Subgenual activation and the finger of blame: individual differences and depression vulnerability

Author for correspondence:
Roland Zahn,
E-mail: roland.zahn@kcl.ac.uk

Karen E Lythe[1], Jennifer A Gethin[1], Clifford I Workman[1,2], Matthew A. Lambon Ralph[1,3], John F.W. Deakin[2], Jorge Moll[4] and Roland Zahn[1,4,5,6]

[1]The University of Manchester & Manchester Academic Health Sciences Centre, School of Psychological Sciences, Neuroscience and Aphasia Research Unit, Manchester, M13 9PL, UK; [2]The University of Manchester & Manchester Academic Health Sciences Centre, Institute of Brain, Behaviour and Mental Health, Neuroscience & Psychiatry Unit, Manchester, M13 9PL, UK; [3]MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road, Cambridge, CB2 7EF, UK; [4]Cognitive and Behavioral Neuroscience Unit, D'Or Institute for Research and Education (IDOR), 22280-080 - Rio de Janeiro, RJ, Brazil; [5]Institute of Psychiatry, Psychology & Neuroscience, Department of Psychological Medicine, Centre for Affective Disorders, King's College London, London, SE5 8AZ, UK and [6]National Service for Affective Disorders, South London and Maudsley NHS Foundation Trust, London, SE5 8AZ, UK

Abstract

Background. Subgenual cingulate cortex (SCC) responses to self-blaming emotion-evoking stimuli were previously found in individuals prone to self-blame with and without a history of major depressive disorder (MDD). This suggested SCC activation reflects self-blaming emotions such as guilt, which are central to models of MDD vulnerability.

Method. Here, we re-examined these hypotheses in an independent larger sample. A total of 109 medication-free participants (70 with remitted MDD and 39 healthy controls) underwent fMRI whilst judging self- and other-blaming emotion-evoking statements. They also completed validated questionnaires of proneness to self-blaming emotions including those related to internal (autonomy) and external (sociotropy) evaluation, which were subjected to factor analysis.

Results. An interaction between group (remitted MDD *v.* Control) and condition (self- *v.* other-blame) was observed in the right SCC (BA24). This was due to higher SCC signal for self-blame in remitted MDD and higher other-blame-selective activation in Control participants. Across the whole sample, extracted SCC activation cluster averages for self- *v.* other-blame were predicted by a regression model which included the reliable components derived from our factor analysis of measures of proneness to self-blaming emotions. Interestingly, this prediction was solely driven by autonomy/self-criticism, and adaptive guilt factors, with no effect of sociotropy/dependency.

Conclusions. Despite confirming the prediction of SCC activation in self-blame-prone individuals and those vulnerable to MDD, our results suggest that SCC activation reflects blame irrespective of where it is directed rather than selective for self. We speculate that self-critical individuals have more extended SCC representations for blame in the context of self-agency.

Introduction

Self-blame and associated feelings, such as guilt and self-worthlessness, play a key role in cognitive (Abramson, Seligman, & Teasdale, 1978; Beck, Rush, Shaw, & Emery, 1979), as well as psychodynamic models of major depressive disorder (MDD) vulnerability (Freud, 1917). This is supported by findings of persistent biases towards blaming oneself relative to others in MDD even on remission of symptoms with no overall increase in negative emotions when controlling for the direction of blame (Green, Moll, Deakin, Hulleman, & Zahn, 2013a; Zahn et al., 2015). Understanding the neurocognitive basis of self-blaming emotions as vulnerability factors for MDD is important for elucidating the link between psychosocial and biological factors predisposing to MDD.

As recently reviewed (Zahn, De Oliveira-Souza, & Moll, 2020), the most reproducible neural correlate of individual differences in proneness to self-blaming emotions such as guilt is a higher subgenual cingulate cortex (SCC) activation for guilt *v.* other-directed anger (Zahn, Moll, et al., 2009). This was true of guilt-prone individuals irrespective of whether they had a history of MDD or not (Green, Lambon Ralph, Moll, Deakin, & Zahn, 2012). SCC activation was also higher in remitted MDD patients when they made decisions to anonymously donate to charity, which could be driven by anticipated guilt (Pulcu et al., 2014). Consistent activation for guilt in the SCC across individuals was, however, found in

CAMBRIDGE
UNIVERSITY PRESS

CrossMark

some studies (Basile et al., 2011; Morey et al., 2012), but not others (Zahn, de Oliveira-Souza, Bramati, Garrido, & Moll, 2009), (Green et al., 2012; Zahn, Moll, et al., 2009). One reason for finding consistent group effects in the SCC could have been the modelling of guilt-intensity at the trial-by-trial level in one of these studies (Morey et al., 2012). This may have led to similar results as our approach of modelling individual differences in guilt-frequency to reveal SCC activation (Green et al., 2012; Zahn, Moll, et al., 2009). We have previously interpreted these results as indicating a selective role for the SCC in self-blaming emotions (Green et al., 2012; Zahn, de Oliveira-Souza, et al., 2009; Zahn, Moll, et al., 2009) and interpreted our finding of SCC activation only in guilt-prone individuals due to higher guilt-selective effect sizes in those individuals rather than a categorical difference between those who are prone to self-blame *v.* those who are not. The alternative explanation of a true lack of self-blame-selective activation in the SCC in a large proportion of people would challenge our previous interpretation by calling into question whether the SCC is associated with self-blame-selective representations. The aim of this study was to examine the latter possibility and to investigate whether individual differences in internal *v.* external-evaluation dependence of self-blaming emotions may explain individual differences in SCC activation.

The literature on self-blaming emotions has primarily focussed on the distinction between depressogenic forms of self-blaming emotions entailing the causal attribution to one's characterological faults (Janoff-Bulman, 1979) and thus hopelessness and helplessness (Abramson et al., 1978), e.g. overgeneral guilt (O'Connor, Berry, Weiss, & Gilbert, 2002), shame (Tangney, Wagner, & Gramzow, 1992), self-disgust/hate (O'Connor et al., 2002; Zahn et al., 2015; Zahn et al., 2015) *v.* adaptive forms entailing self-blame for a specific behaviour, such as differentiated guilt associated with reparative action (Tangney et al., 1992). Despite entailing related constructs, 'autonomy' and 'sociotropy' dimensions of MDD vulnerability have evolved in a largely separate literature and have been validated as being independent of general negative affect (Robins et al., 1994). Beck observed that differences in strivings for being accepted by others (sociotropy) or achievement and self-control (autonomy) render individuals vulnerable to developing depression in response to different types of life events (Clark, Steer, Beck, & Ross, 1995). Sociotropic individuals were thought to typically develop depression after social resource threats, whereas autonomic individuals were deemed more susceptible to threats to their independence (Clark et al., 1995) (i.e. the sense of self-agency). This hypothesis was based on the psychoanalytical literature (Balint, 1959) and Bowlby's concepts of subtyping depression vulnerability (Robins et al., 1994) on the basis of 'anxious attachment' *v.* 'compulsive self-reliance' (Bowlby, 1977) as ways of responding to early attachment threat or loss.

Bowlby's attachment theory discusses the survival benefits of balancing alternating 'attachment behaviour' (i.e. seeking protection by parents) with 'exploratory behaviour' (i.e. developing autonomy) in young offspring across social species (Bowlby, 1977). Developing autonomy from a secure attachment base (Bowlby, 1977) entails developing a healthy sense of agency that attributes blame for specific actions. In contrast, overgeneralised characterological attributions of the causal agency to oneself for negative events were postulated by the revised learned helplessness model to render individuals vulnerable to excessive self-blaming emotions, low self-worth and MDD (Abramson et al., 1978).

Here, we sought to stratify participants with and without a history of MDD according to their proneness to adaptive forms of self-blaming emotions, such as differentiated guilt linked to reparative actions (Tangney et al., 1992), and overgeneralised self-blaming emotions we predicted to be associated with internal evaluation (autonomy), such as overgeneralised guilt (O'Connor et al., 2002) and self-hate/disgust (Green et al., 2013a), as well as those we predicted to be more strongly linked to external evaluation (sociotropy), such as shame (Higgins, 1987). We hypothesised that (1) SCC activation in response to self-blaming-emotion-evoking stimuli is higher in individuals who are prone to self-blaming emotions, in particular, self-hate and guilt as these are thought to rely more on internalised moral norms (Higgins, 1987) and hence a stronger attribution of agency to oneself compared with shame, a feeling linked to external evaluation and uncontrollable factors (Higgins, 1987). This is based on previous evidence for the hypothesis that activation in ventromedial frontal subregions is associated with emotional stimuli that require representing social agency (Zahn, Moll, et al., 2009). We further hypothesised that (2) SCC activation is higher in individuals with remitted MDD who are known to have a largely heightened vulnerability to developing depressive episodes compared with control participants (Eaton et al., 2008). The latter hypothesis was based on the extensive literature on the importance of SCC activation in MDD (Ebert & Ebmeier, 1996; Price & Drevets, 2010; Ressler & Mayberg, 2007; Siegle, Carter, & Thase, 2006), particularly in familial forms (Drevets, Ongur, & Price, 1998) which are associated with guilt-proneness (Leckman et al., 1984), despite our previous failure to find differences between remitted MDD and control participants in a smaller independent sample (Green et al., 2012). To investigate these hypotheses, we used standard scales of proneness to self-blaming emotions (O'Connor et al., 2002; Tangney & Dearing, 2000) and their link with strivings for autonomy and sociotropy (Robins et al., 1994) rather than ratings of stimuli also used during the fMRI scan as in some of our previous studies (Green et al., 2012; Zahn, Moll, et al., 2009).

Method

Participants

Participants were recruited as part of the UK Medical Research Council-funded 'Development of Cognitive and Imaging Biomarkers Predicting Risk of Self-Blaming Bias and Recurrence in Major Depression' project (Lythe et al., 2015). In total, 109 (70 patients with remitted MDD and 39 control) participants were included in the fMRI data analysis for the current study (see online Supplementary Tables 1 & 2 for exclusion reasons) which was approved by the South Manchester National Health Service Research Ethics Committee. Questionnaire measures for individual variability analyses were incomplete for three MDD patients, leaving $n = 67$ MDD and $n = 39$ control participants for this part of our analysis. All participants gave written informed consent and received compensation for time and travel costs. This investigation of fMRI activation data at baseline has not previously been reported, but data were collected as part of a longitudinal study, examining whether self-blame-selective alterations in anterior temporal fMRI connectivity predict subsequent recurrence of depression (see Lythe et al., 2015).

We included people with a diagnosis of MDD in remission for at least 6 months according to the Structured Clinical Interview

for DSM-IV-TR (First, Spitzer, Gibbon, & Williams, 2002) [online Supplementary Table 4, with high inter-rater reliability as reported in (Lythe et al., 2015)] and a current Montgomery Asberg Depression Scale (Montgomery & Åsberg) score <10. Exclusion criteria were current Axis-I disorders including a history of alcohol or substance abuse and past comorbid Axis-I disorders that were the likely cause of depressive symptoms (online Supplementary Table 1 & 2). The control group had no current or past Axis-I diagnoses, and no first-degree history of MDD, bipolar disorder or schizophrenia. Both the MDD and Control groups were psychotropic medication-free, right-handed, native English speakers, with normal or corrected-to-normal vision.

Questionnaire measures

All employed questionnaires have previously been validated and found to show high internal consistency in relevant samples (O'Connor et al., 2002; O'Connor, Berry, Weiss, Bush, & Sampson, 1997; Robins et al., 1994; Tangney et al., 1992; Tangney & Dearing, 2000; Tangney, Stuewig, & Mashek, 2007) and this also pertains to the constructs of 'autonomy' and 'sociotropy' with a recent meta-analysis combining data from 90 studies and 30 372 participants (Yang & Girgus, 2019) using either the Personal Style Inventory employed here or the Sociotropy Autonomy Scale (Clark & Beck, 1991). We computerised these paper-based questionnaires with Excel Macros. We used the Interpersonal Guilt Questionnaire [IGQ-67,(O'Connor et al., 1997)], which includes four subscales: omnipotent responsibility guilt, which arises from exaggerated feelings of responsibility for the wellbeing and happiness of others; survivor guilt, where one feels bad for being better off than others; separation guilt, arising from the fear of harming another by pursuing one's own goals; and self-hate. The Test of Self-Conscious Affect [TOSCA, (Tangney & Dearing, 2000)] was used to measure shame, adaptive guilt, detachment/unconcern and externalisation of blame. Participants completed the Personal Style Inventory – Revised Edition [PSI-II, (Robins et al., 1994)] to obtain measures of sociotropy and autonomy. The sociotropy scale comprises three subscales: 'concern about what people think', 'dependency' and 'pleasing others'. The autonomy scale also consists of three subscales: 'self-criticism', 'need for control' and 'defensive separation'.

fMRI paradigm

As in our previous independent study (Green et al., 2012), participants saw sentences containing social concepts (e.g. 'stingy', 'impatient') describing actions counter to socio-moral values, in either negative or negated positive form. The agent was either the participant [self-agency condition ($n = 90$)] or their best friend [other-agency condition ($n = 90$)]. For example, (participant's name) does act stingily towards (best friend's name). Self- and other-agency conditions contained the same social concepts. Participants were required to report how unpleasant they would feel ('mildly' or 'very') via a button press within 5 s, followed by a jittered inter-trial interval with a mean duration of 4 s. A low-level visual perception baseline condition (null condition) asked the participant to observe rows of asterisks arranged in the same way as the verbal stimuli but required no response ($n = 90$) and was pseudo-randomly interspersed across three runs, the order of which was counterbalanced across participants (details of the fMRI task have previously been reported in Lythe et al. (2015), see Supplementary Methods).

After the scanning session, participants rated the degree of unpleasantness on a 7-point Likert scale (1 = not unpleasant, 7 = extremely unpleasant) associated with each stimulus. In addition, they were asked to 'choose the feeling that they would feel most strongly' from different self- and other-blaming emotions as previously reported (Zahn et al., 2015). Self-blaming and other-blaming emotion trials for the fMRI analysis were defined as those that were perceived as highly unpleasant (those rated post-scanning at *individual* median or above) in the respective self- and other-agency conditions. In addition, participants were asked to 'choose the feeling that they would feel most strongly' from different self-blaming and other-blaming emotions and the results of this 'Value-Related Moral Sentiment Task' have been previously reported (Zahn et al., 2015) to show selective associations of self-blaming emotions with the self-agency and other-blaming emotions with the other-agency conditions.

Image acquisition

An fMRI protocol optimised for the detection of ventral brain regions was used as described previously (Green et al., 2012). T2*-weighted echo-planar images (three runs of 405 volumes with five dummy scans, Repetition Time = 2000 ms) and T1-weighted, magnetisation-prepared, rapid-acquisition gradient-echo structural images were acquired on an MRI scanner (3 T Achieva, Philips, see Supplementary Methods).

Behavioural data analysis

Analyses were carried out in SPSS20 (http://www.ibm.com/software/analytics/spss/) at $\alpha = 0.05$, two-tailed. To reduce the questionnaire variables into uncorrelated factors we used principal components analysis (PCA) with VARIMAX rotation. For the number of factors considered, Eigenvalues >1, Scree plot and interpretability were taken into account. Item loadings with values greater than 0.58 were used to describe the components. Reliable factors were determined as those with at least three loadings above 0.80, or four or more loadings above 0.60 as recommended by Stevens (Stevens, 2009), who points out that those factors are most reliable which have many variable loadings.

Image analysis

Functional images were realigned, unwarped, coregistered to the participant's T1-weighted images, and normalised using the default resulting voxel size of $2 \times 2 \times 2$ mm to the SPM template using nonlinear transformation parameters derived during segmentation of the T1-weighted image, before a smoothing kernel of 6 mm full-width-at-half-maximum was applied (http://www.fil.ion.ucl.ac.uk/spm8/). We used SPM8 rather than SPM12 to keep our analysis comparable with our previous paper describing functional connectivity results in this sample (Lythe et al., 2015).

At the individual level, Blood-Oxygenation-Level-Dependent (BOLD) effects were modelled for self-agency and other-agency conditions and modelling high (median or above across trials for the individual) and low (below median across trials for individual) degrees of unpleasantness of the trials in each condition. Null events and movement parameters (i.e. six parameters describing movement by rotation and translation in three dimensions each) were also included as covariates for the three runs in addition to the established unwarping and realignment algorithms recommended for task-based fMRI in SPM (Andersson,

Hutton, Ashburner, Turner, & Friston, 2001). Root mean squares of the movement parameters did not differ between the two groups for both translation (MDD: 0.33 ± .18; CONTROL: 0.35 ± .18; $t(107) = -0.54$, $p = 0.59$) and rotation (MDD: 0.01 ± .00; CONTROL: 0.01 ± .00; $t(107) = .66$, $p = 0.51$).

We modelled the temporal and spatial derivatives of the haemodynamic response function. All analyses were inclusively masked with a grey matter mask as previously described (Green et al., 2012). At the second level, we used a factorial model with two factors: Group (MDD *v.* Control) and condition (Self- *v.* Other-Blame). *F*-contrasts for main effects of group, condition and their interaction were displayed at $p = 0.005$ (uncorrected voxel-level) and then corrected for family-wise-error (FWE) at the voxel-level at $p = 0.05$ over our *a priori* SCC ROI [MNI coordinates: −4, 23, −5; 6 mm sphere, as used for an independent previous sample (Green et al., 2012)] and the volume of the whole brain. This a priori ROI was previously derived from averaging SCC coordinates across studies linking this region with self-blame and although its peak is slightly left lateralised, the ROI is bilateral. To determine which conditions gave rise to the identified interaction effect and to correlate with behavioural measures, we used MarsBaR (Brett, Anton, Valabregue, & Poline, 2002) to extract the SCC cluster average regression coefficients in each condition (Self- and Other-Blame) *v.* the low-level baseline for each participant and examined these further in SPSS20. The exact anatomical labelling of the peak coordinate was determined by using the MNI to Talairach transform by Brett and identifying the anatomical label in axial, coronal and sagittal sections of the original Talairach atlas in its printed version (Talairach & Tournoux, 1988).

Results

Demographic and clinical data

There were no significant differences between Remitted MDD and Control groups in age, years of education and sex ($t < -1.33$, $p > 0.19$, online Supplementary Table 3). Scores on the Beck Depression Inventory [BDI; (Beck, Steer, & Carbin, 1988)] and Montgomery Asberg Depression Rating Scale [MADRS, (Montgomery & Åsberg, 1979)] were slightly elevated in the Remitted MDD group compared with the Control group, although mean scores for both groups fell below the threshold for mild depressive symptoms (online Supplementary Table 3). In addition, Global Assessment of Functioning scale [GAF; (First et al., 2002)] scores were slightly lower in the Remitted MDD group compared with the Control group, although mean scores for both groups suggested absent or minimal symptoms and good levels of social functioning (online Supplementary Table 3). Groups did not differ on the percentage of trials included in the self- and other-blaming emotion conditions, unpleasantness ratings, or response times during fMRI acquisition ($t < 1.57$, $p > 0.12$, online Supplementary Table 5).

Principal component analysis

Based on Eigenvalues >1, Screeplot and clinical interpretation, a four-factor solution resulted from the principal component analysis and explained 73.78% of the total variance (Table 1). The first factor represented sociotropy/dependency with high loadings on all sociotropy subscales and on omnipotent responsibility and separation guilt. The second factor captured an autonomy/self-

criticism factor, with high loadings on all the autonomy subscales, and on self-hate and survivor guilt. The third factor comprised detachment/unconcern and externalisation from the TOSCA. The fourth factor mainly captured guilt from the TOSCA questionnaire and was labelled as 'adaptive guilt', because of the operationalisation of guilt on the TOSCA as a non-depressogenic behavioural form of self-blame linked to reparative actions. Only the first two factors were considered as reliable based on the number and loading of the components within each of the factors (Stevens, 2009). For our further correlations with fMRI results, we therefore used the two reliable factors and the TOSCA guilt score which was deemed more reliable than the adaptive guilt factor score.

Remitted MDD patients exhibited higher sociotropy/dependency and autonomy/self-criticism factor scores compared with the Control group (Table 2). There were no between-group differences for factors 3 (detachment/externalisation) and 4 (adaptive guilt).

fMRI results

A significant interaction effect between group (Remitted MDD *v.* Control) and condition (self- *v.* other-blaming) was observed in the right SCC (Fig. 1, Table 3) and confirmed for the extracted cluster averages in this region [$F(1107) = 7.65$, $p = 0.007$], with no main effect of agency [$F(1107) = 2.15$, $p = 0.15$] or group [$F(1107) = .03$, $p = 0.86$]. This interaction effect was due to higher SCC signal for self-blame in the Remitted MDD group ($M = 1.05$, S.D. = 7.08) relative to other-blame ($M = 0.16$, S.D. = 6.61), resulting in a positive difference for self-blame *v.* other-blame ($M = 0.89$, S.D. = 5.74, $t = 1.29$, df = 69, $p = 0.20$) in the Remitted MDD group, whilst the Control group showed the reverse pattern of lower SCC signal for self-blame ($M = -1.05$, S.D. = 8.03) relative to other-blame ($M = 1.82$, S.D. = 7.18) resulting in a negative difference for self- *v.* other-blame ($M = -2.87$, S.D. = 8.38, $t = -2.14$, df = 38, $p = 0.04$). The interaction effect results from the significant differences between the groups on these self-blame *v.* other-blame differences ($t = -2.77$, df = 107, $p = 0.007$, mean difference = −3.76, standard error = 1.36). There were no significant main effects or interactions outside the SCC in our whole brain analysis.

Across the whole sample, SCC activity during self- *v.* other-blame was significantly predicted by a linear regression model which included the two reliable factors and the adaptive guilt measure from the TOSCA (Table 4). Interestingly, this prediction was solely driven by autonomy/self-criticism factor scores, and adaptive guilt, with no effect of sociotropy/dependency factor scores (Table 4). To avoid circular analyses, we did not primarily consider group here, because the extracted SCC activation cluster means were already biased by the SPM analysis to find the voxels showing a maximal group × condition interaction. Unsurprisingly, the effects for autonomy/self-criticism on SCC activation disappeared when covarying group ($\beta = 0.10$, $t = 0.82$, $p = 0.42$), because of higher scores on this factor in the MDD group as reported above, with the effects of adaptive guilt remaining ($\beta = 0.25$, $t = 2.7$, $p = 0.009$).

The SCC activation coefficients for self- *v.* other-blame in the MDD group did not correlate with the number of previous MDEs (Spearman's $\rho = -0.02$, $p = 0.89$), or measures of residual symptoms: BDI scores ($\rho = -0.12$, $p = 0.31$), GAF scores ($\rho = 0.15$, $p = 0.21$), or MADRS scores ($\rho = -0.29$, $p = 0.20$). There were also no correlations between the SCC coefficients for

Table 1. Rotated factor values showing loadings of each component

Scale Item	Factors			
	Factor 1	Factor 2	Factor 3	Factor 4
IGQ-67 – Omnipotence responsibility Guilt	**0.75**[a]	0.33	−0.17	0.20
IGQ-67 – Separation Guilt	**0.58**[a]	−0.24	0.01	0.26
IGQ-67 – Self-hate	0.52	**0.61**[a]	−0.07	0.17
IGQ-67 – Survivor Guilt	0.40	**0.60**[a]	−0.16	0.38
TOSCA Shame	0.48	0.52	0.08	0.48
TOSCA Guilt	0.07	0.08	−0.11	**0.93**[a]
TOSCA Detachment	−0.18	−0.24	**0.80**[a]	−0.20
TOSCA Externalisation	0.07	0.15	**0.90**[a]	0.06
PSI-II: Sociotropy – Concern What People Think	**0.85**[a]	0.27	−0.01	0.08
PSI-II: Sociotropy – Dependency	**0.83**[a]	0.19	0.03	−0.14
PSI-II: Sociotropy – Pleasing Others	**0.60**[a]	0.49	−0.12	0.15
PSI-II: Autonomy – Self-criticism	0.47	**0.67**[a]	0.13	0.23
PSI-II: Autonomy – Need for Control	0.19	**0.85**[a]	0.02	−0.02
PSI-II: Autonomy – Defensive Separation	−0.05	**0.90**[a]	−0.07	−0.01

IGQ, Interpersonal Guilt Questionnaire; PSI II, Personal Style Inventory Revised Edition; TOSCA, Test of Self-Conscious Affect.
[a]Factor loadings above threshold (>0.58). Principal Components Analysis using VARIMAX rotation. $n = 106$ ($n = 39$ control and $n = 67$ MDD) were included in this analysis. Principal components analyses are designed to derive uncorrelated factor components which explain the variance contained in the set of variables (Stevens, 2009).

Table 2. Group comparisons on factors derived from principal components analysis

Factor	MDD ($n = 67$)	Control ($n = 39$)	MDD *v.* Control
1 Sociotropy/dependency	0.19 ± 0.96	−0.32 ± 1.00	$t(104) = 2.60, p < 0.01$[a]
2 Autonomy/self-criticism	0.45 ± 0.83	−0.78 ± 0.77	$t(104) = 7.56, p < 0.0001$[a]
3 Detachment/externalisation	−0.12 ± 1.02	0.21 ± 0.93	$t(104) = −1.67, p = 0.10$
4 Adaptive guilt	0.01 ± 1.07	−0.01 ± 0.87	$t(104) = −1.71, p = 0.92$

Data for two MDD participants were missing for the Interpersonal Guilt Questionnaire (IGQ-67) and Test of Self-Conscious Affect (TOSCA), and data for three MDD participants were missing for the Personal Style Inventory (PSI-II).
[a]Significant at $p = 0.05$ threshold, 2-tailed. Means and standard deviations are reported ($M ±$ s.d.).

self- *v.* other-blame in the MDD group for rated unpleasantness or negative affectivity as measured on the Positive and Negative Affect scale (Supplementary results, (Watson, Clark, & Tellegen, 1988)).

Discussion

Our results confirm our predictions of higher self-blame-selective SCC activation in individuals with remitted MDD and those who are prone to self-blaming emotions. As expected (Green et al., 2012; Zahn, Moll, et al., 2009), SCC activation was not associated with unpleasantness of the stimuli or standard measures of negative affectivity (Watson et al., 1988) shown to be highly associated with 'neuroticism' (Clark, Watson, & Mineka, 1994), and can thus not be attributed to negative emotionality overall. Overall, individuals whose self-blaming tendencies were related to internal evaluation and striving for self-agency (i.e. adaptive guilt and maladaptive autonomy) displayed higher levels of self-blame-selective SCC activation, whereas external evaluation-related self-blame (sociotropy) showed no such relationship. Intriguingly, healthy

control individuals exhibited a reversed SCC response, namely its selective activation to other-blaming relative to self-blaming emotions.

Our factor analysis was in keeping with previous validation work (Robins et al., 1994; Yang & Girgus, 2019) that autonomy and sociotropy load onto different components and that adaptive guilt as operationalised by the TOSCA separates from all other measures (Green, Moll, Deakin, Hulleman, & Zahn, 2013a; Tangney et al., 1992). As predicted, self-hate, designed to be unrelated to concern for others (O'Connor et al., 2002), loaded onto the same factor as autonomy measures. Contrary to our predictions, shame did not show the expected stronger associations with sociotropy rather than autonomy. This may be due to the operationalisation of shame on the TOSCA as a characterological form of self-blame associated with feeling like hiding without specifying the external evaluation aspects of shame (Green et al., 2013a). As expected, our MDD group showed higher factor scores for both sociotropy and autonomy factors which is in keeping with their postulated role in MDD vulnerability (Clark et al.,
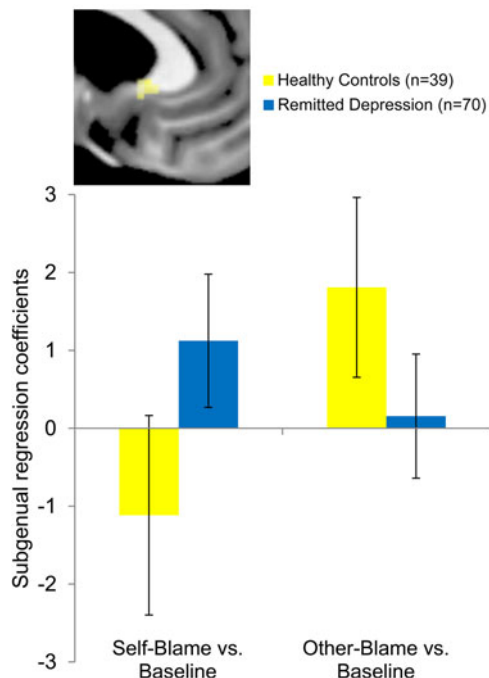
Fig. 1. A cropped section through the right SCC area (BA24) showing an interaction effect between group (MDD *v*. Control) and condition (self- *v*. other-blaming) is displayed using MRIcron (http://people.cas.sc.edu/rorden/mricron/install.html) at an uncorrected voxel-level threshold of *p* = 0.005, with no cluster-size threshold. The activation survived voxel-based familywise error-correction at *p* = 0.05 over our a priori SCC region of interest previously published in an independent sample (Green et al., 2012). As can be seen from the bar charts displaying extracted regression coefficient cluster averages and standard errors, this interaction was due to higher SCC signal for self-blame in the MDD group compared with the Control group and lower SCC signal for other-blame in the MDD group compared with the Control group. There were no main effects of group or condition in the SCC. SCC: subgenual cingulate cortex.

1995; Robins et al., 1994) whilst adaptive guilt was comparable between groups.

Confirming the predictions of our first hypothesis, we found that self-blame-selective SCC activation was higher in self-blame-prone individuals across diagnostic groups, in particular its internal evaluation-related forms such as self-hate and striving for autonomy. This is in keeping with the hypothesis that SCC activations in the detected anterior sector (BA24) may be related to social agency attributions in the context of self-blame (Zahn, Moll, et al., 2009). It is also consistent with the reproducible evidence on SCC activations in guilt-prone individuals without (Green et al., 2012; Zahn, Moll, et al., 2009) or with a history of MDD (Green et al., 2012).

Confirming the predictions of our second hypothesis, self-blame-selective SCC activation was higher in the remitted MDD *v*. Control group suggesting its role in MDD vulnerability. This is in keeping with the extensive literature on abnormalities in SCC activation in current MDD, which has been complicated by considerable variability in findings possibly related to loss of grey matter volume (Drevets, 1998) as well as SCC hyperactivity being stronger in familial *v*. non-familial MDD (Drevets et al., 1998). The localisation of SCC abnormalities reported in MDD varies; our finding of a BA24 activation is adjacent to, but outside the posterior subgenual cortex (BA25) used for deep brain stimulation (Mayberg et al., 2005) and it is likely that posterior and anterior sectors of the subgenual region are functionally

specialised (Zahn et al., 2020). Some authors prefer using the term 'subcallosal cingulate' (Hamani et al., 2011), although Brodmann called his area 25: 'Area Subgenualis' (Judas, Cepanec, & Sedmak, 2012). Despite these variations in terminology we follow Hamnani et al., to suggest treating 'subcallosal' and 'subgenual' as synonymous. It remains to be investigated whether the anterior subgenual cingulate regions (BA24/32) code for causal social agency contexts is due to their closer connection with pregenual anterior cingulate representations shown to correlate with subjective feelings of motor agency (Marchesotti et al., 2017) and emerging evidence for functional subdivisions between pregenual and subgenual areas in social learning (Lockwood & Wittmann, 2018).

Despite demonstrating higher self-blame-selective SCC activation in our MDD group relative to the Control group, we found no association with other indicators of MDD vulnerability, such as the number of previous episodes, or prospective recurrence risk [reported previously (Lythe et al., 2015)]. Further, we showed that self-blame-selective SCC activation was not associated with residual symptoms which indicates it is not directly comparable to SCC hyperactivity found in studies of symptomatic MDD (Drevets, 1998). These findings are most parsimoniously explained by assuming an association of self-blame-selective SCC activation with primary vulnerability factors for MDD, such as a tendency to internalise blame that may be adaptive and lead to prosocial behaviour (Tangney et al., 1992) but could interact with other factors such as specific life events to trigger overgeneralised forms of self-blame such as self-hate. We have previously shown that self-hate correlates with self-blame-selective abnormalities in functional connectivity between the SCC and the right anterior temporal cortex in MDD (Green et al., 2012) reflecting a lack of conceptual-emotional integration as a source of overgeneralisation (Green et al., 2013). This is in keeping with the notion that differentiated interpretations of social behaviour require integration of conceptual information about the social meaning of a situation as represented in the right superior anterior temporal lobe (Pobric, Lambon Ralph, & Zahn, 2016; Skipper, Ross, & Olson, 2011; Zahn et al., 2007; Zahn et al., 2017) with agency-context-related information in the SCC (Green et al., 2010).

The intriguing finding of other-blame-selective activation of the SCC in control participants necessitates a re-interpretation of our previous findings on the functional role of the SCC and its importance for self-blaming emotions. This result shows that failures in previous studies to detect SCC activation in response to self-blaming emotions without modelling individual differences in proneness to such emotions (Green et al., 2012; Moll et al., 2007; Zahn, de Oliveira-Souza, et al., 2009; Zahn, Moll, et al., 2009) were likely due to the fact that the SCC's role in self-blaming emotions is not selective for the self as previously asserted (Zahn, Moll, et al., 2009), but that it is equally important for other-blame-related emotions such as anger/indignation towards others. This is because one would be unlikely to find a functional specialisation of a brain region in one part of the population and the opposite function in another part of the population. To explain the individual differences in the direction of selectivity for either self- or other-blame in the SCC's response, it is most likely that shared cognitive/emotional components that are required for both self- and other-blame are recruited to varying degrees in different individuals depending on the context of self- *v*. other-agency.

Given that medial frontal activations were shown to increase when judging life events that had been more frequently

Table 3. Factorial model for fMRI activation in remitted MDD and Control group

| | | | MNI Peak Coordinates | | | | | |
| | | | | | | | | |
Hemisphere	Region	Cluster size	Brodmann area	x	y	z	F value	Voxel-based FWE-corrected p value
Group × Condition interaction effect								
Right	Anterior subgenual cingulate cortex	11	24	6	22	−2	9.46	0.05[a]

[a]Using our a priori subgenual cingulate region of interest [6 mm radius sphere around centre coordinate: MNI $x = -4$, $y = 23$, $z = -5$ (Green et al., 2012)] for multiple comparison correction. There were no main effects of agency or group in this region. No voxels survived voxel-based FWE-correction over the whole brain at $p = 0.05$ for main effects or interactions. FWE, familywise error; MNI, Montreal Neurological Institute. Talairach coordinates for the MNI peak were derived using Brett's MNItoTAL formula (Talairach $x = 6$, $y = 21$, $z = -3$) and referenced in the printed Talairach atlas (Talairach & Tournoux, 1988). Please note that although our a priori ROI's peak is in the left hemisphere, it includes the right SCC as well and interestingly, only a right hemispheric peak coordinate survived multiple comparison correction in this analysis.

Table 4. Individual differences in proneness to self-blame predict SCC activation

	B	S.E.	Beta	Statistic	p value
Overall model	n/a	n/a	n/a	F = 4.78	0.004**
Constant	−14.7	5.3	n/a	t = −2.76	0.007**
Autonomy/self-criticism	1.5	0.65	0.22	t = 2.33	0.02*
Adaptive Guilt (TOSCA)	0.27	0.65	0.26	t = 2.73	0.007**
Sociotropy/Dependency	0.32	0.12	0.04	t = .42	0.67

Cluster averages for the fMRI activation for self- v. other-blame were extracted for each individual and used as an outcome variable in a linear regression model in SPSS $n = 106$ ($n = 39$ control and $n = 67$ MDD). As predictor variables in this model, we used the identified two reliable principal components from our factor analysis of standard questionnaire measures of proneness to self-blaming emotions and whether people are more prone to blame themselves when evaluating themselves (autonomy) or when being evaluated by others (sociotropy), as well as the adaptive guilt measure from the TOSCA questionnaire. To avoid circular analyses, we did not primarily consider group here, because the extracted SCC activation cluster means were already biased by the SPM analysis to find the voxels showing a maximal group × condition interaction. Importantly, the factor analysis of questionnaire measures was independent of the fMRI analysis. * = significant at $p = 0.05$, ** = significant at $p = 0.01$. Please note that all predictor variables were modelled together and that betas therefore reflect partial effects adjusted for the other predictors in the model. As per the design, the Autonomy and Sociotropy factors were uncorrelated (Pearson's $r = 0$) and adaptive guilt as measured on the TOSCA showed no correlation with autonomy or sociotropy factors ($r < 0.08$, $p < 0.43$, $n = 106$).

encountered (Krueger et al., 2007), it is reasonable to assume that personal familiarity with blame attributions in either self- or other-agency contexts leads to more extended representations in the anterior SCC which are specific to agency context (self v. other). Based on the evidence on the externalisation of blame as a protective factor for self-esteem and against MDD in healthy populations (Mezulis, Abramson, Hyde, & Hankin, 2004), we speculate that their other-blame-selective SCC activation reflects their stronger familiarity with blame externalisation rather than internalisation compared with the MDD group. In contrast, people with MDD are more familiar with internalising blame to themselves as evidenced by persistent self-blaming biases (Green et al., 2013a; Zahn et al., 2015). This interpretation would also account for reproducible associations of proneness to self-blaming emotions with SCC activation as discussed above and is compatible with reduced SCC activation in psychopathy (Decety, Skelly, & Kiehl, 2013; Harenski & Hamann, 2006) which entails a lack of guilt (Hare, 2003).

On a more cautionary note, the following limitations of our study need to be discussed: to avoid multiple comparisons this study did not examine other regions of interest such as the frontopolar cortex which is the most reproducible region consistently activated for guilt across subjects (Moll et al., 2007; Zahn, Moll, et al., 2009), (Basile et al., 2011; Kédia, Berthoz, Wessa, Hilton, & Martinot, 2008; Morey et al., 2012; Seara-Cardoso et al., 2016; Takahashi et al., 2004) but is unlikely to be related to blame as it is also reproducibly found in fMRI studies of compassion (Immordino-Yang, McColl, Damasio, & Damasio, 2009), (Moll et al., 2007), (Fehse, Silveira, Elvers, & Blautzik, 2015; Kédia et al., 2008), which does not entail blame, compared against equally unpleasant and complex emotions. The non-selective

activation of the frontopolar cortex for prosocial feelings requiring the anticipation of complex consequences of actions/events such as compassion and guilt is in keeping with impairments in guilt and compassion (Moll et al., 2011) as well as selective impairments in the knowledge of long-term consequences of social behaviour in neurodegenerative lesions of the frontopolar cortex (Zahn et al., 2017). It is also important to note that this study deliberately focussed on people with MDD as their main diagnosis who were fully remitted and so our results may not generalise to patients with chronic MDD and co-morbid anxiety disorders.

Conclusions

Despite confirming the prediction of SCC activation in self-blame-prone individuals and those with remitted MDD, our results suggest that SCC activation is associated with blame irrespective of direction rather than selective for the self as previously argued. We speculate that patients with remitted MDD and those prone to self-blame have a more extended representation of blame-related information in the SCC in the context of self- v. other-agency with the opposite pattern occurring in healthy controls at low MDD risk which could explain these findings. These SCC representations may relate to the causal agency which would explain higher self-blame-selective SCC signal in individuals striving for autonomy. Future studies in people at familial risk of MDD prior to their first episode are needed to confirm our interpretation that self-blame-selective SCC activation is associated with primary vulnerability to MDD, and may interact with other factors, such as life events, to result in overgeneralised self-blame that was previously associated with changes in SCC functional connectivity rather than activation (Green et al., 2012).

References

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned help-lessness in humans - critique and reformulation. *Journal of Abnormal Psychology*, 87(1), 49–74.

Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *Neuroimage*, 13(5), 903–919. doi:10.1006/nimg.2001.0746.

Balint, M. (1959). *Thrills and regressions*. London: Maresfield Library.

Basile, B., Mancini, F., Macaluso, E., Caltagirone, C., Frackowiak, R. S., & Bozzali, M. (2011). Deontological and altruistic guilt: Evidence for distinct neurobiological substrates. *Human Brain Mapping*, 32(2), 229–239.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100.

Bowlby, J. (1977). The making and breaking of affectional bonds. I. Aetiology and psychopathology in the light of attachment theory. An expanded version of the Fiftieth Maudsley Lecture, delivered before the Royal College of Psychiatrists, 19 November 1976. *British Journal of Psychiatry*, 130, 201–210.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Clark, D. A., & Beck, A. T. (1991). Personality-Factors in dysphoria - a psycho-metric refinement of Beck Sociotropy-Autonomy Scale. *Journal of Psychopathology and Behavioral Assessment*, 13(4), 369–388. doi:10.1007/Bf00960448.

Clark, D. A., Steer, R. A., Beck, A. T., & Ross, L. (1995). Psychometric char-acteristics of revised sociotropy and autonomy scales in college students. *Behaviour Research and Therapy*, 33(3), 325–334.

Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology*, 103(1), 103–116. doi:10.1037/0021-843x.103.1.103.

Decety, J., Skelly, L. R., & Kiehl, K. A. (2013). Brain response to empathy-eliciting scenarios involving pain in incarcerated individuals with psychopathy. *JAMA Psychiatry*, 70(6), 638–645. doi:10.1001/jamapsychiatry.2013.27.

Drevets, W. C. (1998). Functional neuroimaging studies of depression: The anatomy of melancholia. *Annual Review of Medicine*, 49, 341– +. doi:10.1146/annurev.med.49.1.341.

Drevets, W. C., Ongur, D., & Price, J. L. (1998). Neuroimaging abnormalities in the subgenual prefrontal cortex: Implications for the pathophysiology of familial mood disorders. *Molecular Psychiatry*, 3(3), 220–226. doi:10.1038/sj.mp.4000370.

Eaton, W. W., Shao, H., Nestadt, G., Lee, H. B., Bienvenu, O. J., & Zandi, P. (2008). Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry*, 65(5), 513–520.

Ebert, D., & Ebmeier, K. P. (1996). The role of the cingulate gyrus in depres-sion: From functional anatomy to neurochemistry. *Biological Psychiatry*, 39(12), 1044–1050.

Fehse, K., Silveira, S., Elvers, K., & Blautzik, J. (2015). Compassion, guilt and innocence: An fMRI study of responses to victims who are responsi-ble for their fate. *Social Neuroscience*, 10(3), 243–252. doi:10.1080/17470919.2014.980587.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition.* (*SCID-I/P*). New York, NY: Biometrics Research, New York State Psychiatric Institute.

Freud, S. (1917). Trauer und melancholie. *Zeitschrift fuer Aerztliche Psychoanalyse*, 4(6), 288–301.

Green, S., Lambon Ralph, M. A., Moll, J., Deakin, J. F. W., & Zahn, R. (2012). Guilt-Selective functional disconnection of anterior temporal and subgenual cortices in Major depressive disorder. *Archives of General Psychiatry*, 69(10), 1014–1021.

Green, S., Lambon Ralph, M. A., Moll, J., Zakrzewski, J., Deakin, J. F., Grafman, J., … Zahn, R. (2013). The neural basis of conceptual-emotional integration and its role in major depressive disorder. *Social Neuroscience*, 8(5), 417–433.

Green, S., Moll, J., Deakin, J. F., Hulleman, J., & Zahn, R. (2013a). Proneness to decreased negative emotions in major depressive disorder when blaming others rather than oneself. *Psychopathology*, 46(1), 34–44. doi:10.1159/000338632.

Green, S., Ralph, M. A. L., Moll, J., Stamatakis, E. A., Grafman, J., & Zahn, R. (2010). Selective functional integration between anterior temporal and dis-tinct fronto-mesolimbic regions during guilt and indignation. *Neuroimage*, 52(4), 1720–1726.

Hamani, C., Mayberg, H., Stone, S., Laxton, A., Haber, S., & Lozano, A. M. (2011). The subcallosal cingulate gyrus in the context of Major depression. *Biological Psychiatry*, 69(4), 301–308. doi:10.1016/j.biopsych.2010.09.034.

Hare, R. D. (2003). *The hare psychopathy checklist-revised* (2nd ed.). Toronto: Multi-Health Systems.

Harenski, C. L., & Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *Neuroimage*, 30(1), 313–324.

Higgins, E. T. (1987). Self-discrepancy - a theory relating self and affect. *Psychological Review*, 94(3), 319–340.

Immordino-Yang, M. H., McColl, A., Damasio, H., & Damasio, A. (2009). Neural correlates of admiration and compassion. *Proceedings of the National Academy of Sciences of the United States of America*, 106(19), 8021–8026. doi:10.1073/pnas.0810363106.

Janoff-Bulman, R. (1979). Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of Personality and Social Psychology*, 37(10), 1798.

Judas, M., Cepanec, M., & Sedmak, G. (2012). Brodmann's map of the human cerebral Cortex - or Brodmann's maps? *Translational Neuroscience*, 3(1), 67–74. doi:10.2478/s13380-012-0009-x.

Kédia, G., Berthoz, S., Wessa, M., Hilton, D., & Martinot, J.-L. (2008). An agent harms a victim: A functional magnetic resonance imaging study on specific moral emotions. *Journal of Cognitive Neuroscience*, 20(10), 1788–1798.

Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., … Grafman, J. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), 20084–20089.

Leckman, J. F., Caruso, K. A., Prusoff, B. A., Weissman, M. M., Merikangas, K. R., & Pauls, D. L. (1984). Appetite disturbance and excessive guilt in major depression - use of family study data to define depressive subtypes. *Archives of General Psychiatry*, 41(9), 839–844.

Lockwood, P. L., & Wittmann, M. K. (2018). Ventral anterior cingulate cortex and social decision-making. *Neuroscience and Biobehavioral Reviews*, 92, 187–191. doi:10.1016/j.neubiorev.2018.05.030.

Lythe, K. E., Moll, J., Gethin, J. A., Workman, C., Green, S., Lambon Ralph, M. A., … Zahn, R. (2015). Self-blame-selective hyperconnectivity between anterior temporal and subgenual cortices and prediction of recurrent depressive episodes. *JAMA Psychiatry*, 72(11), 1–8.

Marchesotti, S., Martuzzi, R., Schurger, A., Blefari, M. L., del Millán, J. R., Bleuler, H., & Blanke, O. (2017). Cortical and subcortical mechanisms of brain-machine interfaces. *Human Brain Mapping*, 38(6), 2971–2989. doi:10.1002/hbm.23566.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., … Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, 45(5), 651–660. doi:10.1016/j.neuron.2005.02.014.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, 130(5), 711–747. doi:10.1037/0033-2909.130.5.711.

Moll, J., de Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M., Paiva, M. L., … Grafman, J. (2007). The self as a moral agent: Linking the neural bases of social agency and moral sensitivity. *Social Neuroscience*, 2(3-4), 336–352. doi:10.1080/17470910701392024.

Moll, J., Zahn, R., de Oliveira-Souza, R., Bramati, I. E., Krueger, F., Tura, B., … Grafman, J. (2011). Impairment of prosocial sentiments is associated with frontopolar and septal damage in frontotemporal dementia. *Neuroimage*, 54(2), 1735–1742.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134(4), 382–389.

Morey, R. A., McCarthy, G., Selgrade, E. S., Seth, S., Nasser, J. D., & LaBar, K. S. (2012). Neural systems for guilt from actions affecting self versus others. *Neuroimage*, 60(1), 683–692.

O'Connor, L. E., Berry, J. W., Weiss, J., Bush, M., & Sampson, H. (1997). Interpersonal guilt: The development of a new measure. *Journal of Clinical Psychology*, 53(1), 73–89. doi:10.1002/(Sici)1097-4679(199701)53:1 < 73::Aid-Jclp10>3.0.Co;2-I.

O'Connor, L. E., Berry, J. W., Weiss, J., & Gilbert, P. (2002). Guilt, fear, submission, and empathy in depression. *Journal of Affective Disorders*, 71(1–3), 19–27.

Pobric, G., Lambon Ralph, M. A., & Zahn, R. (2016). Hemispheric specialization within the superior anterior temporal Cortex for social and nonsocial concepts. *Journal of Cognitive Neuroscience*, 28(3), 351–360. doi:10.1162/jocn_a_00902.

Price, J. L., & Drevets, W. C. (2010). Neurocircuitry of mood disorders. *Neuropsychopharm*, 35(1), 192–216. doi:10.1038/npp.2009.104.

Pulcu, E., Zahn, R., Moll, J., Trotter, P. D., Thomas, E. J., Juhasz, G., … Elliott, R. (2014). Enhanced subgenual cingulate response to altruistic decisions in remitted major depressive disorder. *Neuroimage Clinical*, 4, 701–710. doi:10.1016/j.nicl.2014.04.010.

Ressler, K. J., & Mayberg, H. S. (2007). Targeting abnormal neural circuits in mood and anxiety disorders: From the laboratory to the clinic. *Nature Neuroscience*, 10(9), 1116–1124. doi:10.1038/nn1944.

Robins, C. J., Ladd, J., Welkowitz, J., Blaney, P. H., 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(4), 277–300. doi:10.1007/Bf02239408.

Seara-Cardoso, A., Sebastian, C. L., McCrory, E., Foulkes, L., Buon, M., Roiser, J. P., & Viding, E. (2016). Anticipation of guilt for everyday moral transgressions: The role of the anterior insula and the influence of interpersonal psychopathic traits. *Scientific Reports*, 6, 36273. doi:10.1038/srep36273.

Siegle, G. J., Carter, C. S., & Thase, M. E. (2006). Use of FMRI to predict recovery from unipolar depression with cognitive behavior therapy. *American Journal of Psychiatry*, 163(4), 735–738. doi:10.1176/appi.ajp.163.4.735.

Skipper, L. M., Ross, L. A., & Olson, I. R. (2011). Sensory and semantic category subdivisions within the anterior temporal lobes. *Neuropsychologia*, 49(12), 3419–3429. doi:10.1016/j.neuropsychologia.2011.07.033.

Stevens, J. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Routledge, Taylor & Francis Group.

Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: An fMRI study. *Neuroimage*, 23(3), 967–974.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system : An approach to cerebral imaging.* Stuttgart; New York: Georg Thieme.

Tangney, J. P., & Dearing, R. L. (2000). *The test of self-conscious affect- 3 (TOSCA-3)*. Fairfax: George Mason University.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, 58, 345–372.

Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology*, 101(3), 469–478.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070.

Yang, K., & Girgus, J. S. (2019). Are women more likely than men are to care excessively about maintaining positive social relationships? A meta-analytic review of the gender difference in sociotropy. *Sex Roles*, 81(3), 157–172. doi:10.1007/s11199-018-0980-y.

Zahn, R., de Oliveira-Souza, R., Bramati, I., Garrido, G., & Moll, J. (2009). Subgenual cingulate activity reflects individual differences in empathic concern. *NeuroscienceLetters*, 457(2), 107–110.

Zahn, R., De Oliveira-Souza, R., & Moll, J. (2020). Moral motivation and the basal forebrain. *Neuroscience and Biobehavioral Reviews*, 108, 207–217. doi:10.1016/j.neubiorev.2019.10.022.

Zahn, R., Green, S., Beaumont, H., Burns, A., Moll, J., Caine, D., … Ralph, M. A. L. (2017). Frontotemporal lobar degeneration and social behaviour: Dissociation between the knowledge of its consequences and its conceptual meaning. *Cortex*, 93, 107–118. doi:10.1016/j.cortex.2017.05.009.

Zahn, R., Lythe, K. E., Gethin, J. A., Green, S., Deakin, J. F., Workman, C., & Moll, J. (2015). Negative emotions towards others are diminished in remitted major depression. *European Psychiatry*, 30(4), 6. doi:10.1016/j.eurpsy.2015.02.005.

Zahn, R., Lythe, K. E., Gethin, J. A., Green, S., Deakin, J. F., Young, A. H., & Moll, J. (2015). The role of self-blame and worthlessness in the psychopathology of major depressive disorder. *Journal of Affective Disorders*, 186, 337–341. doi::10.1016/j.jad.2015.08.001.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences of the USA*, 104(15), 6430–6435.

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The neural basis of human social values: Evidence from functional MRI. *Cerebral Cortex*, 19(2), 276–283. doi:10.1093/cercor/bhn080.

SUPPLEMENTARY MATERIAL

Subgenual activation and the finger of blame: individual differences and depression vulnerability

Karen E Lythe, Ph.D.[1], Jennifer A Gethin, Ph.D..[1], Clifford I Workman, Ph.D.[1,2], Matthew A. Lambon Ralph, FRSLT, Ph.D.[1, 3], John F.W. Deakin, F.R.C.Psych., Ph.D.[2], Jorge Moll, M.D. Ph.D.[4],

Roland Zahn, M.D.[1,4,5,6]

[1]The University of Manchester & Manchester Academic Health Sciences Centre, School of

Psychological Sciences, Neuroscience and Aphasia Research Unit, Manchester, M13 9PL,UK

[2]The University of Manchester & Manchester Academic Health Sciences Centre,

Institute of Brain, Behaviour and Mental Health, Neuroscience & Psychiatry Unit,

Manchester, M13 9PL, UK

[3]MRC Cognition and Brain Sciences Unit, University of Cambridge, 15 Chaucer Road, Cambridge, CB2 7EF, UK

[4]Cognitive and Behavioral Neuroscience Unit, D'Or Institute for Research and Education

(IDOR), 22280-080 - Rio de Janeiro, RJ, Brazil

[5]Institute of Psychiatry, Psychology & Neuroscience, Department of Psychological Medicine, Centre for Affective Disorders,

King's College London, London, SE5 8AZ, UK

[6]National Service for Affective Disorders, South London and Maudsley NHS Foundation Trust, London SE5 8AZ

*Corresponding author

Dr Roland Zahn (see address above)

E-mail: roland.zahn@kcl.ac.uk

Phone: 0044-(0)20 7848 0348

Fax: 0044-(0)20 7848 0298

Supplementary Methods

Image acquisition details

An fMRI protocol optimised for detection of ventral brain regions was used as described previously (Green, Lambon Ralph, Moll, Deakin, & Zahn, 2012). T2*-weighted echo-planar images (3 runs of 405 volumes with 5 dummy scans) were acquired on an MRI scanner (3T Achieva, Philips) with an 8-channel head coil, 3mm section thickness, ascending continuous acquisition parallel to the anterior to posterior commissural line, 35-40 slices depending on the participant's head, repetition time=2000 milliseconds, echo time=20.5 milliseconds, field of view=220x220x120mm, acquisition matrix=80 x 80 voxels, reconstructed voxel size=2.29x2.29x3mm, and sensitivity encoding factor=2, enabling dynamic stabilisation to correct for signal drift.

T1-weighted, magnetization-prepared, rapid-acquisition gradient-echo structural images were obtained using 160 axial slices; 9mm slice thickness; repetition time=8.4ms; echo time=3.9ms; field of view=240x191x144mm; acquisition matrix=256x163 voxels; reconstructed voxel size=.94 x.94x.9mm; flip angle: 8°.

fMRI stimulus design details

As previously described (Green et al., 2012; Lythe et al., 2015; R. Zahn et al., 2015), participants were shown written descriptions of negative interactions between themselves and their best friends in which either they (self-agency condition, n = 90) or their best friend (other-agency condition, n = 90) acted counter to social and moral values. Statements read as ''[participant's best friend's name] does act e.g. tactlessly towards [participant's name]'' (other-agency condition, 90 items), ''[participant's name] does act e.g. tactlessly towards [participant's best friend's name]'' (self-agency condition, 90 items). The same social concepts (e.g. 'tactless', 'generous') were used in the self- and other-agency conditions, and 50% of the

stimuli used negative concepts (e.g. 'tactless') while 50% used negated positive concepts (e.g. 'not generously'). After the scan, participants were shown the stimuli again in a fully randomised order, but required to select the feeling that they felt was the best label for the emotion that they would experience most strongly in response to the social violation. The choice of feelings included shame, guilt, indignation/anger towards oneself, indignation/anger towards best friend, contempt/disgust towards oneself, contempt/disgust towards best friend, no feeling, or other feeling. Participants also rated how strongly they would experience unpleasant feelings as a result of the behaviour using a 1–7 visual analogue Likert scale (1 = not unpleasant, 7 = extremely unpleasant). This "Value-related Moral Sentiment Task" was based on an earlier version and details about the stimulus selection and design have also been described in (Green, Moll, Deakin, Hulleman, & Zahn, 2013; R. Zahn et al., 2007; Roland Zahn et al., 2009).

For the fMRI analysis, as explained in (Lythe et al., 2015), we used a broader definition of self-blaming and other-blaming emotions in the present study compared with our previous studies specifically investigating self-agency-related guilt vs. other-agency-related indignation/anger towards others (Green et al., 2012; Roland Zahn et al., 2009), which makes the results not directly comparable. The approach in the current study, however, increased the simplicity and power of our analysis for future applications and was justified by our finding that two important self-blaming emotions (shame and guilt) showed no BOLD activation differences in the SCC in a secondary data analysis (Pulcu et al., 2014). The change in approach was also to avoid large individual differences in the number of trials underpinning the computation of the BOLD response for self-blaming and other-blaming emotions, which may have confounded our previous analyses (Green et al., 2012; Roland Zahn et al., 2009).

Supplementary Results

There were also no correlations between the SCC coefficients for self- versus other-blame in the MDD group for rated unpleasantness or negative affectivity as measured on the Positive and Negative Affect scale (Watson, Clark, & Tellegen, 1988) in the self-blame (unpleasantness: $\rho=-.05$, $p=.68$; negative affectivity: $\rho=-.09$, $p=.36$) and other-blame conditions (unpleasantness: $\rho=-.15$, $p=.21$; negative affectivity: $\rho=-.03$, $p=.78$).

Supplementary Table 1. Exclusion reasons for volunteers following phone pre-screening. This table has previously been published in JAMA Psychiatry (doi:10.1001/jamapsychiatry.2015.1813)

Exclusion reason	*N*
MRI contraindications	77
Psychiatric disorders other than MDD	54
Current antidepressants or other centrally active medications	52
Withdrawal after telephone pre-screening	33
Not meeting full screening criteria for MDD	30
Family history of MDD/bipolar/schizophrenia (*Control* group)	26
Substance or alcohol abuse	23
Current antihypertensive or statin medications	20
Left-handed	20
Non-native English speaker	19
Thyroid function problems	19
Fulfilling criteria for current MDD	13
History of cancer	7
Not remitted for long enough (<6 months)	7
Epilepsy	5
No reason recorded	5
Other general medical conditions	5
Diabetes	4
Out of age range (18 – 65 years)	4
Excluded because of age-matching (*Control* group)	3
Multiple sclerosis	3
History of stroke	1
Vitamin D deficiency	1
Total excluded after phone pre-screening	**431**

In total, 707 people participated in the phone pre-screening interview, 276 passed this screening with 184 in the remitted MDD and 92 in the *Control* group and were invited for the first study day on which a full clinical interview was administered. Of these, 202 (138 individuals pre-screened as remitted MDD and 64 pre-screened as control participants) were reachable, able and willing to be seen on the first study day after reading the participant information sheet sent to them.

Supplementary Table 2. Exclusion reasons for participants following clinical interview.
This table has been adapted from a previously published one in JAMA Psychiatry
(doi:10.1001/jamapsychiatry.2015.1813)

Clinical group and exclusion reason	N
MDD group	
Fulfilling criteria for a bipolar disorder	6
Fulfilling criteria for current social anxiety disorder	6
Not meeting full criteria for MDD	5
Fulfilling criteria for past substance abuse	4
Not remitted for long enough (<6 months)	3
Residual symptoms of post-traumatic stress disorder	3
Probable personality disorders	2
Fulfilling criteria for current generalized anxiety disorder	1
MRI contraindications	1
Withdrawal after the clinical interview	1
Total MDD excluded after clinical interview	32
Control group	
Probable or definite positive first degree family history of MDD	4
Fulfilling criteria for a past MDE lasting less than two months	1
Fulfilling criteria for current adjustment disorder	1
Fulfilling criteria for current MDD	1
Fulfilling criteria for current social anxiety disorder	1
Non-native English speaker	1
Past depressive episode not fulfilling criteria for a past MDE	1
Total Control excluded after the clinical interview	10

After the clinical interview on the first study day, 160 participants were enrolled in the study (106 MDD and 54 Control participants). 144 participants completed the second study day which included the MRI scan (10/106 MDD and 6/54 were unable to schedule the second session). fMRI data for 138/144 participants were collected, with 6/144 participants not completing the fMRI acquisitions. Of the 138 participants for which fMRI data were collected, 91 were in the MDD group and 47 in the Control group. Data for 4/138 participants were excluded from the fMRI analysis due to abnormal images (3 MDD, 1 Control). 25/134 participants (18/88 MDD and 7/46 Control) were excluded entirely from fMRI analysis due to head movement and/or signal loss. fMRI data for 109 participants (70 MDD and 39 Control) had good signal coverage and mild movement (movement of up to 6mm translation and 2°rotation) and were included in the analyses.

Supplementary Table 3. Demographic variables and basic clinical characteristics

	MDD (*n*=70)	Control (*n*=39)	MDD vs Control comparison
Age	34.1 ± 12.0	33.4 ± 13.2	$t(107) = .32, p = .75$
Years of education	16.7 ± 2.4	17.4 ± 2.6	$t(107) = -1.33, p = .19$
Sex	20 male	15 male	$x^2 (1, N = 109) = 1.12, p = .29$
BDI score*	4.4 ± 4.2	1.0 ± 1.8	$t(101) = 5.87, p < .0001$
MADRS*	1.2 ± 1.5	.6 ± 1.2	$t(94) = 2.17, p = .03$
GAF*	85.0 ± 5.8	88.9 ± 2.8	$t(105) = -4.73, p < . 0001$

BDI, Beck Depression Inventory; MADRS, Montgomery-Åsberg Depression Rating Scale; GAF, Global

Assessment of Functioning Scale. *Significant at *p* < .05 threshold, 2-tailed. Means and standard deviations are

reported (*M ± SD*).

Supplementary Table 4. Clinical characteristics of the remitted MDD group
This table has been adapted from a previously published one in JAMA Psychiatry
(doi:10.1001/jamapsychiatry.2015.1813)

	MDD (*N*=70)
Past MDD subtype	
With melancholic features	35/70
With atypical features	8/70
No specific subtype	27/70
Number of previous MDEs	
1	16
2	21
3	14
4	6
5	7
6 or more	6
Average number of previous MDEs	3.24 ± 3.02 (range: 1-18)
Last MDE details	
Average length of MDE (months)	14.38 ± 17.93 (range: 1-96)
Average time in remission (months)	27.30 ± 23.95 (range: 5.5-140)
Severe depressive episode*	56/70
Moderate depressive episode*	14/70
No psychotropic medication since (months)	36.99 ± 57.94 (range: 0-372)
Previous medication	
SSRI antidepressant	54/70
SNRI antidepressant	3/70
Tricyclic antidepressant	3/70
Mirtazapine	1/70
Unknown class of antidepressant	6/70
No antidepressant medication	12/70
Benzodiazepines	2/70
Quetiapine	1/70
Previous CBT	12/70
Previous counselling	26/70
Self-guided CBT using internet or books	4/70
Previous suicide attempts	0.19 ± 0.39 (range: 0-1)
Life-time axis-I co-morbidity*	
Panic disorder with agoraphobia	1/70
Bulimia nervosa	1/70
Posttraumatic Stress Disorder	1/70
No life-time co-morbidity	67/70
Family history	
First degree relative with MDD	41/70
No family member with history of MDD	21/70
First degree relative with schizophrenia or bipolar disorder	7/70
Unknown	1/70

*According to ICD-10 criteria. **All co-morbid disorders were fully remitted at the time of study and none were likely to be the primary cause of the depressive episodes. MDD subtype classification was based on adapting the SCID-I for DSM-IV-TR to allow lifetime assessment of subtypes. All participants had stopped medication well before the required washout phase. Means and standard deviations are reported (M ± SD), or number of cases. CBT, cognitive behavioural therapy; MDE, major depressive episode; SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin norepinephrine reuptake inhibitor.

Supplementary Table 5. Ratings and response times for highly unpleasant self- and other-blaming emotion trials

	MDD (*n*=70)	Control (*n*=39)	MDD vs Control comparison
Frequency (% trials)			
Self-blaming emotion	59.6 ± 7.5	59.4 ± 12.7	*t*(53) = 0.09, *p* = .93
Other-blaming emotion	57.4 ± 6.8	57.6 ± 13.6	*t*(49) = -0.08, *p* = .93
Rated unpleasantness			
Self-blaming emotion	4.9 ± 1.0	4.3 ± 1.0	*t*(107) = 1.57, *p* = .12
Other-blaming emotion	4.6 ± 0.9	4.3 ± 1.0	*t*(107) = 1.42, *p* = .16
Response times (ms)			
Self-blaming emotion	2342 ± 476	2371 ± 424	*t*(106) = -0.31, *p* = .76
Other-blaming emotion	2383 ± 461	2379 ± 460	*t*(106) = 0.04, *p* = .97

There were no between-group differences on any of the above measures at p=0.05, 2-sided. Response time data were missing for one MDD participant. Means and standard deviations are reported (M ± SD).

Supplementary References

Green, S., Lambon Ralph, M. A., Moll, J., Deakin, J. F. W., & Zahn, R. (2012). Guilt-Selective Functional Disconnection of Anterior Temporal and Subgenual Cortices in Major Depressive Disorder. *Archives of General Psychiatry, 69*(10), 1014-1021.

Green, S., Moll, J., Deakin, J. F. W., Hulleman, J., & Zahn, R. (2013). Proneness to Decreased Negative Emotions in Major Depressive Disorder when Blaming Others rather than Oneself. *Psychopathology, 46*(1), 34-44.

Lythe, K. E., Moll, J., Gethin, J. A., Workman, C., Green, S., Lambon Ralph, M. A., . . . Zahn, R. (2015). Self-Blame-Selective Hyperconnectivity Between Anterior Temporal and Subgenual Cortices and Prediction of Recurrent Depressive Episodes. *JAMA Psychiatry, 72*(11), 1-8.

Pulcu, E., Lythe, K., Elliott, R., Green, S., Moll, J., Deakin, J. F., & Zahn, R. (2014). Increased amygdala response to shame in remitted major depressive disorder. *PloS One, 9*(1), e86900. doi:10.1371/journal.pone.0086900

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Zahn, R., Lythe, K. E., Gethin, J. A., Green, S., Deakin, J. F., Workman, C., & Moll, J. (2015). Negative emotions towards others are diminished in remitted major depression. *European Psychiatry, 30*(4), 6. doi:10.1016/j.eurpsy.2015.02.005

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proc Natl Acad Sci USA, 104*(15), 6430-6435.

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The Neural Basis of Human Social Values: Evidence from Functional MRI. *Cerebral Cortex, 19*(2), 276-283. doi:10.1093/cercor/bhn080