Psychological Medicine (2016), **46**, 1289–1300. © Cambridge University Press 2016
doi:10.1017/S0033291715002834

ORIGINAL ARTICLE

Neural correlates of social exchanges during the Prisoner's Dilemma game in depression

V. B. Gradin[1]*, A. Pérez[1], J. A. Macfarlane[2], I. Cavin[2], G. Waiter[3], E. B. Tone[4], B. Dritschel[5], A. Maiche[1] and J. D. Steele[6]

[1] *Faculty of Psychology, CIBPsi, Universidad de la República, Montevideo, Uruguay*
[2] *Medical Physics, NHS Tayside, University of Dundee, Dundee, UK*
[3] *Aberdeen Biomedical Imaging Centre, University of Aberdeen, Aberdeen, UK*
[4] *Department of Psychology, Georgia State University, Atlanta, GA, USA*
[5] *Department of Psychology, University of St Andrews, Fife, UK*
[6] *School of Medicine (Neuroscience), University of Dundee, Ninewells Hospital and Medical School, Dundee, UK*

Background. Depression is a disabling disorder that significantly impacts on the interpersonal functioning of individuals. However, little is known about the neural substrates of such difficulties. In the last few years neuroeconomics, which combines imaging with multiplayer behavioural economic paradigms, has been used to study the neural substrates of normal and abnormal interpersonal interactions.

Method. This study used functional magnetic resonance imaging to investigate neural activity in unmedicated depressed participants ($n = 25$) and matched healthy controls ($n = 25$). During scanning, participants played a behavioural economic game, the Prisoner's Dilemma. In this game, the participant and a co-player independently choose either to cooperate or not cooperate with each other.

Results. Depressed participants reported higher levels of negative feelings (betrayal, guilt) during the game than did controls. Neural activation was compared between 'imbalanced' events [when one of the players cooperated and the other defected ('CD' and 'DC')] and 'draw' events [when both players either cooperated or defected ('CC' and 'DD')]. Participants preferentially activated the anterior insula and the dorsolateral prefrontal cortex (DLPFC), a region implicated in cognitive control and regulation of emotions. Importantly, compared to controls depressed participants showed reduced activation in the left DLPFC, with the extent of signal reduction correlating with increased self-report feelings of guilt associated with DC outcomes.

Conclusions. Our findings suggest that depression is associated with reduced activation of the DLPFC during social events that involve unreciprocated cooperation. This abnormality may underlie anomalies in cognitive control and top-down regulation of emotions during challenging social exchanges.

Received 31 August 2015; Revised 26 November 2015; Accepted 27 November 2015; First published online 14 January 2016

Key words: Depression, fMRI, neuroeconomics, Prisoner's Dilemma

Introduction

Depression is a common and disabling disorder that can profoundly affect how an individual interacts with others. People experiencing depression report difficulties maintaining and enjoying relationships, less supportive social networks, less active social lives, excessive reassurance seeking, poor intimate relationships and in general, more maladaptive and less satisfactory social interactions (Billings *et al.* 1983; Fredman *et al.* 1988; Hirschfeld *et al.* 2000; Segrin, 2000; Papakostas *et al.* 2004).

Despite the importance of interpersonal difficulties in psychiatric disorders such as depression, the neurobiology of such impairments remains largely understudied, partly due to difficulties in recreating and quantifying interpersonal exchanges (King-Casas & Chiu, 2012). In the last few years, neuroeconomic approaches (Glimcher & Rustichini, 2004) which combine interactive behavioural economic tasks with neuroimaging, have been used to study interpersonal functioning and its neural substrates in clinical populations (Rilling *et al.* 2007; King-Casas *et al.* 2008; Hasler, 2011; McClure-Tone *et al.* 2011; King-Casas & Chiu, 2012; Gradin *et al.* 2015). Behavioural economic tasks involve multiplayer interactive scenarios that allow quantification of social exchanges and the study of social concepts such as fairness, cooperation, and trust.

* Address for correspondence: Dr V. Gradin, Faculty of Psychology, Center for Investigation of Basic Psychology (CIBPsi), Universidad de la República, Montevideo 11200, Uruguay.
(Email: victoriagradin@gmail.com)

A well known behavioural economic task is the Prisoner's Dilemma (PD) game (Axelrod & Hamilton, 1981). The PD allows examination of social relationships based on cooperative and uncooperative behaviours. In each round two players independently choose to either 'cooperate' with or 'defect' from each other. There are four possible outcomes: both players cooperate (CC), one of the players cooperates and the other defects (CD, DC) or both players defect (DD). Depending on the outcome, each player is awarded a sum of money (see the payoff matrix, Fig. 1*a*).

Neuroimaging studies using economic behavioural tasks have reported that *positive* social exchanges activate regions of the reward circuitry, much as non-social rewards do. For example, studies using the Ultimatum Game – a task where participants accept or reject monetary offers made by others – have shown that fair offers typically activate the striatum in the receiver (Tabibnia *et al.* 2008; Crockett *et al.* 2013; Gradin *et al.* 2015). Similarly, studies using the PD have reported striatal activation in response to reciprocated cooperation (Rilling *et al.* 2002). In contrast, unfair offers during the Ultimatum Game have been reported to activate regions implicated in processing aversive emotions and salience detection (anterior insula), cognitive conflict (dorsal anterior cingulate and dorsomedial prefrontal cortex), cognitive control and emotion regulation [dorsolateral prefrontal cortex (DLPFC)] (Sanfey *et al.* 2003). Unreciprocated cooperation during the PD has also been linked to anterior insula activation (Rilling *et al.* 2008). In addition, it has been reported that the PD is associated with increased activation of the DLPFC when compared to the 'stag hunt game', possibly because the PD places higher cognitive control demands on participants (Emonds *et al.* 2012), although other work has shown that the stag hunt game can involve relatively demanding mental computations (Yoshida *et al.* 2008, 2010).

Human studies of brain function in depression have reported a number of abnormalities. First, several studies indicate reduced activation in reward-related brain regions, particularly the striatum, in depression; this reduction may be linked to anhedonia (Eshel & Roiser, 2010; Gradin *et al.* 2011; Zhang *et al.* 2013). In addition, a recent study using the Ultimatum Game reported reduced striatal responses to increasing fairness of offers in depression, suggesting diminished responsiveness not only to material rewards but also to social rewards (Gradin *et al.* 2015).

Depression has also been linked to emotion regulation models. These models hypothesize that depression is associated with hyperactivity of limbic regions that are involved in detecting emotions (bottom-up processes), and also with abnormal functioning of regions higher in the cognitive hierarchy, such as the DLPFC, resulting in abnormalities in control and regulation of emotions (top-down processes) (Gotlib & Hamilton, 2008; Disner *et al.* 2011; Rive *et al.* 2013).

In this study, we investigated neural activation in unmedicated depressed participants and healthy controls while they played the PD. Based on evidence that neural dysfunction in reward-related regions characterizes depressed individuals, we first hypothesized that depressed participants would show diminished striatal responses to reciprocated cooperation in comparison to controls. Second, based on evidence of neural anomalies in regions implicated in both bottom-up and top-down emotion processes in depressed individuals, we predicted that unreciprocated cooperation (CD) would be associated with enhanced activity in regions such as the insula, that are involved in processing emotionally salient aversive stimuli, and abnormal activity in regions such as the DLPFC, that are involved in top-down regulation of emotions.

Finally, we investigated emotional and neural responses during outcomes in which the participant defected while the co-player cooperated (DC). While this type of outcome may trigger some positive emotions (as the participant receives the highest payoff) it may also trigger the negative emotion of guilt (Rilling *et al.* 2007). Since excessive feelings of guilt are a core symptom of depression (APA, 2013), it was hypothesized that DC outcomes would lead to enhanced feelings of guilt in depression. At the neural level, it was hypothesized that, like unreciprocated cooperation (CD), DC outcomes could be associated with hyperactivity of emotion detection regions and with abnormal activation of control and emotion regulation regions such as the DLPFC.

Method

Participants

The study was approved by the local Research Ethics Committee and written informed consent was obtained from all participants. Data were acquired from 25 participants meeting criteria for an episode of DSM-IV depression and 25 healthy controls. The study was advertised within the universities of Dundee and St Andrews, UK. Potential participants were invited to self-nominate either for the depression or control group. Applicants were invited to a recruitment session (approximately 3–7 days before scanning) and were screened for depression and other psychiatric symptoms using the Mini International Neuropsychiatric Interview (MINI Plus v. 5.0) and symptom burden quantified using the Beck Depression Inventory (BDI; Beck *et al.* 1961). Inclusion criteria for the depression group were: satisfying DSM-IV criteria for a major

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

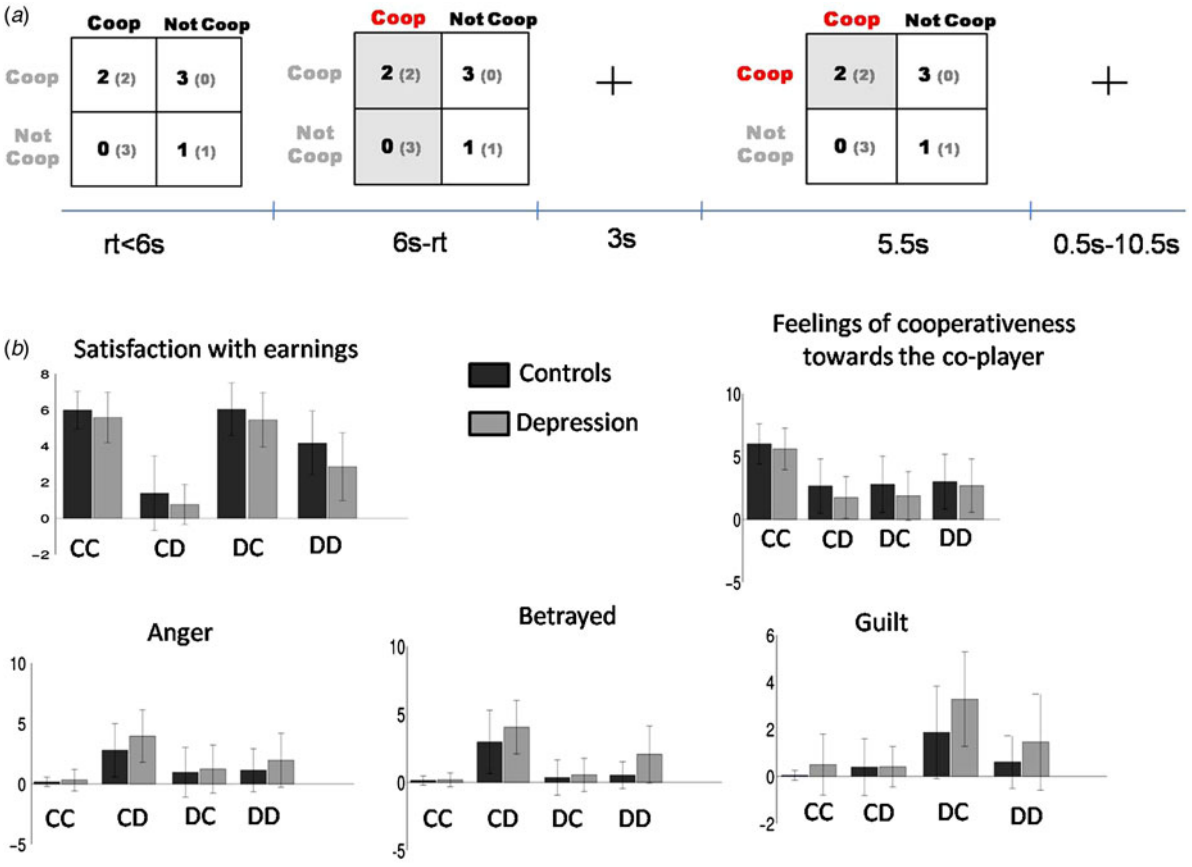

Fig. 1. The Prisoner's Dilemma (PD) game and emotional results. (*a*) On each trial the participant and a (supposed) co-player make a simultaneous and independent decision regarding whether to cooperate or not cooperate (defect) with each other. Depending on their decisions they receive a payoff. At the beginning of the trial the participant sees the payoff matrix displayed on the screen. The columns of the matrix represent the participant's choices and the rows correspond to the co-player's choices. Whether the cooperative or not cooperative choice appears in the left or right column was randomized across trials. In the payoff matrix, numbers in bold/light grey correspond to the participant/co-player payoffs. Once the participant makes his/her choice the selected column of the matrix turns yellow. At the end of the trial, the payoff matrix is shown with only one cell highlighted, indicating the outcome of the trial. rt, reaction time; s, seconds. (*b*) Emotional responses to each of the PD game outcomes. Error bars denote standard deviations

depressive disorder plus a score $\geqslant$16 in the BDI and at least 3 weeks of not taking antidepressant medication. Participants in the control group had no current or past history of depression or any other psychiatric disorder.

Participants in the depression and control groups were matched on the basis of gender, age, years of education, and estimated pre-morbid IQ according to the National Adult Reading Test (NART; Nelson & Wilson, 1991) (Table 1).

Clinical ratings

Prior to scanning participants were assessed for symptom severity. Participants completed the (BDI, Beck *et al.* 1961), the Hamilton Depression/Anxiety scale (HAMD/HAMA, Hamilton, 1959, 1960), the Montgomery–Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979), the Spielberger State Anxiety scale (Spielberger, 1983), the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), the Positive Affect Negative Affect Scale (PANAS; Watson *et al.* 1988) and the Snaith–Hamilton hedonia scale (Snaith *et al.* 1995). The HAMD/HAMA and the MADRS were undertaken by a rater (V.B.G.). Between the recruitment and scanning sessions, participants completed the Sociotropy Autonomy Scale (SAS, Beck *et al.* 1983), the Personal Style Inventory (PSI; Robins *et al.* 1994), the Childhood Trauma Questionnaire (CTQ; Bernstein *et al.* 2003) and the Inventory of Interpersonal Problems (IIP, Horowitz *et al.* 1993).

Prisoner's Dilemma

While in the scanner, participants played the PD game (Fig. 1*a*). Before scanning, participants were shown

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

Table 1. *Participants' details*

	Control	Depression	Significance
n	25	25	
Female/male	17/8	17/8	N.S.
Age, years	25.44 ± 5.02	25.48 ± 5.52	$p = 0.98$, N.S.
NART	123.76 ± 2.82	124.28 ± 2.05	$p = 0.46$, N.S.
Education, years	16.52 ± 3.02	17.26 ± 2.93	$p = 0.38$, N.S.
BDI	0.40 ± 0.76	28.80 ± 9.06	$p < 0.001$
HAMD	0.16 ± 0.47	12.44 ± 4.23	$p < 0.001$
MADRS	0.48 ± 0.82	20.80 ± 6.97	$p < 0.001$
HAMA	0.44 ± 0.71	9.28 ± 4.17	$p < 0.001$
Spielberger State Anxiety	25.60 ± 3.79	48.48 ± 10.62	$p < 0.001$
RSES	25.40 ± 3.48	9.20 ± 3.82	$p < 0.001$
PANAS positive affect	38.96 ± 4.29	18.24 ± 4.78	$p < 0.001$
PANAS negative affect	11.92 ± 2.40	25.64 ± 6.43	$p < 0.001$
Snaith–Hamilton	4.12 ± 3.40	20.12 ± 4.53	$p < 0.001$
SAS sociotropy	57.92 ± 11.79	80.56 ± 17.76	$p < 0.001$
SAS autonomy	66.88 ± 13.49	67.16 ± 14.54	$p = 0.9$, N.S.
PSI sociotropy	82.32 ± 15.21	104.84 ± 15.36	$p < 0.001$
PSI autonomy	73.56 ± 16.12	94.56 ± 10.95	$p < 0.001$
CTQ	5.68 ± 0.96	9.67 ± 2.75	$p < 0.001$
IIP	54.84 ± 28.16	110.36 ± 27.31	$p < 0.001$

NART, National Adult Reading Test; BDI, Beck Depression Inventory; HAMD/HAMA, Hamilton Depression/Anxiety scale; MADRS, Montgomery–Asberg Depression Rating Scale; RSES, Rosenberg Self-Esteem Scale; PANAS, Positive Affect Negative Affect Scale; SAS, Sociotropy Autonomy Scale; PSI, Personal Style Inventory; CTQ, Childhood Trauma Questionnaire; IIP, Inventory of Interpersonal Problems; N.S., no significant difference between groups.

Values are mean ± S.D.

p values of the independent samples *t* test are provided.

how to play the PD (Supplementary material). Participants were told that they would be playing a game with a co-player who was outside the scanner room. It was explained that on each trial, both players would have to make simultaneous and independent decisions regarding whether to cooperate or not cooperate with each other. Depending on their decisions they would both receive earnings on each round. If they both cooperated they would both earn £2; if one cooperated and the other did not they would earn £0 and £3, respectively; if neither cooperated they would both earn £1. Participants were told that at the end of the game they would be paid a percentage of the money they had accumulated during the game and that the other player would also be similarly paid. In reality, participants played the PD against a pre-programmed algorithm (Rilling *et al.* 2002; McClure *et al.* 2007). This deception was necessary in order to minimize differences among participants in the experience of the PD while ensuring ecological validity.

The PD algorithm (McClure *et al.* 2007) generates each response based on the participant's choices on the prior two rounds (Supplementary material). A higher frequency of volunteer cooperation in the prior two rounds elicited a higher probability for a cooperative response, while a higher frequency of participant defection in the previous two rounds elicited a higher probability of a defection response. Following McClure *et al.* (2007), the algorithm was designed so that the participant would also experience periodic defection or 'betrayal'. Specifically, the algorithm had a 50% chance of defecting after four consecutive mutual cooperation trials. This effect was introduced as previous PD studies have reported that participants otherwise engage in mutual cooperation during much of the game (Rilling *et al.* 2002). This pattern of play would prevent participants from experiencing cooperation-defection outcomes in an adequate number of trials for statistical analysis. Participants played two sessions of the PD in the scanner. Each session lasted ~11.5 min and had 38 trials. The inter-trial timing variation ('jitter') was determined using 'Optseq' (http://surfer.nmr.mgh.harvard.edu/optseq/).

After scanning, participants completed a questionnaire that assessed their perceptions and emotional reaction to each of the PD outcomes (Supplementary material). Specifically, participants rated on nine-point Likert scales their satisfaction with their earnings, as well as their feelings of cooperativeness, anger, betrayal and guilt.

After the experiment, participants were debriefed regarding the cover story. All participants believed the cover story. No participants reported being unhappy regarding the deception. In the scanner, before playing the PD, participants played another behavioural economic task, the Ultimatum Game (Gradin *et al.* 2015). Participants were paid according to their earnings in both games with an average of £17.

Behavioural and emotional analysis

Emotional ratings were analysed using a three-way ANOVA with factors emotion, PD outcome and group. An ANOVA with factors outcome and group was used to analyse the number of occurrences of each outcome type as well as transition probabilities. The Greenhouse-Geisser correction was used for non-sphericity.

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

Neuroimaging analysis

For blood oxygen level-dependent response imaging, T2* weighted gradient echo-planar images were obtained using a 3-T Siemens Magnetom Trio Tim MRI scanner with a 12-channel head coil (see the Supplementary Material for further details on data acquisition and preprocessing). SPM8 (http://www.fil.ion.ucl.ac.uk/spm) was used for analyses.

For the first level analysis, an event related design was used which modelled neural activation at the decision (when the participant selected a column of the payoffs matrix) and outcome (when the feedback screen with the final payoff was presented) times. Specifically, six regressors were defined: decision C, decision D, and outcomes CC, CD, DC and DD. Six head motion realignment parameter estimates were included as covariates of no interest. Regressors of interest were convolved with the SPM8 haemodynamic response function without time or dispersion derivatives. Contrast images of interest were taken to second-level analyses and *within-* and *between*-group activations explored using one-sample and two-sample *t* tests.

For the depression group, we tested for significant correlations between neural activity and self-reported emotional rating scores. This correlational analysis was limited to the regions of interest where activation differed between groups. The dependent variable in this analysis was the mean value of the parameter estimates across voxels within the regions that showed between-group differences.

Unless otherwise stated, all analysis regions are reported as significant at a whole brain $p < 0.05$ cluster level. This was achieved by using parameters identified with Monte Carlo simulations: a simultaneous requirement for a voxel threshold of $p < 0.01$ and a minimum cluster size of 68 continuous voxels (Slotnick *et al.* 2003). All images are presented at this threshold.

Results

Two control and three depression data sets were excluded from analyses as these participants did not experience all four PD outcome types during each scanning session.

Clinical ratings

Depressed participants scored higher than controls on measures of depression (BDI, HAMD, MADRS), anxiety (HAMA, Spielberger State Anxiety scale), negative affect (PANAS), and anhedonia (Snaith–Hamilton scale). The depression group also scored significantly higher in sociotropy (SAS, PSI), autonomy (PSI), child abuse and neglect (CTQ) and interpersonal problems (IIP). The depressed group had lower mean scores than controls on measures of self-esteem (RSES) and positive affect (PANAS) (see Table 1 for all comparisons).

Emotional responses

After scanning, participants rated their perceptions of and emotional reactions to each of the PD outcomes. A three-way ANOVA identified significant main effects for emotion ($F_{2.18,93.77} = 85.99$, $p < 0.001$], outcome ($F_{2.5,107.50} = 5.186$, $p = 0.004$), a significant emotion×group interaction ($F_{2.18,93.77} = 6.83$, $p = 0.001$) and a significant emotion×outcome interaction ($F_{6.30,270.91} = 65.18$, $p < 0.001$). Follow-up analyses included independent ANOVAs for each emotion category with factors outcome and group. Each outcome type was associated with specific emotional reactions (Fig. 1*b*, Supplementary Table S1), consistent with previous work (Rilling *et al.* 2007). Specifically, CC outcomes were associated with satisfaction with earnings and feelings of cooperativeness; CD outcomes with feelings of anger and betrayal; DC outcomes with guilt; DD outcomes with intermediate levels of all emotions.

For satisfaction with earnings there was a significant effect of group, with depressed participants reporting less satisfaction than controls ($F_{1,43} = 10.67$, $p = 0.002$). Regarding feelings of cooperativeness and anger, there was no significant effect of group or a significant interaction with outcome type. For betrayal there was a significant effect of group ($F_{1,43} = 5.47$, $p = 0.024$), with depressed participants reporting higher levels of betrayal than controls. There was also a significant group×outcome-type interaction ($F_{2,3,99.52} = 3.46$, $p = 0.029$). Exploration of this interaction indicated that depressed participants reported significantly more betrayal than controls on DD outcomes ($p = 0.004$); and there were no significant between-group differences on any other outcomes. Finally, for guilt there was a significant effect of group ($F_{1,43} = 5.54$, $p = 0.023$), with depressed participants reporting higher levels than controls. There was also a non-significant interaction ($F_{2.44,104.87} = 2.45$, $p = 0.08$), which might be considered a trend. Decomposition of this interaction indicated that depressed participants reported higher levels of guilt than controls on DC outcomes ($p = 0.022$), not differentiating on all other outcomes. In summary, depressed participants reported less positive and more negative feelings in response to the PD game than controls.

Behavioural analyses

A mixed ANOVA with factors outcome and group was used to analyse the number of occurrences of each

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

outcome type. There was a significant effect of outcome ($F_{1.62,69.64} = 18.27$, $p < 0.001$), with CC and DD outcomes occurring more frequently than CD and DC outcomes. There were no significant group or interaction effects. We also analysed transition probabilities (i.e. the probability of cooperating following a specific outcome in the previous trial). This analysis identified a significant effect of outcome ($F_{2.58,110} = 50.18$, $p < 0.001$), with participants being more likely to cooperate after CC outcomes, followed by DC, CD, and DD outcomes, respectively. There was no significant group or interaction effect (Supplementary Table S1). Controls and depressed participants did not differ on earnings during the game. We examined reaction times for cooperation and defection following co-player cooperation or defection and having group as a factor. This analysis yielded no significant main effect for group or significant interactions with the group factor.

Neuroimaging analyses

To detect brain regions involved in reward processing during the PD, we analysed the contrast of reciprocated *v.* unreciprocated cooperation (CC > CD) (Rilling *et al.* 2004). For this contrast (Supplementary Table S2), across all participants (Fig. 2*a*) and also in the control group alone (Fig. 2*b*), we found activations extending through the nucleus accumbens and dorsal caudate, consistent with previous studies (Rilling *et al.* 2002, 2004). At the same level of significance, no activation was observed in the striatum in the depression group (Fig. 2*c*), nor were there significant between-group differences in this region. Similar results were obtained when considering the contrast [(CC + DC) > (CD + DD)] (i.e. every time a co-player cooperates *v.* every time a co-player does not cooperate) in order to examine responses to rewarding feedback *v.* unrewarding feedback during the task (Supplementary Fig. S1).

We also examined brain activity in response to unreciprocated *v.* reciprocated cooperation (CD > CC). Across all participants this analysis yielded activity in the bilateral DLPFC and the left anterior insula. No significant between-group differences in activation were observed in *a priori* regions of interest (Supplementary Fig. S2 and Table S3). Next, we examined activation associated with events in which the co-player cooperated while the participant defected; mutually cooperative trials served as a baseline (contrast: DC > CC). Across all participants this contrast showed significant activation in the bilateral DLPFC and bilateral insula (Supplementary Fig. S3 and Table S4), with no significant between-group differences observed in any regions of interest.

As both CD and DC outcomes activated a network comprising the DLPFC and insula, we pooled these events in a single contrast [(CD + DC) > (CC + DD)] (Rilling *et al.* 2002). That is, we compared the outcomes in which one of the players did not cooperate *v.* the outcomes in which both players either cooperated or defected. As noted in prior research (Rilling *et al.* 2002), CD and DC outcomes are typically aversive to at least one of the players and so are unlikely to be repeated, while CC and DD outcomes are more likely to repeat in a stable manner. Across all participants, this contrast elicited significant activations in the bilateral DLPFC, bilateral anterior insula and dorsomedial prefrontal cortex (Table 2, Fig. 3*a*). Controls also showed activations across these regions (Fig. 3*b*). Importantly, while depressed participants did show activation in the same network (Fig. 3*c*), they showed significantly diminished activation in the left DLFPC (Fig. 3*d*). This between-group difference in the left DLPFC was driven by the combination of reduced responses to CD and DC outcomes in the depressed group (Fig. 3*e*). Given the evidence supporting a role for the DLPFC in regulating emotions (Rive *et al.* 2013) we investigated whether activity in the left DLPFC correlated with self-reported emotional ratings in response to CD (anger and betrayal) and DC (guilt) outcomes. This analysis showed that in the depression group, diminished activation in the left DLPFC correlated with increasing self-reported ratings of guilt in response to DC outcomes ($r_{22} = -0.42$, $p = 0.05$; Fig. 3*f*). No significant correlations were found for anger and betrayal ratings.

Discussion

This study investigated behavioural, emotional, and neural responses during the PD game in adults with unmedicated depression.

At a neural level, it was found that during imbalanced (CD and DC) *v.* draw outcomes (CC and DD), depressed volunteers showed diminished activation in the left DLPFC compared to controls. The DLPFC has been implicated in processes of reasoning and higher cognition such as working memory, cognitive control (Miller & Cohen, 2001; D'Esposito & Postle, 2015), and also in the regulation of emotions (Okon-Singer *et al.* 2015). Of relevance here, the DLPFC has been found to activate in response to unfair relative to fair offers during the Ultimatum Game (Sanfey *et al.* 2003). This preferential DLPFC activation was interpreted as relating to the higher cognitive demands imposed by the unfair *v.* fair offers (Sanfey *et al.* 2003). Similarly, the DLPFC activation observed in our study was specifically associated with CD and DC outcomes. While CC and DD outcomes represent

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

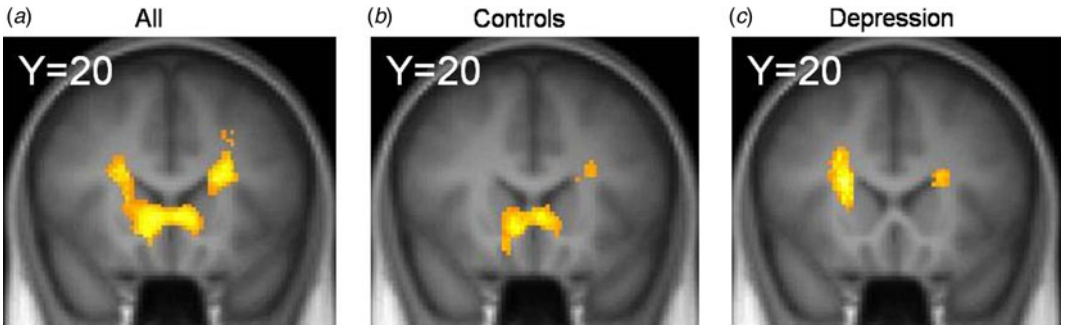

Fig. 2. Neural responses to reciprocated *v.* unreciprocated cooperation. Neural responses to reciprocated *v.* unreciprocated cooperation (CC > CD) across all participants (*a*), in controls (*b*) and in depression (*c*).

a draw, are stable, and tend to be repeated (Rilling *et al.* 2002), CD and DC outcomes are associated with negative emotions (anger and betrayal in one case, guilt in the other) and are more likely to lead to altered behaviour. If after a block of mutual cooperation the participant finds themself with a CD outcome, they may be more likely to choose to defect. Analogously, if after consecutive DD trials a DC outcome occurs, the participant will have to decide whether to follow the co-player signal and move to cooperation. Thus, the DLPFC activation found during these events may relate to higher cognitive demands placed by these outcomes in terms of emotion regulation and decision making.

Within this framework, diminished DLPFC activation during CD and DC outcomes in depressed volunteers suggests fewer cognitive resources for dealing with these events in terms of emotion regulation and decision making. Abnormal functioning of the DLPFC in depression is consistent with previous findings. Depression has been associated with reduced grey matter volume (Li *et al.* 2010) and abnormally low levels of resting state activity (Galynker *et al.* 1998; Mayberg *et al.* 1999) in the DLPFC. It has also been reported that damage to the DLPFC confers vulnerability to depression (Koenigs *et al.* 2008).

It has been hypothesized, that abnormal functioning of the DLPFC in depression may be associated with dysfunction in top-down regulation of emotion (Gotlib & Hamilton, 2008; Disner *et al.* 2011). Research using several paradigms lends support to this perspective. For example, one study showed decreased DLPFC activation in depression while participants had to ignore fear stimuli, as well as on post-error trials, suggesting impaired top-down control over affective interference and an impairment in making post-error cognitive adjustments (Fales *et al.* 2008). A second study found decreased DLPFC activation during reversal learning in depression (Remijnse *et al.* 2009). In a third study, participants with a history of depression failed to activate the DLPFC when they heard critical remarks from their own mothers (Hooley *et al.* 2005).

Findings do not point uniformly, however, to a consistent association between depression and attenuated DLPFC activity; indeed, several studies have yielded evidence of DLPFC hyperactivity in depression (Strigo *et al.* 2008; Frodl *et al.* 2009; Etkin & Schatzberg, 2011). According to a recent review on emotion regulation (Rive *et al.* 2013), whether the DLPFC overactivates or underactivates in depression depends on whether the emotion regulation process occurs in an automatic or voluntary manner. Studies using tasks that engage *automatic* emotion regulation (Frodl *et al.* 2009; Etkin & Schatzberg, 2011) have reported hyperactivity of the DLPFC in depression, possibly related to the need for additional resources in order to override strong bottom-up emotional influences. In contrast, studies using tasks that demand *voluntary* emotion regulation (Fales *et al.* 2008; Remijnse *et al.* 2009), reported decreased DLPFC activity in depression, suggesting a failure in recruitment of cognitive resources for cognitive control and regulation.

Rive *et al.* (2013) have proposed that during early automatic stages of emotion regulation, depressed subjects may be capable of regulating emotions, but only with the recruitment of additional lateral prefrontal regions. However, during explicit voluntary control, when the emotional experience is already ongoing, this strategy of additional recruitment may fail, as reflected by abnormally reduced activity in lateral prefrontal cortices. Studies of voluntary emotion regulation have used tasks that involve learning from feedback or reappraisal (Rive *et al.* 2013). In our study, the PD implies learning from feedback which may be consistent with reports of diminished DLPFC activation in depression feedback studies (Fales *et al.* 2008; Remijnse *et al.* 2009).

Cognitive theories of depression (Beck, 1979) propose that a core feature of the illness is a bias towards negativity in the processing of information, with

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

Table 2. *Within-group and between-group brain activations during the outcomes in which one player cooperated while the other did not v. the times in which both cooperated or defected (contrast [(CD + DC) > (CC + DD)])*

	BA	Cluster size	x	y	z	*T*
Activation for the contrast [(CD + DC) > (CC + DD)]						
All subjects						
L frontal lobe, middle frontal gyrus	9	2615	−40	10	34	4.73
L anterior insula		–[a]	−34	24	−4	6.65
R frontal lobe, precentral gyrus	9	2813	40	6	32	5.45
R anterior insula		–[a]	32	24	−2	5.12
Frontal lobe, medial frontal gyrus	8	1589	−8	28	50	3.97
Superior midbrain, thalamus		2524	−4	−18	−4	5.22
L occipital lobe, superior occipital gyrus	19	1026	−40	−80	22	3.79
R occipital lobe, superior occipital gyrus	19	3045	34	−76	24	3.85
Control group						
L frontal lobe, precentral gyrus	9	2672	−40	14	40	6.41
L anterior insula		–[a]	−36	24	−6	5.09
R frontal lobe, middle frontal gyrus	9	1708	42	8	44	5.14
R anterior insula		–[a]	34	20	−6	4.69
Frontal lobe, medial frontal gyrus	9	844	2	48	34	4.02
Thalamus		2782	6	−24	2	7.24
Parietal lobe, precuneus	7	142	0	−70	46	3.69
Parietal lobe, precuneus	31	1623	−12	−72	20	3.62
R parietal lobe, superior parietal lobule	7	–[a]	30	−64	56	3.29
Depression group						
L frontal lobe, inferior frontal gyrus	9	111	−42	4	32	3.63
R frontal lobe, inferior frontal gyrus	45, 9	357	50	24	24	3.79
L anterior insula		526	−38	18	−12	5.09
R anterior insula		110	32	28	0	3.24
Superior midbrain		228	6	−14	−8	4.94
R temporal lobe, fusiform gyrus	37	221	46	−54	−12	3.60
Control > Depression						
L frontal lobe, middle frontal gyrus	9	136	−36	24	32	3.20
Posterior thalamus		322	0	−28	6	4.04
Cerebellum		256	14	−54	−6	3.06
Depression > Control						
No significant activations						

L/R, Left/right; BA, Brodmann area.

Coordinates (x, y, z) reported in MNI space.

[a] Indicates that the peak belongs to the same cluster as the peak above.

All results significant at $p < 0.05$ cluster extent corrected across the whole brain.

depressed individuals selectively attending to and encoding negative events while filtering out positive information. This bias may decrease the experience of positive emotions while enhancing the feeling of negative emotions (Disner *et al.* 2011). Consistent with this, in our study depressed volunteers reported decreased satisfaction with earnings, as well as increased feelings of betrayal and guilt in response to the PD game. Our finding of heightened negative emotions in depressed participants is consistent with two previous PD studies. In one study, it was found that depressed participants reported feelings of self-devaluation, sadness and helplessness regarding exchanges during a modified version of the PD (Hokanson *et al.* 1980). In a second study (McClure *et al.* 2007), it was found that adolescent girls with anxiety and/or depression reported higher levels of anger towards the co-player. It is possible that the reduced DLPFC activation observed in the depression group underlies abnormalities in emotion regulation leading to the observed enhanced negative feelings. Consistent with this, reduced left DLPFC in depression correlated with increased feelings of guilt in response to DC outcomes.

It was not observed that depressed participants differed from controls in reward related activation in the striatum in response to mutual cooperation. While

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

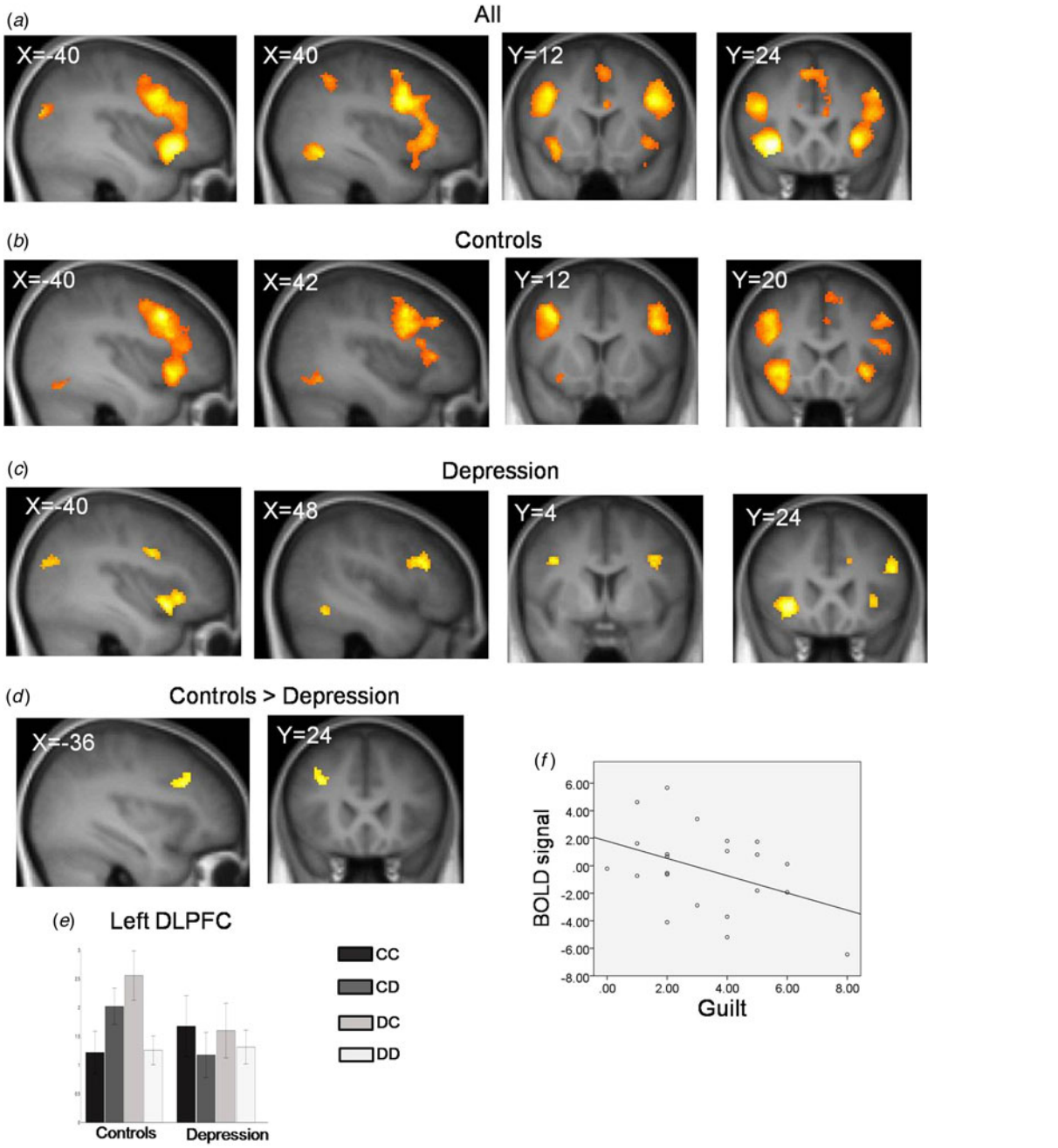

Fig. 3. Neural responses during events in which one player cooperated while the other did not *v.* events where both players cooperated or defected (contrast [(CD + DC) > (CC + DD)]). Neural responses across all participants (*a*), in controls (*b*) and in depression (*c*). Controls exhibited stronger responses in the left dorsolateral prefrontal cortex (DLPFC) than depressed participants (*d*). (*e*) Mean value of parameter estimates across voxels within a sphere of diameter 10 mm centred at peak coordinates (MNI: −36, 24, 32) of the left DLPFC. Error bars denote standard error of the mean. (*f*) Correlation within the depression group. *X* axis: self-reported feelings of guilt in response to DC outcomes during the PD game; *Y* axis: mean value of parameter estimates for the contrast [(CD + DC) > (CC + DD)] across voxels in the left DLPFC region where depressed participants differed from controls.

depressed volunteers had a weaker striatal response to mutual cooperation than controls, the between-group difference in this region did not pass our significance threshold. Of note, a previous study using the same participants as in the current study (Gradin *et al.* 2015), showed significantly diminished striatal activation in response to increasing fairness of offers during the Ultimatum Game in depression. Larger studies

https://doi.org/10.1017/S0033291715002834 Published online by Cambridge University Press

should investigate reward-linked brain activation in depression using the PD and other interactive paradigms. Similarly, depressed participants did not differ from controls in emotion/salience detection regions such as the insula. As above, further work needs to address the function of these regions in depression in the context of social interaction paradigms.

Of note, while depressed participants differed from controls in emotional and neural responses to the PD, the two groups did not differ in behaviour. Two previous PD studies have examined the behaviour of depressed populations. One study (Hokanson *et al.* 1980) used a modified version of the PD in which each player's relative power was manipulated. Results showed that when depressed individuals were in a controlling role, the pattern of play in the PD was relatively exploitive and non-cooperative. In contrast, another study using the PD (McClure *et al.* 2007) found that adolescents with anxiety/depression were more likely than controls to cooperate following co-player cooperation, suggesting a stronger need for maintenance of positive social interactions. Similarly to what is observed using the PD, studies using the Ultimatum Game have shown inconsistent results in depression reporting either increased, decreased or unchanged rejection rates to unfair offers (Harle *et al.* 2010; Destoop *et al.* 2012; Scheele *et al.* 2013; Gradin *et al.* 2015). As has been noted (Gradin *et al.* 2015; Pulcu & Elliott, 2015; Wang *et al.* 2015), these studies indicate that is not simple to predict depressed behaviour in the context of economic social exchange paradigms, and that further work is needed in order to investigate whether specific depression subtypes can be characterized by more consistent patterns of behaviour.

A possible limitation of the study relates to the use of a university sample which may limit generalizability of the results. This recruitment method was applied in order to facilitate recruitment of unmedicated depressed participants, avoiding a potential medication confound.

In summary, this study investigated patterns of emotional, behavioural, and neural responses in unmedicated depressed and control participants during social exchanges in the PD. In comparison to controls, the depressed group reported decreased levels of satisfaction with earnings and increased levels of betrayal and guilt feelings. Depressed participants also showed diminished DLPFC activation during exchanges in which one player cooperated and the other defected *v.* the events in which both players cooperated or defected. This abnormality in the DLPFC of depressed individuals may contribute to impairments in cognitive control and top down regulation of emotion during social situations that involve unreciprocated cooperation.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291715002834.

Acknowledgements

This study was funded by the University of Dundee, the Scottish Mental Health Research Network, Dr Kathleen White's Clinical Neuroscience Research Endowment Fund and a posdoctoral fellowship from ANII (National Agency for Investigation and Innovation, Uruguay) to V.B.G. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. We would like to thank Christine Matthews, Mairi Stirling, Craig Adams and Fiona Grant for their support with obtaining ethics approval, the recruitment and data management. We also thank the staff from the Advanced Intervention Service (AIS) for valuable discussions regarding the study design. Finally, we thank all the participants in this study

J.D.S. has received research funding via an honorarium associated with a lecture from Wyeth and an unrestricted educational grant from Schering-Plough.

Declaration of Interest

None.

References

APA (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. Washington, DC: American Psychiatric Association.

Axelrod R, Hamilton WD (1981). The evolution of cooperation. *Science* **211**, 1390–1396.

Beck AT (1979). *Cognitive Theory of Depression*. J. Wiley and Sons: New York.

Beck AT, Epstein N, Harrison RP, Emery G (1983). *Development of the Sociotropy-Autonomy Scale: a Measure of Personality Factors in Psychopathology*. Unpublished manuscript, University of Pennsylvania: Piladelphia.

Beck AT, Ward HC, Mendelson M (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.

Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, Stokes J, Handelsman L, Medrano M, Desmond D, Zule W (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse and Neglect* **27**, 169–190.

Billings AG, Cronkite RC, Moos RH (1983). Social-environmental factors in unipolar depression: comparisons of depressed patients and nondepressed controls. *Journal of Abnormal Psychology* **92**, 119–133.

Crockett MJ, Apergis-Schoute A, Herrmann B, Lieberman MD, Muller U, Robbins TW, Clark L (2013). Serotonin modulates striatal responses to fairness and retaliation in humans. *Journal of Neuroscience* **33**, 3505–3513.

D'Esposito M, Postle BR (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology* **66**, 115–142.

Destoop M, Schrijvers D, De Grave C, Sabbe B, De Bruijn ER (2012). Better to give than to take? Interactive social decision-making in severe major depressive disorder. *Journal of Affective Disorders* **137**, 98–105.

Disner SG, Beevers CG, Haigh EA, Beck AT (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience* **12**, 467–477.

Emonds G, Declerck CH, Boone C, Vandervliet EJ, Parizel PM (2012). The cognitive demands on cooperation in social dilemmas: an fMRI study. *Social Neuroscience* **7**, 494–509.

Eshel N, Roiser JP (2010). Reward and punishment processing in depression. *Biological Psychiatry* **68**, 118–124.

Etkin A, Schatzberg AF (2011). Common abnormalities and disorder-specific compensation during implicit regulation of emotional processing in generalized anxiety and major depressive disorders. *American Journal of Psychiatry* **168**, 968–978.

Fales CL, Barch DM, Rundle MM, Mintun MA, Snyder AZ, Cohen JD, Mathews J, Sheline YI (2008). Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression. *Biological Psychiatry* **63**, 377–384.

Fredman L, Weissman MM, Leaf PJ, Bruce ML (1988). Social functioning in community residents with depression and other psychiatric disorders: results of the New Haven Epidemiologic Catchment Area Study. *Journal of Affective Disorders* **15**, 103–112.

Frodl T, Scheuerecker J, Albrecht J, Kleemann AM, Muller-Schunk S, Koutsouleris N, Moller HJ, Bruckmann H, Wiesmann M, Meisenzahl E (2009). Neuronal correlates of emotional processing in patients with major depression. *World Journal of Biological Psychiatry* **10**, 202–208.

Galynker II, Cai J, Ongseng F, Finestone H, Dutta E, Serseni D (1998). Hypofrontality and negative symptoms in major depressive disorder. *Journal of Nuclear Medicine* **39**, 608–612.

Glimcher PW, Rustichini A (2004). Neuroeconomics: the consilience of brain and decision. *Science* **306**, 447–452.

Gotlib IH, Hamilton JP (2008). Neuroimaging and Depression. *Current Directions in Psychological Science* **17**, 159–163.

Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, Reid I, Hall J, Steele JD (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* **134**, 1751–1764.

Gradin VB, Perez A, MacFarlane JA, Cavin I, Waiter G, Engelmann J, Dritschel B, Pomi A, Matthews K, Steele JD (2015). Abnormal brain responses to social fairness in depression: an fMRI study using the Ultimatum Game. *Psychological Medicine* **45**, 1241–1251.

Hamilton M (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology* **32**, 50–55.

Hamilton M (1960). Rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* **23**, 56–62.

Harle KM, Allen JJ, Sanfey AG (2010). The impact of depression on social economic decision making. *Journal of Abnormal Psychology* **119**, 440–446.

Hasler G (2011). Can the neuroeconomics revolution revolutionize psychiatry? *Neuroscience and Biobehavioral Reviews* **36**, 64–78.

Hirschfeld RM, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, Moller HJ, Healy D, Baldwin D, Humble M, Versiani M, Montenegro R, Bourgeois M (2000). Social functioning in depression: a review. *Journal of Clinical Psychiatry* **61**, 268–275.

Hokanson JE, Sacco WP, Blumberg SR, Landrum GC (1980). Interpersonal behavior of depressive individuals in a mixed-motive game. *Journal of Abnormal Psychology* **89**, 320–332.

Hooley JM, Gruber SA, Scott LA, Hiller JB, Yurgelun-Todd DA (2005). Activation in dorsolateral prefrontal cortex in response to maternal criticism and praise in recovered depressed and healthy control participants. *Biological Psychiatry* **57**, 809–812.

Horowitz LM, Rosenberg SE, Bartholomew K (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. *Journal of Consulting and Clinical Psychology* **61**, 549–560.

King-Casas B, Chiu PH (2012). Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biological Psychiatry* **72**, 119–125.

King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR (2008). The rupture and repair of cooperation in borderline personality disorder. *Science* **321**, 806–810.

Koenigs M, Huey ED, Calamia M, Raymont V, Tranel D, Grafman J (2008). Distinct regions of prefrontal cortex mediate resistance and vulnerability to depression. *Journal of Neuroscience* **28**, 12341–12348.

Li CT, Lin CP, Chou KH, Chen IY, Hsieh JC, Wu CL, Lin WC, Su TP (2010). Structural and cognitive deficits in remitting and non-remitting recurrent depression: a voxel-based morphometric study. *Neuroimage* **50**, 347–356.

Mayberg HS, Liotti M, Brannan SK, McGinnis S, Mahurin RK, Jerabek PA, Silva JA, Tekell JL, Martin CC, Lancaster JL, Fox PT (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *American Journal of Psychiatry* **156**, 675–682.

McClure EB, Parrish JM, Nelson EE, Easter J, Thorne JF, Rilling JK, Ernst M, Pine DS (2007). Responses to conflict and cooperation in adolescents with anxiety and mood disorders. *Journal of Abnormal Child Psychology* **35**, 567–577.

McClure-Tone EB, Nawa NE, Nelson EE, Detloff AM, Fromm SJ, Pine DS, Ernst M (2011). Preliminary findings: neural responses to feedback regarding betrayal and cooperation in adolescent anxiety disorders. *Developmental Neuropsychology* **36**, 453–472.

Miller EK, Cohen JD (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* **24**, 167–202.

Montgomery SA, Asberg M (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* **134**, 382–389.

Nelson HE, Wilson JR (1991). *The Revised National Adult Reading Test – Test Manual*. NFER-Wilson: Winsor.

Okon-Singer H, Hendler T, Pessoa L, Shackman AJ (2015). The neurobiology of emotion-cognition interactions: fundamental questions and strategies for future research. *Frontiers in Human Neuroscience* **9**, 58.

Papakostas GI, Petersen T, Mahal Y, Mischoulon D, Nierenberg AA, Fava M (2004). Quality of life assessments in major depressive disorder: a review of the literature. *General Hospital Psychiatry* **26**, 13–17.

Pulcu E, Elliott R (2015). Neural origins of psychosocial functioning impairments in major depression. *Lancet Psychiatry* **2**, 835–843.

Remijnse PL, Nielen MM, van Balkom AJ, Hendriks GJ, Hoogendijk WJ, Uylings HB, Veltman DJ (2009). Differential frontal-striatal and paralimbic activity during reversal learning in major depressive disorder and obsessive-compulsive disorder. *Psychological Medicine* **39**, 1503–1518.

Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C (2002). A neural basis for social cooperation. *Neuron* **35**, 395–405.

Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA, Lilienfeld SO (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry* **61**, 1260–1271.

Rilling JK, Goldsmith DR, Glenn AL, Jairam MR, Elfenbein HA, Dagenais JE, Murdock CD, Pagnoni G (2008). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* **46**, 1256–1266.

Rilling JK, Sanfey AG, Aronson JA, Nystrom LE, Cohen JD (2004). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* **15**, 2539–2543.

Rive MM, van Rooijen G, Veltman DJ, Phillips ML, Schene AH, Ruhe HG (2013). Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. *Neuroscience and Biobehavioral Reviews* **37**, 2529–2553.

Robins CJ, Ladd J, Welkowitz J, Blaney PH, Diaz R, Kutcher G (1994). The personal style inventory: preliminary validation studies of new measures of sociotropy and autonomy. *Journal of Psychopathology and Behavioral Assessment* **16**, 277–300.

Rosenberg M (1965). *Society and the Adolescent Self-Image*. Princeton University Press: Princeton, NJ.

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* **300**, 1755–1758.

Scheele D, Mihov Y, Schwederski O, Maier W, Hurlemann R (2013). A negative emotional and economic judgment bias in major depression. *European Archives of Psychiatry and Clinical Neuroscience* **263**, 675–683.

Segrin C (2000). Social skills deficits associated with depression. *Clinical Psychology Review* **20**, 379–403.

Slotnick SD, Moo LR, Segal JB, Hart J Jr. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research* **17**, 75–82.

Snaith RP, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *British Journal of Psychiatry* **167**, 99–103.

Spielberger CD (1983). *State-Trait Anxiety Inventory for Adults, STAI Form Y-1*. Mind Garden: Palo Alto.

Strigo IA, Simmons AN, Matthews SC, Craig AD, Paulus MP (2008). Association of major depressive disorder with altered functional brain response during anticipation and processing of heat pain. *Archives of General Psychiatry* **65**, 1275–1284.

Tabibnia G, Satpute AB, Lieberman MD (2008). The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science* **19**, 339–347.

Wang Y, Yang LQ, Li S, Zhou Y (2015). Game theory paradigm: a new tool for investigating social dysfunction in major depressive disorders. *Frontiers in Psychiatry* **6**, 128.

Watson D, Clark LA, Tellegen A (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology* **54**, 1063–1070.

Yoshida W, Dolan RJ, Friston KJ (2008). Game theory of mind. *PLOS Computational Biology* **4**, e1000254.

Yoshida W, Seymour B, Friston KJ, Dolan RJ (2010). Neural mechanisms of belief inference during cooperative games. *Journal of Neuroscience* **30**, 10744–10751.

Zhang WN, Chang SH, Guo LY, Zhang KL, Wang J (2013). The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *Journal of Affective Disorders* **151**, 531–539.