.T., Mars, R.B., Coles, M.G.H., Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, **7**(5), 549–54.
- Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, **40**, 655–64.
- Yoshida, K., Saito, N., Iriki, A., Isoda, M. (2012). Social error monitoring in macaque frontal cortex. *Nature Neuroscience*, **15**(9), 1307–12.
- Yu, H., Hu, J., Hu, L., Zhou, X. (2014). The voice of conscience: Neural bases of interpersonal guilt and compensation. *Social Cognitive and Affective Neuroscience*, **9**, 1150–8.