Neural responses to dynamic multimodal stimuli and pathology-specific impairments of social cognition in schizophrenia and depression

Christina Regenbogen, Thilo Kellermann, Janina Seubert, Daniel A. Schneider, Raquel E. Gur, Birgit Derntl, Frank Schneider and Ute Habel

Background
Individuals with schizophrenia and people with depression both show abnormal behavioural and neural responses when perceiving and responding to emotional stimuli, but pathology-specific differences and commonalities remain mostly unclear.

Aims
To directly compare empathic responses to dynamic multimodal emotional stimuli in a group with schizophrenia and a group with depression, and to investigate their neural correlates using functional magnetic resonance imaging (fMRI).

Method
The schizophrenia group (n = 20), the depression group (n = 24) and a control group (n = 24) were presented with portrait-shot video clips expressing emotion through three possible communication channels: facial expression, prosody and content. Participants rated their own and the actor's emotional state as an index of empathy.

Results
Although no group differences were found in empathy ratings, characteristic differences emerged in the fMRI activation patterns. The schizophrenia group demonstrated aberrant activation patterns during the neutral speech content condition in regions implicated in multimodal integration and formation of semantic constructs. Those in the depression group were most affected during conditions with trimodal emotional and trimodal neutral stimuli, in key regions of the mentalising network.

Conclusions
Our findings reveal characteristic differences in patients with schizophrenia compared with those with depression in their cortical responses to dynamic affective stimuli. These differences indicate that impairments in responding to emotional stimuli may be caused by pathology-specific problems in social cognition.

Declaration of interest
None.

Empathy is a social cognitive ability that can be conceptualised in different ways,[1] but most definitions include 'processes whereby one person can come to know the internal state of another and can be motivated to respond with sensitive care [. . .]'.[1] This partial sharing of emotions requires not only intact emotion recognition abilities but also adequate perspective-taking and theory-of-mind[2] abilities. Emotions can be transmitted through social cues emerging from several sensory channels, including facial expressions, prosody and speech content, and their integration[3] has beneficial effects on both emotion processing[4] and on empathy.[5] When cues present emotion in two but not a third channel (for example prosody and facial expression signal sadness, but speech content remains neutral), understanding the situation requires more effort and empathy ratings decrease.[6] Diminished empathy often accompanies psychiatric disorders[7] such as schizophrenia and major depression. Despite distinct psychopathologies, both disorders overlap in long-lasting and multifaceted deficits in social cognition,[8,9] which are evident already on a basic sensory level. Both disorders show impaired recognition of emotions presented via facial cues and via prosody (schizophrenia,[10,11] depression[12,13]); in addition, both show aberrant functional activation patterns in the superior temporal gyri towards emotional prosody and in the fusiform gyri towards emotional facial expressions (schizophrenia,[14,15] depression[16,17]). Whether these functional similarities are the result of similar underlying pathological mechanisms, however, remains poorly understood. When perceiving and responding to social situations, patients with depression demonstrate lower self-report empathy but higher galvanic skin conductance levels,[18] which are associated with limbic–cortical dysfunctions.[19] They have also been shown to overidentify with emotions and display higher levels of personal distress,[20] yet at the same time, blunted affect and anhedonia.[18] Impaired social cognition in schizophrenia, on the other hand, has been linked to both negative symptoms such as diminished emotional experience and avolition,[21] but also to positive symptoms such as delusions, hallucinations and disorganised speech and behaviour. Their difficulties are most profound during processing of complex mental states, including their own (for example appraisal and regulation[22]) and those of others (empathy[23]). Potential precursors of these difficulties may be found in patients' impairments to establish coherent constructs,[24] deal with ambiguity,[25] ignore irrelevant information[26] and form associative memory traces.[27,28] Direct comparisons of social cognition skills between these two patient groups are required to directly test the specificity of these differences.

Although comparable deficits have been reported in individuals with schizophrenia and those with depression during audiovisual processing presentations of static stimuli,[29,30] these fail to convey evaluative aspects of empathy to the same extent as natural situations do,[31] and may have led to systematic underestimations of group differences.[32,33] This study therefore used a novel set of stimuli to investigate the neural correlates of *dynamic* stimulus presentation in these two patients groups. Participants were presented with portrait-shot video clips expressing emotion through three possible communication channels: facial expression, prosody and content. We pursued two interrelated aims. The first aim was to compare brain activation and behavioural indices of empathy for both trimodal and bimodal stimuli in a group with

schizophrenia and a group with depression. Trimodal stimuli, rich in emotional information, should facilitate understanding and appropriate responses to social situations. Bimodal stimuli, on the other hand, introduce mismatching information that requires integration of matching channels while ignoring irrelevant ones. We expected that (a) patients' empathy ratings towards trimodal emotional stimuli would lie within the normal range, but would be more strongly affected than healthy controls by the absence of emotional information in one channel.[34] On a neural level, we hypothesised that (b) patients with schizophrenia would show abnormal activation to bimodal stimuli in multisensory integration areas, for example in the intraparietal and superior temporal sulci,[35] but also in the hippocampus,[36] because of difficulty ignoring irrelevant information[26] and establishing coherent constructs.[24] Patients with depression were expected to show aberrant brain activation towards social stimuli because of limbic-cortical dysfunctions.[19,37,34] The second aim was to test whether corresponding emotional information on multiple sensory channels would lead to the same characteristic activation patterns in patients as previously reported in healthy controls. We hypothesised that (c) multisensory emotional facilitation would be reflected in signal increases in sensory cortices, i.e. increased activation in primary and secondary visual cortices for facial expressions, in primary and secondary auditory cortex for prosody, as well as in left-lateralised parietotemporal stream for speech content in the trimodal compared with the bimodal conditions. Alternatively, an absence of such effects would indicate failure to benefit from multimodality.[29,30]

Method

Participants

Twenty patients with paranoid schizophrenia (schizophrenia group, mean age 37.30 years, s.d = 8.44) and 24 age- and education-matched patients with major depression (depression group, mean age 36.42 years, s.d. = 12.01) were included in the study. Patients were recruited from in- and out-patient treatment facilities of the Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University. Clinical diagnoses were confirmed by C.R. and D.A.S. using the Structured Clinical Interview for DSM Disorders (SCID).[38] Symptom severity was assessed using the Positive and Negative Syndrome Scale (PANSS[39]) in the schizophrenia group and the Beck Depression Inventory (BDI-II[40]) and the Hamilton Rating Scale for Depression (HRSD-17[41]) in the depression group. All participants in both groups showed moderate to severe symptomatology (online Table DS1). The 24 healthy comparison participants (control group, mean age 35.25 years, s.d. = 9.80) comprised an age- and education-matched sub-data-set.[6] They were screened for a lifetime DSM-IV diagnosis (SCID-light[38]), neurological illness or current substance misuse. All participants were right-handed[42] and fulfilled magnetic resonance scanning criteria.

Participants were tested with a standard neuropsychological battery that is described in the online supplement. They further completed an alexithymia questionnaire (Toronto Alexithymia Scale TAS-20[43]) and two empathy scales, the German version of the Interpersonal Reactivity Index (IRI[44]) as well as the E-Scale.[45] Average scores were analysed with one-factorial ANOVAs with group as between-participant factor. *Post hoc* pair-wise comparisons were Bonferroni-corrected ($P_{Bonf} < 0.017$).

Empathy assessment

Stimuli consisted of 96 thoroughly evaluated[5] video clips (duration, mean 10.93 s, s.d = 0.93), alternating a male ($n = 44$)

or female ($n = 52$) protagonist who told a self-related story of disgusting, fearful, happy, sad or neutral situations. Prior to filming, all actors had been asked to imagine their story as vividly as possible, and to remember an emotionally corresponding life experience. Clips existed in six experimental conditions comprising 16 videos each and did not differ in length ($F(5,90) = 1.26$, $P = 0.29$). We refrained from analysing emotion-specific aspects and collapsed results across emotions (four clips in each condition).

Experimental conditions reflected the grouping of emotional information carried by the three channels. 'Trimodal emotional' included emotion in story content, facial expressions and prosody. 'Trimodal neutral' included neutral story content, facial expressions and prosody. In the bimodal emotional conditions 'neutral prosody', 'neutral face' or 'neutral speech', two channels were emotional (with the same target emotion), whereas the third was neutral. 'Foreign language' consisted of a trimodal emotional clip, spoken in Slavic languages that were incomprehensible to all participants. Synchronicity between the dubbed audio files and video streams (for conditions in which prosody and facial expression were incongruent; neutral prosody and neutral facial expression) was maximised with Digidesign ProTools for Windows (http://www.avid.com).

In the magnetic resonance scanner, participants were instructed to imagine the presented actors to be close friends. After each clip they rated the other's as well as their own emotional state and its intensity via right-hand button presses (index and middle finger) on a seven-point scale from 'very negative' on the left side of the scale to 'very positive' on the right side within a fixed interval of 4.5 s (online Fig. DS1). Matching ratings between the rating for 'other' and the valence of the target emotion represented emotion recognition ('other'). Matching ratings between the participant's own affective state and the target emotion represented affective responses ('self'). Overlap between both ratings and the target valence represented task empathy (for example, a sad video clip evokes a response in the same direction (one, two or three finger presses towards 'very negative') for the 'self' and 'other' assessment). Our empathy definition is closely related to empathic accuracy,[46] however, it further demands that the receiver shares the recognised emotion (see also de Vignemont & Singer;[47] empathy is an affective state isomorphic to another person's state, elicited by that person; however, we did not enquire whether participants actually attributed their affective state to have been elicited by the other person.) To test hypothesis (a) empathy and its components (emotion recognition and affective responses) were tested in three separate group × condition generalised estimating equations (GEEs) accounting for binomial data distribution and non-spherical errors. (The condition trimodal neutral was not analysed because responses to this condition did not fall under the definition of empathy.) *Post hoc* pair-wise comparisons were Bonferroni-corrected ($P_{Bonf} < 0.005$).

Cognitive impairments in psychiatric disorders have been regarded as an accompanying vulnerability factor and endopheno-type of not only individuals with acute schizophrenia, but also patients in remission or relatives.[48] Correlation analyses were therefore performed between task empathy and cognitive measures (Wortschatztest (WST), Trail Making Test (TMT-A, TMT-B) Regensburger Word Fluency Test (RWT), working memory, short-term memory), and between task empathy and emotional self-report measures (Vienna Emotion Recognition Task (VERT-K), IRI, E-Scale), as well as clinical parameters in patients (depression group: BDI, HRSD, TAS; schizophrenia group: PANSS-Positive, PANSS-Negative, TAS). All correlations were two-tailed Bonferroni-corrected Pearson correlations.

Functional magnetic resonance imaging

Whole brain analyses

Detailed parameters of the 3 Tesla functional magnetic resonance (fMRI) measurements and data preprocessing can be found in the online supplement. On a single-participant level, seven onset regressors (one for each stimulus presentation by condition and one for the rating scale presentation) were created by convolving the respective box-car functions with the canonical haemo-dynamic response function (HRF). Realignment parameters were included as covariates of no interest and the session mean was regressed on a constant term. A 128 s high-pass-filter was applied and serial auto-correlations were accounted for by including a first order autoregressive model.

A mixed-effects 3×6 general linear model (GLM) (group $\times$ condition) was used for group-level inference in SPM8 (Wellcome Department of Cognitive Neurology, London). A random factor modelled participant-specific variance. Deviations from sphericity were corrected by variance components assuming a compound symmetry structure for within-participant measures and heteroscedasticity between participants and conditions.

All contrasts (for description of contrasts 1–5, see online Appendix) were calculated within this GLM and thresholded with a combined height and extent threshold technique based on Monte-Carlo simulations carried out in 3DClustSim in AFNI2011.[49] Based on an uncorrected threshold of $P < 0.001$ and the spatial properties of the residual image, an extent threshold of $k = 125$ voxels was estimated using 100 000 iterations. This corresponded to a family-wise error of $P < 0.05$ at the cluster-level. Post hoc pair-wise comparisons compared the mean activations between clusters (adjusted for effects of interest) with t-tests (Bonferroni-corrected) in SPSS version 20 for Windows.

Comparing trimodal and bimodal conditions

In order to compare brain activation for both trimodal and bimodal stimuli in patients and controls we tested the full interaction group (3) $\times$ condition (6) (contrast 1).

Comparing trimodal emotional and trimodal neutral conditions

We used the planned interaction between group and condition, comparing trimodal emotional to trimodal neutral (contrast 2) to specifically test differences between congruent emotional and neutral stimuli.

Both contrasts should test differences between patients and controls in trimodal and bimodal emotional stimulus processing (hypothesis (b)), contrast 1 allowing for all kinds of possible differences, contrast 2 tailored to directly compare emotional and neutral stimuli.

Channel-sensitive contrasts

Finally, we analysed planned T-contrasts emotional prosody (trimodal emotional > neutral prosody), emotional face (trimodal emotional > neutral face) and emotional speech (trimodal emotional > neutral speech) separately for each participant to assess the neural contributions of dynamic emotion in single channels. To test potential differences between patients and controls regarding multisensory emotional facilitation (hypothesis (c)) planned F-contrasts were conducted (contrasts 3, 4, and 5).

Results

Behavioural empathy

The GEE testing empathy (Fig. 1, online Table DS2) revealed a significant main effect of condition (Wald $\chi^2(4) = 222.06$, $P < 0.001$) and a trendwise interaction between group and condition (Wald $\chi^2(8) = 15.26$, $P = 0.051$). The main effect of group (Wald $\chi^2(2) = 0.14$, $P = 0.93$) was not significant.

Explorative post hoc comparisons of the interaction effect showed that it was not driven by differences between the three groups in any condition, but by within-group effects. In the control group, empathy was significantly higher in the trimodal emotional condition compared with all other conditions except neutral prosody (neutral facial expression: $t(23) = 5.15$; neutral speech: $t(23) = 8.71$; foreign language: $t(23) = 6.70$, all $Ps < 0.001$). This effect was replicated in the depression group, but also included neutral prosody ($t(23) = 4.45$; neutral facial expression: $t(23) = 4.40$; $Ps = 0.001$; neutral speech: $t(23) = 7.61$; foreign language: $t(23) = 7.21$, $Ps < 0.001$). In the schizophrenia group, only those conditions in which speech content was neutral ($t(19) = 6.34$) or foreign ($t(19) = 6.81$, $Ps < 0.001$) prompted lower empathy ratings. Rating results regarding 'other' and 'self' are detailed in the online supplement.

Functional imaging data

Comparing trimodal and bimodal conditions

The F-contrast group $\times$ condition (contrast 1) yielded activation differences in four areas: the left hippocampus, supplementary motor area, as well as the right intraparietal sulcus and middle temporal gyrus (Fig. 2, Table 1). Post hoc pair-wise comparisons were conducted in each area separately to further break down these effects. In the hippocampus, higher activation was observed in the schizophrenia group in the trimodal neutral condition compared with both the control ($t(42) = 3.26$, $P = 0.002$) and depression group ($t(42) = 3.57$, $P = 0.001$). The schizophrenia group also showed less activation in response to neutral speech trials compared with both the control group ($t(42) = -6.08$, $P < 0.001$) and depression group ($t(42) = -3.40$, $P = 0.002$).

In the left supplementary motor area, the depression group showed significantly higher activation during trimodal emotional trials compared with the control group ($t(46) = 3.54$, $P = 0.001$).
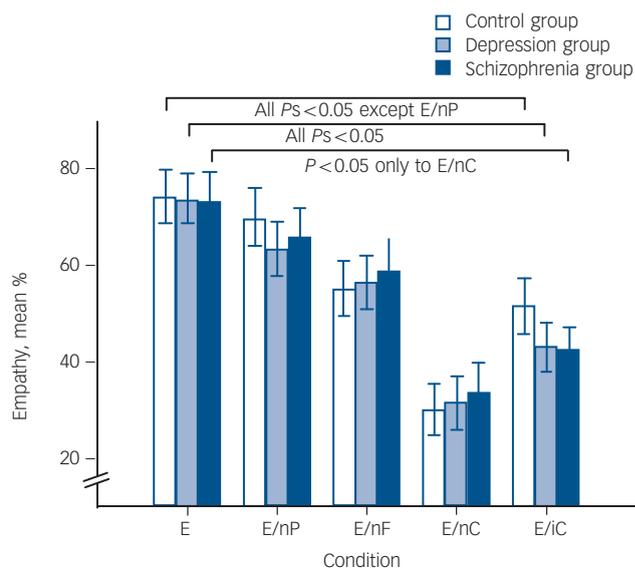

Fig. 1 Participants' average ratings on recognising and feeling the target valence and intensity (empathy) are indicated by bars for each group.

Post hoc comparisons following a generalised estimating equation (GEE) were calculated within each group. The lines represent significant differences between condition trimodal emotional (E) and the remaining conditions. Bars indicate s.e.m. EnP, neutral prosody; EnF, neutral facial expression; EnC, neutral speech; EiC, foreign language.
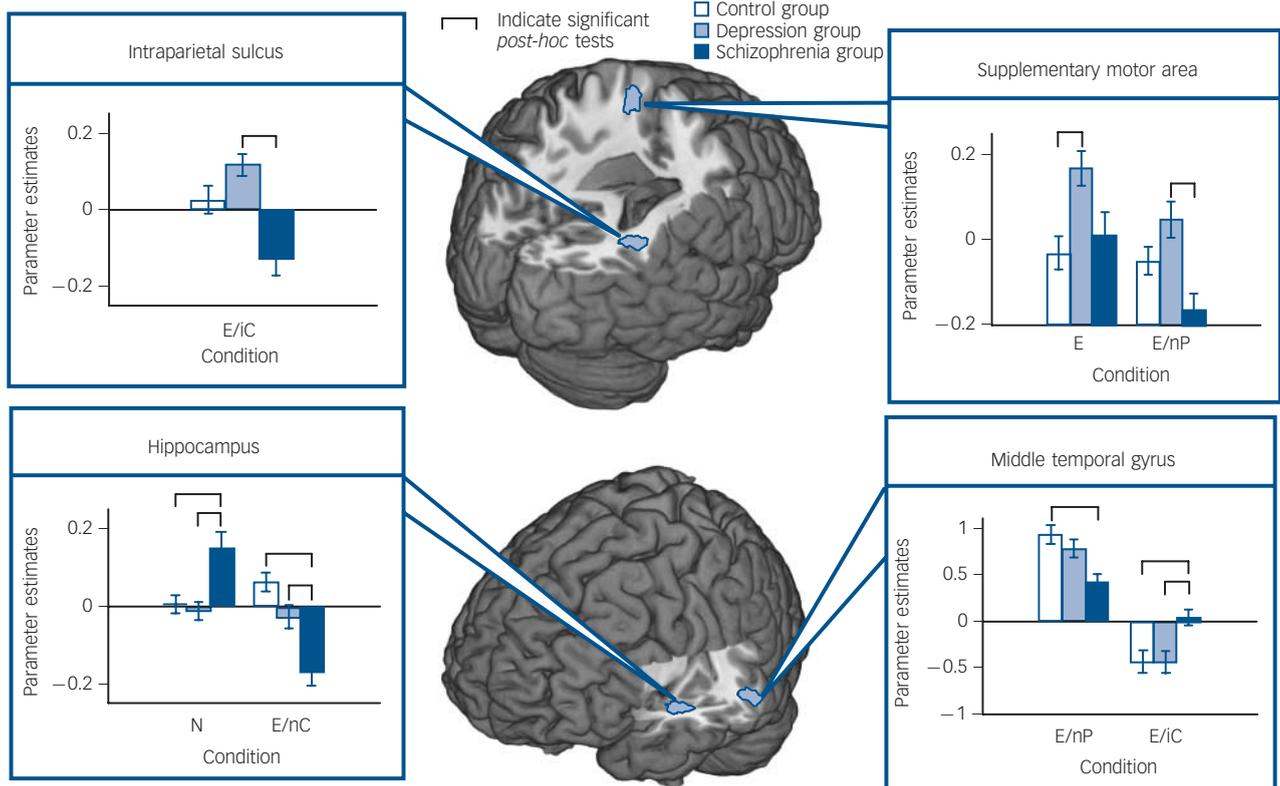
Fig. 2 Full interaction of group × condition (contrast 1, all conditions, random-effects general linear model, $F > 3.04$, Monte-Carlo-cluster-level corrected, $P < 0.05$, $k > 125$, coordinates see Table 1).

Significant *post hoc* comparisons of parameter estimates (mean, s.e.m., arbitrary units) are displayed via bar charts and were calculated within each group and condition. In the schizophrenia group: lower activation in neutral speech (E/nC) in the hippocampus, in foreign language (E/iC) in the intraparietal sulcus and in neutral prosody (E/nP) in the supplementary motor area; higher activation in trimodal neutral (N) in the hippocampus; less deactivation in foreign language (E/iC) in the middle temporal gyrus. In the depression group: higher activation in trimodal emotional (E) in the supplementary motor area.

Table 1 Functional magnetic resonance imaging activation of the full interaction (contrast 1) between participant group and communication conditions ($F > 3.04$, Monte-Carlo-cluster-level-corrected, $P < 0.05$, masked inclusively with contrast 1, $P < 0.05$, uncorrected, to limit the results cluster-corrected whole-brain results) as well as the planned interaction between group and trimodal emotional *v.* trimodal neutral (contrast 2)[a]

Contrast	Hemisphere	Size	F	x	y	z
1 Full interaction						
Hippocampus	Left	566	6.04	−35	−18	−15
Middle temporal gyrus	Left	197	3.86	−65	−41	−8
Supplementary motor area	Left	169	4.41	−17	−3	60
Intraparietal sulcus	Right	185	5.28	34	−46	22
2 Trimodal emotional *v.* trimodal neutral						
Posterior cingulate cortex	Left	201	11.70	−6	−30	50
Temporoparietal junction	Right	154	10.42	47	−35	23

a. Stereotaxic coordinates of local maxima of activation are expressed as *x, y, z* values in Montreal Neurological Institute space.

They also showed higher activation in the neutral prosody condition compared with the schizophrenia group ($t(42) = 3.60$, $P = 0.001$). In the intraparietal sulcus, the schizophrenia group showed lower activation in the foreign language condition compared with the depression group ($t(42) = -4.87$, $P < 0.001$). In the middle temporal gyrus, the schizophrenia group showed lower activation in the neutral prosody condition compared with the control group ($t(42) = -3.84$, $P < 0.001$) and higher activation in the condition foreign language compared with the control ($t(42) = 3.23$, $P = 0.002$) and the depression group ($t(42) = 3.31$, $P = 0.002$).

Comparing trimodal emotional and trimodal neutral conditions

The planned *F*-contrast group × condition (contrast 2) showed activation in the left temporoparietal junction and the left posterior cingulate cortex (Fig. 3). *Post hoc* comparisons in the temporoparietal junction showed lower activation for the depression group than the control group in the trimodal neutral condition ($t(46) = -3.68$, $P = 0.001$). In the posterior cingulate cortex, higher activation for the depression group in the trimodal emotional condition than for the control ($t(46) = 3.51$,
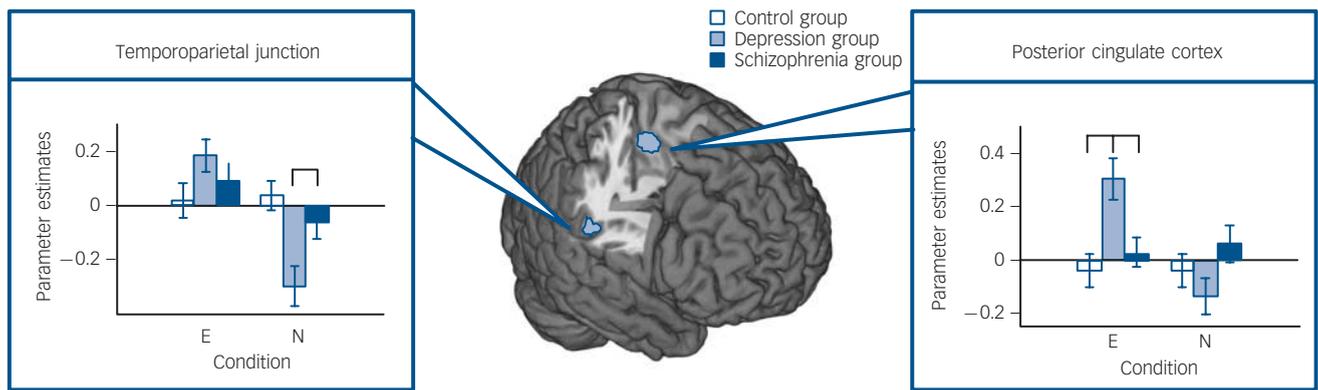
Fig. 3 Multimodal emotional *v.* multimodal neutral (contrast 2). Planned interaction of group condition (conditions trimodal emotional (E) and trimodal neutral (N), random-effects general linear model, $F > 7.03$, Monte-Carlo-cluster-level corrected, $P < 0.05$, $k > 125$, coordinates see Table 1).

Significant *post hoc* comparisons of parameter estimates (mean, s.e.m., arbitrary units) are displayed via bar charts. In the depression group: lower activation in trimodal neutral (N) in the temporoparietal junction; higher activation in trimodal emotional (E) in posterior cingulate cortex.

$P = 0.002$) or the schizophrenia group ($t(42) = 2.78$, $P = 0.008$) were observed.

Channel-sensitive contrasts

Within each participant group, *T*-contrasts comparing the trimodal emotional condition with each bimodal emotional condition revealed activation in areas responsible for processing the respective sensory modality (online Table DS3–5, online Fig. DS2). Bilateral emotion-related enhancements of auditory cortex activation and inferior parietal lobe activation were present for the emotional prosody subtraction contrast; bilateral occipito-temporal and middle occipital gyrus activation enhancement was present in the emotional face subtraction contrast. Finally, left angular gyrus and superior parietal lobe, middle temporal gyrus and middle and inferior frontal gyrus activation was enhanced in response to the emotional speech subtraction contrast.

Group-comparisons of each channel-contrast (contrasts 3–5) showed significant group differences only for neutral speech in the hippocampus, supplementary motor area and putamen/caudate nucleus (Table 2). The extracted activation did not survive Bonferroni-correction.

Neuropsychology and empathy questionnaires

The ANOVA testing alexithymia (TAS, $F(2,64) = 5.03$, $P = 0.01$) indicated higher TAS scores in the depression group than in the control group: ($t(46) = 3.17$, $P = 0.003$). Another significant main effect for group was found in one empathy scale (IRI:

$F(2,64) = 3.92$, $P = 0.03$) with *post hoc* comparisons showing lower scores in the schizophrenia group than in the depression ($t(42) = -2.73$, $P = 0.009$) or the control group ($t(42) = -2.26$, $P = 0.03$), but this was below the Bonferroni-correction threshold ($P < 0.017$). A complete overview on all neuropsychological variables and empathy questionnaires can be found in online Table DS1.

No correlation between task empathy and cognitive measures, and between task empathy and emotional self-report or clinical parameters survived Bonferroni-correction (full results can be found in the online supplement).

Discussion

Using a novel set of dynamic multisensory stimuli, our study directly compared neural responses to emotional stimuli between participants with schizophrenia and participants with depression in an effort to identify commonalities and differences in the underlying mechanisms contributing to disturbed social cognition in these patient groups.

Although empathy ratings in neither patient group were significantly different from healthy controls, characteristic differences emerged in the bimodal emotional conditions. In combination with the activation patterns from the neuroimaging analyses, our findings link problems in social cognition primarily to impaired processing of semantic context in the schizophrenia group, whereas the patterns observed in the depression group indicate contributions of abnormal activity in mentalising networks.

Table 2 Functional magnetic resonance imaging activation of trimodal emotional *v.* neutral prosody (contrast 3), trimodal emotional *v.* neutral facial expression (contrast 4), trimodal emotional *v.* neutral speech content (contrast 5) ($Fs > 7.29$, Monte-Carlo-cluster-level-corrected, $P < 0.05$)[a]

Contrast	Hemisphere	Size	F	x	y	z
3 Trimodal emotional *v.* neutral prosody						
No significant brain regions	–	–	–	–	–	–
4 Trimodal emotional *v.* neutral facial expression						
No significant brain regions	–	–	–	–	–	–
5 Trimodal emotional *v.* neutral speech content						
Supplementary motor area	Left	173	14.26	−18	−5	62
Putamen/caudatus nucleus	Right	358	9.84	21	14	−5
Hippocampus	Left	204	9.55	−36	−21	−17

a. Stereotaxic coordinates of local maxima of activation are expressed as *x, y, z* values in Montreal Neurological Institute space.

Behavioural and neural markers of empathy assessed with multimodal stimuli

Our study demonstrated similar behavioural responses of empathy during dynamic trimodal stimulus presentation in the schizophrenia, depression and control groups, as well as an absence of previously documented activation impairments of patients in the superior temporal[14] or fusiform gyri[15,17] to emotional prosody or facial expression alone (contrasts 3 and 4, see online Appendix). In all participant groups, emotional information in any sensory channel enhanced activation in sensory-specific areas relative to emotionally neutral cues, replicating our previous findings in healthy participants;[6] activation to emotional prosody was mainly found in primary and secondary auditory cortices,[50,51] activation to emotional faces in the bilateral fusiform cortices,[52] and to emotional speech in areas of the semantic language stream[53] and cerebellum, representing content predictability.[54]

Although a direct comparison with static stimuli is not possible, we tentatively attribute the lack of observed behavioural impairments to potential benefits of our more naturalistic and dynamic stimulus material.[31,55] This view is strengthened by the observed robust emotion-enhanced activations in basic sensory processing areas in both patient populations (Fig. DS2). Direct comparisons between static and dynamic material should confirm these potential benefits in future studies. Further, it should be noted that all conditions included at least bimodal emotions; our definition of 'single channel' here relies on the subtraction of bimodal from trimodal conditions. This relative difference fails to capture potential supra-additive effects that may occur in the trimodal condition, which may have influenced activity in modality-respective areas.

In line with our findings in healthy participants,[5,6] patients demonstrated behavioural empathy reductions in the bimodal emotional conditions. Whereas neutrality of any channel affected the control and depression groups, lower empathy ratings in the schizophrenia group were only prompted by the neutral speech and foreign language conditions. Given that empathic appraisal[47] depends on the successful joint evaluation of emotional cues and contextual information, *what* was being said needed to be ignored in these two conditions to elicit a correct empathic response. We suggest that healthy participants focused on the emotional channels and suppressed or ignored neutral speech content.[6] Such a strategy would be consistent with moderately high emotion recognition accuracy (70–80%), yet lower affective involvement (approximately 30%) resulting in reduced empathy in this condition compared with the trimodal emotional one.

Patients with schizophrenia demonstrate difficulty integrating ambiguous social constructs

In the schizophrenia group, behavioural empathic responses remained stable between trimodal and most bimodal conditions. Lower empathy responses were limited to conditions where speech content did not contain or signal emotion (neutral speech, foreign language). This response pattern contrasted with both the control and depression groups, who were equally affected by neutrality of any other sensory channel, and it suggests a focus on concrete contextual information in the schizophrenia group during the formation of an empathic response, while neglecting paraverbal information. This interpretation is supported by the neuro-imaging results, which show characteristic hypoactivation patterns during the two conditions in which speech content did not match facial expression and prosody: hippocampal activation was reduced during neutral speech relative to both other groups, whereas intraparietal sulcus activity was decreased during foreign speech (Fig. 2).

Slightly enhanced hippocampal activation in the control group in contrast during neutral speech may reflect successful emotional understanding and association of two emotional channels and ignoring the third. Together with dorsal striatum (our *post hoc* comparisons in the striatum did not survive Bonferroni-correction; however, given the empirical background on hippocampal–striatum co-work in construct formation, it seemed appropriate to include this in the discussion) and thalamic nuclei,[56] the hippocampus serves associative memory encoding.[57] In the context of multimodal processing, its sensitivity to semantic variations between sensory modalities has been demonstrated.[36] In contrast, the hypoactivation of the hippocampus observed in the schizophrenia group may therefore reflect either failure to make an effort to associate incoming information or failure to ignore it because of its irrelevance.[26] Hypoactivation of the intraparietal sulcus during the foreign language condition may also indicate unsuccessful integration of the emotional paraverbal percept with verbal information, or, alternatively, failure to concentrate on the paraverbal percept and ignore mismatching semantic information.[58] Of interest, the schizophrenia group further showed higher hippocampal activity than both other groups during the trimodal neutral condition, which may indicate inadequate emotional hyperarousal to potentially ambiguous sources.[59] This hypothesis is supported by a lack of deactivation in the middle temporal gyrus; as a key area in meaningful speech comprehension,[60] its continued activation may indicate an increased but unsuccessful effort to infer emotion from neutral content. In the absence of a corresponding behavioural deficit, however, the functional relevance of this enhanced activity remains speculative and needs to be further explored.

Depression is associated with inadequate activation related to mentalising

Although behavioural empathy ratings in the depression group did not significantly differ from the control group, these patients consistently showed hypoactivation in the posterior cingulate cortex and supplementary motor area in trimodal emotional clips, as well as hypoactivation in the temporoparietal junction compared with the control group during the trimodal neutral condition. All of these regions play key roles in the brain's mentalising network; in particular, the supplementary motor area and the posterior cingulate cortex are part of a medial fronto–parietal axis responsible for automatic processing of internally oriented personal information such as thoughts, feelings and experience.[61] The temporoparietal junction, on the other hand, forms part of a lateral fronto–temporoparietal axis, which is externally oriented and attributed to controlled appraisal of situational information, such as labelling another's facial expression. Hyperactivations of the medial fronto–parietal axis (supplementary motor area and posterior cingulate cortex) have been linked to hyperresponsiveness to emotions,[34] whereas hypoactivation of regions on the lateral fronto–temporo–parietal axis (temporoparietal junction) have been described before as correlates of impaired cognitive control.[62]

Further, the depression group showed higher alexithymia (TAS) ratings than the control group, indicating 'difficulties in assessing and describing feelings'.[63] These ratings were negatively correlated with task empathy, and may constitute state-dependent indicators[64] of difficulties with reflections on ones' internal affective state and active reappraisal of emotions within a social context. Both of these functions depend on adequate involvement of the cortical networks for mentalising, as described above. However, as the correlation did not survive Bonferroni-correction and depressive symptom severity was only weakly associated with

alexithymia (BDI: $r = 0.23$, HRSD: $r = 0.14$) these results should be interpreted with caution.

Limitations

Although our experimental design was optimised to evaluate responses to a diverse battery of stimulus material of long duration between clinical populations, it also poses some inherent shortcomings which need to be considered. To reduce magnetic resonance-measurement time in this clinical sample, we opted to use a simple bipolar scale to assess emotional valence. Although this procedure does not permit assessment of emotion-specific recognition accuracy, our previous findings show that 'errors' made by both healthy participants[5] and patients with depression[34] are rarely the result of confusion between emotional qualities, but more commonly, a failure to perceive emotion altogether, a confusion which is still possible on the reduced scale.

Averaging across different emotions may further have obscured important aspects of differential pathological processing. Given that this is the first study using dynamic multimodal stimulus material to compare groups, it was our specific goal to identify generalisable systematic group differences. Although the diversity of emotional situations increased the ecological validity of our stimulus battery, the dominance of negatively valenced stimuli over positively valenced ones may have induced a valence-specific bias, which needs to be evaluated by future studies.

Finally, although sample sizes of 20 participants have been reported to show sufficient power for fMRI studies,[65] adequate criticism on underpowered fMRI studies and low reliability[66,67] should be considered. Future studies with larger participant samples and *a priori* power estimations[68] are desirable to increase sensitivity to differences with smaller effect sizes.

Implications

Using a dynamic multimodal task to compare impairments in social cognition between patients with schizophrenia and patients with depression, we identified specific impairments that characterised each patient group. The schizophrenia group showed aberrant activation when speech content was neutral or foreign as a neural correlate of an inability to integrate ambiguous constructs. In the depression group, aberrant hypo- and hyperactivation was observed for multimodal emotional and neutral stimuli in mentalising regions. Differences in behavioural indices of empathy, or in neural activations in basic sensory processing areas were not observed. Taken together, these results demonstrate that comparisons of functional activation patterns during the perception of dynamic social stimuli constitute a promising and sensitive measure for the identification of pathology-specific problems in social cognition. Future research should disentangle the dynamic routes associated with the formation of empathy in order to elucidate the stage at which receiving, managing and communicating emotions goes awry in psychiatric disorders.

Funding

Christina Regenbogen, Dr. rer. medic., Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, and Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany; **Thilo Kellermann**, Dr. rer. medic., Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, and JARA Translational Brain Medicine, Jülich, Germany; **Janina Seubert**, Dr. rer. medic., Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; **Daniel A. Schneider**, MSc, Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, and JARA Translational Brain Medicine, Jülich, Germany; **Raquel E. Gur**, MD, PhD, Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania, USA; **Birgit Derntl**, Dr. rer. nat., Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen and JARA Translational Brain Medicine, Jülich, Germany; **Frank Schneider**, MD, Dr. rer. soc., Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, JARA Translational Brain Medicine, Jülich, Germany, and Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania, USA; **Ute Habel**, Dr. rer. soc., Department of Psychiatry, Psychotherapy, and Psychosomatics, Medical School, RWTH Aachen University, Aachen, and JARA Translational Brain Medicine, Jülich, Germany

Correspondence: Christina Regenbogen, Department of Clinical Neuroscience, Karolinska Institutet, Nobels väg 9, 17177 Stockholm, Sweden. Email: christina.regenbogen@ki.se

First received 9 Dec 2013, final revision 24 Jul 2014, accepted 6 Aug 2014

References

1 Batson CD. These things called empathy: eight related but distinct phenomena. In *The Social Neuroscience of Empathy* (eds J Decety, W Ickes): 3–16. MIT Press, 2009.

2 Frith CD, Frith U. Interacting minds: a biological basis. *Science* 1999; **286**: 1692–5.

3 Stein BE. Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp Brain Res* 1998; **123**: 124–35.

4 Seubert J, Kellermann T, Loughead J, Boers F, Brensinger C, Schneider F, et al. Processing of disgusted faces is facilitated by odor primes: a functional MRI study. *NeuroImage* 2010; **53**: 746–56.

5 Regenbogen C, Schneider DA, Finkelmeyer A, Kohn N, Derntl B, Kellermann T, et al. The differential contribution of facial expressions, prosody, and speech content to empathy. *Cogn Emot* 2012; **26**: 995–1014.

6 Regenbogen C, Schneider DA, Gur RE, Schneider F, Habel U, Kellermann T. Multimodal human communication: targeting facial expressions, speech content and prosody. *Neuroimage* 2012; **60**: 2346–56.

7 Smith MJ, Horan WP, Karpouzian TM, Abram SV, Cobia DJ, Csernansky JG. Self-reported empathy deficits are uniquely associated with poor functioning in schizophrenia. *Schizophr Res* 2012; **137**: 196–202.

8 Kanske P, Heissler J, Schonfelder S, Wessa M. Neural correlates of emotion regulation deficits in remitted depression: the influence of regulation strategy, habitual regulation use, and emotional valence. *Neuroimage* 2012; **61**: 686–93.

9 Kee KS, Green MF, Mintz J, Brekke, JS. Is emotion processing a predictor of functional outcome in schizophrenia? *Schizophr Bull* 2003; **29**: 487–97.

10 Hoekert M, Kahn RS, Pijnenborg M, Aleman A. Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. *Schizophr Res* 2007; **96**: 135–45.

11 Kohler CG, Walker JB, Martin EA, Healey KM, Moberg PJ. Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr Bull* 2010; **36**: 1009–19.

12 Bourke C, Douglas K, Porter R. Processing of facial emotion expression in major depression: a review. *Aust NZ J Psychiatry* 2010; **44**: 681–96.

13 Peron J, El Tamer S, Grandjean D, Leray E, Travers D, Drapier D, et al. Major depressive disorder skews the recognition of emotional prosody. *Prog Neuropsychopharmacol Biol Psychiatry* 2011; **35**: 987–96.

14 Leitman DI, Wolf DH, Laukka P, Ragland JD, Valdez JN, Turetsky BI, et al. Not pitch perfect: sensory contributions to affective communication impairment in schizophrenia. *Biol Psychiatry* 2011; **70**: 611–8.

15 Li H, Chan RC, McAlonan GM, Gong QY. Facial emotion processing in schizophrenia: a meta-analysis of functional neuroimaging data. *Schizophr Bull* 2010; **36**: 1029–39.

16 Chechko N, Augustin M, Zvyagintsev M, Schneider F, Habel U, Kellermann T. Brain circuitries involved in emotional interference task in major depression disorder. *J Affect Disord* 2013; **149**: 136–45.

17 Stuhrmann A, Suslow T, Dannlowski U. Facial emotion processing in major depression: a systematic review of neuroimaging findings. *Biol Mood Anxiety Disord* 2011; **1**: 10.

18 Falkenberg I, Kohn N, Schoepker R, Habel U. Mood induction in depressive patients: a comparative multidimensional approach. *PLoS One* 2012; **7**: e30016.

19 Mayberg, HS. Limbic-cortical dysregulation: a proposed model of depression. *J Neuropsychiatry Clin Neurosci* 1997; **9**: 471–81.

20 Thoma P, Zalewski I, von Reventlow HG, Norra C, Juckel G, Daum, I. Cognitive and affective empathy in depression linked to executive control. *Psychiatry Res* 2011; **189**: 373–8.

21 Tandon R, Gaebel W, Barch DM, Bustillo J, Gur RE, Heckers S, et al. Definition and description of schizophrenia in the DSM-5. *Schizophr Res* 2013; **150**: 3–10.

22 Kimhy D, Vakhrusheva J, Jobson-Ahmed L, Tarrier N, Malaspina D, Gross JJ. Emotion awareness and regulation in individuals with schizophrenia: implications for social functioning. *Psychiatry Res* 2012; **200**: 193–201.

23 Derntl B, Finkelmeyer A, Toygar TK, Hulsmann A, Schneider F, Falkenberg DI, et al. Generalized deficit in all core components of empathy in schizophrenia. *Schiz Res* 2009; **108**: 197–206.

24 Servan-Schreiber D, Cohen JD, Steingard S. Schizophrenic deficits in the processing of context. A test of a theoretical model. *Arch Gen Psychiatry* 1996; **53**: 1105–12.

25 Salisbury, D. N400 to lexical ambiguity and semantic incongruity in schizophrenia. *Int J Psychophysiol* 2010; **75**: 127–32.

26 Ravizza SM, Robertson LC, Carter CS, Nordahl TE, Salo RE. Is filtering difficulty the basis of attentional deficits in schizophrenia? *Psychiatry Res* 2007; **151**: 201–9.

27 Hall J, Whalley HC, Marwick K, McKirdy J, Sussmann J, Romaniuk L, et al. Hippocampal function in schizophrenia and bipolar disorder. *Psychol Med* 2010; **40**: 761–70.

28 Whalley HC, Papmeyer M, Sprooten E, Lawrie SM, Sussmann JE, McIntosh AM. Review of functional magnetic resonance imaging studies comparing bipolar disorder and schizophrenia. *Bipolar Disord* 2012; **14**: 411–31.

29 Seubert J, Loughead J, Kellermann T, Boers F, Brensinger CM, Habel U. Multisensory integration of emotionally valenced olfactory-visual information in patients with schizophrenia and healthy controls. *J Psychiatry Neurosci* 2010; **35**: 185–94.

30 Muller VI, Cieslik EC, Kellermann TS, Eickhoff SB. Crossmodal emotional integration in major depression. *Soc Cogn Affect Neurosci* 2014; **9**: 839–48.

31 Garrido-Vasquez P, Jessen S, Kotz SA. Perception of emotion in psychiatric disorders: on the possible role of task, dynamics, and multimodality. *Soc Neurosci* 2011; **6**: 515–36.

32 Tomlinson EK, Jones CA, Johnston RA, Meaden A, Wink B. Facial emotion recognition from moving and static point-light images in schizophrenia. *Schizophr Res* 2006; **85**: 96–105.

33 Kan Y, Mimura M, Kamijima K, Kawamura M. Recognition of emotion from moving facial and prosodic stimuli in depressed patients. *J Neurol Neurosurg Psychiatry* 2004; **75**: 1667–71.

34 Schneider D, Regenbogen C, Kellermann T, Finkelmeyer A, Kohn N, Derntl B, et al. Empathic behavioral and physiological responses to dynamic stimuli in depression. *Psychiatry Res* 2012; **200**: 294–305.

35 Freiherr J, Lundstrom JN, Habel U, Reetz K. Multisensory integration mechanisms during aging. *Front Hum Neurosci* 2013; **7**: 863.

36 Gottfried JA, Dolan RJ. The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 2003; **39**: 375–86.

37 Siegle GJ, Thompson W, Carter CS, Steinhauer SR, Thase ME. Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biol Psychiatry* 2007; **61**: 198–209.

38 Wittchen HU, Zaudig M, Fydrich T. *Strukturiertes klinisches Interview für DSM-IV: SKID (Structural Clinical Interview for DSMIV: SCID)*. Hogrefe verlag, 1997.

39 Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; **13**: 261–76.

40 Hautzinger M, Bailer M, Worall H, Keller F. *BDI Beck-Depressions-Inventar Testhandbuch* (2nd edn). Verlag Hans Huber, 1995.

41 Hamilton, M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23**: 56–62.

42 Oldfield, RC. The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 1971; **9**: 97–113.

43 Parker JD, Taylor GJ, Bagby RM. The 20-Item Toronto Alexithymia Scale. III. Reliability and factorial validity in a community population. *J Psychosom Res* 2003; **55**: 269–75.

44 Paulus C. *Der Saarbrücker Persönlichkeitsfragebogen (SPF-IRI) zur Messung von Empathie – V. 5.0.* 2006.

45 Leibetseder M, Laireiter AR, Riepler A, Köller T. E-Skala: Fragebogen zur erfassung von empathie – beschreibung und psychometrische eigenschaften [Questionnaire to assess empathy – description and psychometric properties]. *Z Differentielle Diagnostische Psychol* 2001; **22**: 70–85.

46 Ickes W. Empathic accuracy. *J Pers* 1993; **61**: 587–610.

47 de Vignemont F, Singer T. The empathic brain: how, when and why? *Trends Cogn Sci* 2006; **10**: 435–41.

48 Nuechterlein KH, Subotnik KL, Ventura J, Green MF, Gretchen-Doorly D, Asarnow RF. The puzzle of schizophrenia: tracking the core role of cognitive deficits. *Dev Psychopathol* 2012; **24**: 529–36.

49 Cox, RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 1996; **29**: 162–73.

50 Ethofer T, Van De Ville D, Scherer K, Vuilleumier P. Decoding of emotional information in voice-sensitive cortices. *Curr Biol* 2009; **19**: 1028–33.

51 Frühholz S, Ceravolo L, Grandjean D. Specific brain networks during explicit and implicit decoding of emotional prosody. *Cereb Cortex* 2012; **22**: 1107–17.

52 Kanwisher N. Domain specificity in face perception. *Nat Neurosci* 2000; **3**: 759–63.

53 Beaucousin V, Lacheret A, Turbelin MR, Morel M, Mazoyer B, Tzourio-Mazoyer N. FMRI study of emotional speech comprehension. *Cereb Cortex* 2007; **17**: 339–52.

54 Moberget T, Gullesen EH, Andersson S, Ivry RB, Endestad T. Generalized role for the cerebellum in encoding internal models: evidence from semantic processing. *J Neurosci* 2014; **34**: 2871–8.

55 Van den Stock J, de Jong SJ, Hodiamont PP, de Gelder B. Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Soc Neurosci* 2011; **6**: 537–47.

56 LeDoux, JE. Emotional memory systems in the brain. *Behav Brain Res* 1993; **58**: 69–79.

57 Sperling RA, Bates JF, Cocchiarella AJ, Schacter DL, Rosen BR, Albert MS. Encoding novel face-name associations: a functional MRI study. *Hum Brain Mapp* 2001; **14**: 129–39.

58 Grefkes C, Fink GR. The functional organization of the intraparietal sulcus in humans and monkeys. *J Anat* 2005; **207**: 3–17.

59 Haralanova E, Haralanov S, Beraldi A, Moller HJ, Hennig-Fast K. Subjective emotional over-arousal to neutral social scenes in paranoid schizophrenia. *Eur Arch Psychiatry Clin Neurosci* 2012; **262**: 59–68.

60 Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann N Y Acad Sci* 2010; **1191**: 62–88.

61 Lieberman, MD. Social cognitive neuroscience: a review of core processes. *Annu Rev Psychol* 2007; **58**: 259–89.

62 Cusi AM, Nazarov A, Holshausen K, Macqueen GM, McKinnon MC. Systematic review of the neural basis of social cognition in patients with mood disorders. *J Psychiatry Neurosci* 2012; **37**: 154–69.

63 Taylor GJ, Bagby RM, Parker JDA. The alexithymia construct. A potential paradigm for psychosomatic medicine. *Psychosomatics* 1991; **32**: 153–64.

64 Honkalampi K, Hintikka J, Laukkanen E, Lehtonen J, Viinamaki H. Alexithymia and depression: a prospective study of patients with major depressive disorder. *Psychosomatics* 2001; **42**: 229–34.

65 Thirion B, Pinel P, Tucholka A, Roche A, Ciuciu P, Mangin JF, et al. Structural analysis of fMRI data revisited: improving the sensitivity and reliability of fMRI group studies. *IEEE Trans Med Imaging* 2007; **26**: 1256–69.

66 Brandt DJ, Sommer J, Krach S, Bedenbender J, Kircher T, Paulus FM, et al. Test-retest reliability of fMRI brain activity during memory encoding. *Front Psychiatry* 2013; **4**: 163.

67 Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013; **14**: 365–76.

68 Mumford JA. A power calculation guide for fMRI studies. *Soc Cogn Affect Neurosci* 2012; **7**: 738–42.

EXTRA CONTENT ONLINE