Neural correlates of moral reasoning in autism spectrum disorder

Karla Schneider,1,2 Katharina D. Pauly,1,2 Anna Gossen,1,2 Lea Mevissen,1 Tanja M. Michel,1,4 Ruben C. Gur,3 Frank Schneider,1,2,4 and Ute Habel1,2

1Department of Psychiatry, Psychotherapy, and Psychosomatic Medicine, Medical School, RWTH Aachen University, 52074 Aachen, Germany, 2JARA-Brain Jülich Aachen Research Alliance, Translational Brain Medicine, Forschungszentrum Jülich GmbH, 3Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia Veterans Administration Medical Center, and 4Department of Psychiatry and Psychotherapy, Medical School, University of Rostock, 18147 Rostock

In our study, we tried to clarify whether patients with autism spectrum disorder (ASD) reveal different moral decision patterns as compared to healthy subjects and whether common social interaction difficulties in ASD are reflected in altered brain activation during different aspects of moral reasoning. 28 patients with high-functioning ASD and 28 healthy subjects matched for gender, age and education took part in an event-related functional magnetic resonance imaging study. Participants were confronted with textual dilemma situations followed by proposed solutions to which they could agree or disagree. On a neural level, moral decision making was associated with activation in anterior medial prefrontal regions, the temporo-parietal junction and the precuneus for both groups. However, while patients and healthy controls did not exhibit significant behavioral differences, ASD patients showed decreased activation in limbic regions, particularly the amygdala, as well as increased activation in the anterior and the posterior cingulate gyrus during moral reasoning. Alterations of brain activation in patients might thus indicate specific impairments in empathy. However, activation increases in brain regions associated with the ‘default mode network’ and self-referential cognition also provide evidence for an altered way of patients’ cerebral processing with regard to decision making based on social information.

Keywords: autism spectrum disorder; decision making; morality; neuroimaging; fMRI

INTRODUCTION

Although it is a relatively new field of research, already a wealth of studies approaching behavioral characteristics and neural underpinnings of moral cognition has accumulated (e.g. Moll et al., 2002; Schleim et al., 2010; Bzdok et al., in press). In solving moral tasks, the ventromedial prefrontal cortex (VMPFC), lateral orbitofrontal cortex, temporal and limbic regions seem to play a key role (Moll et al., 2002; Hiraishi et al., 2007; Young et al., 2007; Zahn et al., 2009; Schleim et al., 2010).

Additionally, a recent study (Young et al., 2010) could show a reduced capacity of moral decision making based on the comprehension of the mental states of other individuals, especially in case of attempted harms, if the right temporo-parietal junction (TPJ) was temporarily inhibited by transcranial magnetic stimulation in healthy participants.

Moral reasoning as a sophisticated element of social cognition requires a broad understanding of social contexts. Correspondingly, mental disorders, which are related to deficits in empathy and Theory of Mind (ToM), might interfere with adequate moral decision making (de Oliveira-Souza et al., 2008). For autism spectrum disorders (ASDs), this linkage was already suggested in a few behavioral studies (Blair, 1996; Moran et al., 2011; Zalla et al., 2011; Shulman et al., 2012) and—to the authors’ knowledge— in only one (case) study using functional magnetic resonance imaging (fMRI; Hiraishi et al., 2007). Particularly, Moran et al. (2011) revealed subtle weaknesses in moral decisions in patients with ASD with regard to the ability to differentiate between intentional and attempted harms. Likewise, Zalla et al. (2011) reported impairments of ASD patients with regard to provide welfare-based moral justifications.

According to the main diagnostic classification systems of ICD-10 and DSM-IV, ASDs as pervasive developmental disorders are characterized by the three core features of deficits in communication, repetitive stereotyped behavior and impaired social interaction.

While the literature is indecisive regarding the question of global ToM- and empathy-deficits in ASD (Frith and Happé, 1994; Castelli, 2005; Silani et al., 2008; Senju et al., 2009; Bird et al., 2010; Chevallier, 2012), there seems to be substantial convergence that also high-functioning adults with ASD exhibit at least subtle deficits in empathy (Schulte-Rüther et al., 2010; Zalla et al., 2011) and/or ToM (Kana et al., 2009; Senju et al., 2009; Moran et al., 2011), which also seem to affect patients’ ability to make adequate moral decisions (Hiraishi et al., 2007; Moran et al., 2011; Zalla et al., 2011; Shulman et al., 2012).

Furthermore, most of the recent studies examining morality-associated aspects of social cognition, such as ToM (Kana et al., 2009) and empathy (Schulte-Rüther et al., 2010), revealed dysfunctions in respective brain networks in ASD patients, which have been related to patients’ social impairments (Lombardo et al., 2011).

Anatomical and functional alterations were mainly found in the prefrontal cortex (PFC), particularly the medial PFC and in the TPJ (Kana et al., 2009; Martineau et al., 2010; Schulte-Rüther et al., 2010; Lombardo et al., 2011). Common dysfunctions in these areas in ASD might well come along with deficits in moral reasoning (Blair, 1996; Moran et al., 2011; Zalla et al., 2011; Shulman et al., 2012), and hence result in altered brain activation during moral tasks.

Nonetheless, so far studies on moral understanding in clinical samples mainly focused on patients with psychopathy (e.g. de Oliveira-Souza et al., 2008), who were said to exhibit a contrariwise profile of empathy deficits as compared to patients with ASD, namely intact cognitive elements, but impairments in affective aspects of empathy (Blair, 2008; Krippi and Karim, 2011). However, at the same time potential overlaps between Asperger’s syndrome and sociopathy/psychopathy were suggested (Anckarsäter et al., 2008). More, though there is no evidence for a general overrepresentation of patients with
ASD in forensic settings, a first long-term study on deviant behavior in pervasive developmental disorders (Mouridsen et al., 2008) revealed a tendency in patients with ASD to particular criminal delinquency, such as arson (Haskins and Silva, 2006; Mouridsen et al., 2008, in line with Allen et al., 2008; Anckarsäter et al., 2008). Yet, the underlying factors and potential relationships are still unclear, but might be partly related to changes in moral reasoning and its underlying neural networks.

Hence, a better understanding of moral reasoning in patients with ASD could provide helpful new approaches for specific therapeutic interventions and was the underlying reason for our exploration.

In order to investigate different specific aspects of moral reasoning with potentially different empathic requirements (Greene et al., 2004), we developed a set of dilemmas comprising two moral categories: social–ethical dilemmas (SE) and individual gain vs collective losses (IND) dilemmas.

Behaviorally, we expected longer reaction times in ASD patients due to difficulties in executive functions such as decision making (Channon et al., 2001) and in information processing (O’Connor and Kirk, 2008).

There is no prior fMRI research on ASD and moral decision making. However, given the widely suggested impairments of patients with ASD in ToM and empathy and the proposed relation of morality and both of these aspects of social cognition, we expected that patients would reveal decreased brain activation, mainly in the medial PFC and the TPJ, during moral decision making and additionally reduced activation in emotion-related areas (e.g. amygdala). As we hence hypothesized that conceivable empathy and ToM deficits in patients with ASD would affect their ability of moral decision making, we expected this effect to be strongest in the case of the more personal dilemmas of the IND condition (see also Greene and Haidt, 2002; Greene et al., 2004) which particularly require to (theoretically) put oneself into direct confrontation with another subject (Zalla et al., 2011).

**MATERIALS AND METHODS**

**Participants**

Thirty patients (17 males) with high-functioning ASD were recruited at the in- and outpatient facilities of the RWTH Aachen University Hospital and at local self-help groups and therapy centers. In line with recent studies (Via et al., 2011) and with the upcoming DSM-V and ICD-11, we refer to ‘autism spectrum disorders’ without distinguishing between different autistic subgroups. Patients were either diagnosed at a special autism facility at the Department of Psychiatry, Psychotherapy and Psychosomatic Medicine of the RWTH Aachen University Hospital (n = 9) or had already been diagnosed with ASD by an experienced psychiatrist elsewhere (n = 21).

Participants had to meet the following criteria for inclusion: age between 18 and 55 years, no neurological disorders, no MRI contraindications (metal implants, tattoos, pregnancy, etc.). Patients with current psychiatric comorbidities were excluded. All participants were native German speakers and screened for mental disorders by means of the Structured Clinical Interview for DSM-IV (SCID; Wittchen et al., 1997). fMRI data of two male patients were excluded afterward due to movement artifacts (n = 1) or the fact that the diagnosis of Asperger’s syndrome (n = 1) could not be confirmed. One male and one female healthy participant had to be excluded due to an incidental finding of brain abnormality and due to movement artifacts, respectively. Thus, 28 autistic patients were included in the final analysis as well as 28 gender-, age- and education-matched healthy controls (HCs) (see Table 1). Four of the ASD patients and two healthy subjects were left-handed. All of the other participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Five patients were medicated with selective serotonin reuptake inhibitors (SSRIs), one with a tricyclic antidepressant (see Supplementary Table S1).

**Table 1 Demographical, neuropsychological and psychopathological data**

<table>
<thead>
<tr>
<th></th>
<th>HC  Mean (±s.d.)</th>
<th>ASD Mean (±s.d.)</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>n = 28/28</td>
<td>31.29 (±9.03)</td>
<td>31.39 (±8.97)</td>
<td>–0.05</td>
<td>54</td>
</tr>
<tr>
<td>Education (years)</td>
<td>n = 28/28</td>
<td>12.89 (±0.32)</td>
<td>12.68 (±0.82)</td>
<td>1.29</td>
<td>54</td>
</tr>
<tr>
<td>IQ</td>
<td>n = 28/27</td>
<td>114.04 (±9.55)</td>
<td>109.07 (±9.12)</td>
<td>1.97</td>
<td>53</td>
</tr>
<tr>
<td>Digit span forward</td>
<td>n = 27/19</td>
<td>9.30 (±1.44)</td>
<td>9.05 (±2.09)</td>
<td>0.47</td>
<td>44</td>
</tr>
<tr>
<td>Digit span backward</td>
<td>n = 27/19</td>
<td>9.93 (±7.32)</td>
<td>8.32 (±2.40)</td>
<td>0.92</td>
<td>44</td>
</tr>
<tr>
<td>Lexical fluency (n = 26/19)</td>
<td>16.65 (±5.34)</td>
<td>16.42 (±6.56)</td>
<td>0.13</td>
<td>43</td>
<td>0.896</td>
</tr>
<tr>
<td>Lexical flexibility (n = 26/19)</td>
<td>16.62 (±4.88)</td>
<td>16.84 (±4.18)</td>
<td>–0.16</td>
<td>43</td>
<td>0.871</td>
</tr>
<tr>
<td>Semantic fluency (n = 25/19)</td>
<td>17.68 (±6.41)</td>
<td>18.05 (±5.76)</td>
<td>–0.24</td>
<td>42</td>
<td>0.813</td>
</tr>
<tr>
<td>Semantic flexibility (n = 25/19)</td>
<td>17.28 (±2.85)</td>
<td>17.67 (±4.46)</td>
<td>–0.35</td>
<td>41</td>
<td>0.730</td>
</tr>
<tr>
<td>Difference score TMT-B–TMT-A (s)</td>
<td>14.69 (±8.67)</td>
<td>18.05 (±13.50)</td>
<td>–1.02</td>
<td>43</td>
<td>0.320</td>
</tr>
<tr>
<td>AQ (n = 28/26)</td>
<td>10.07 (±5.93)</td>
<td>37.69 (±7.85)</td>
<td>–14.66</td>
<td>52</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>TAS20 (n = 28/25)</td>
<td>38.93 (±10.10)</td>
<td>59.92 (±11.57)</td>
<td>–7.00</td>
<td>51</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>NSAD (n = 27/26)</td>
<td>4.75 (±2.50)</td>
<td>4.36 (±2.02)</td>
<td>1.20</td>
<td>54</td>
<td>0.241</td>
</tr>
</tbody>
</table>

Independent two-sample t-test: P < 0.05, Bonferroni-corrected: *P < 0.004.

HC = healthy control subjects; ASD = patients with autism spectrum disorders; s.d. = standard deviation; AQ = autism quotient; NSAD = scale for socially desirable behavior; TAS20 = Toronto Alexithymia Scale; TMT-A/B = trail making test.

ASD traits of patients were assessed using the Autism Diagnostic Observation Schedule—Generic (ADOS-G, Module 4; Lord et al., 1999). ASD patients had an averaged ADOS-G score of 7.26 (±4.35; see Supplementary Table S2). Since the sensitivity of the ADOS-G in its current version has recently been questioned for ASD in adults (Bastaansen et al., 2011; Lai et al., 2011), we also included patients who did not meet the ADOS criteria for ASD (cut-off = 7). However, all patients fulfilled the cut-off for ASD according to the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001; for further details see Section 2 in the Supplementary Data and Supplementary Table S2a and S2b). The Hamilton Depression Scale (Hamilton, 1960) revealed no current depressivity in ASD patients with a mean score of 2.94 (±3.85).

Participants further completed the Toronto Alexithymia Scale (TAS20; Bagby et al., 1994). In order to control for variance due to potential cognitive performance differences, a neuropsychological test battery was applied encompassing a German vocabulary test for crystalline intelligence estimation (‘Wortschatztest’—WST; Schmidt and Metzler, 1992), a test on lexical and semantic verbal fluency (‘Regensburger Wortflüssigkeitstest’—RWT; Aschenbrenner et al., 2001), the German version of the digit span of the Wechsler Intelligence Scale for Adults (WIE; Von Aster et al., 2006) and the Trail Making Test (TMT-A/B; Reitan, 1958). Patients’ emotion recognition performance was tested using the Penn Emotion Recognition Test (PERT40; Kohler et al., 2004). Finally, the tendency for social desirable behavior was assessed (‘Deutsche Kurzskala zur Erfassung des Bedürfnisses nach sozialer Anerkennung’; NSAd; Stocké, 2009).

For an overview of the demographical and clinical data, see Table 1 (for additional data on the applied personality tests see Supplementary Tables S3 and S4).

After complete description of the study, written informed consent of each participant was obtained. All participants were paid for their participation after completing the study. The local Institutional Review Board of the Medical Faculty of RWTH Aachen University approved the protocol and the design fulfilled the guidelines of the Declaration of Helsinki (2008).

**Task**

Participants were asked to answer 60 short textual dilemma situations depicting close-to-reality situations, among them 20 social-ethical
Table 2 Dilemma examples for each moral condition

<table>
<thead>
<tr>
<th>Social-ethical (SE)</th>
<th>Individual gain vs collective losses (IND)</th>
<th>High-level baseline (BL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>You want to have a child. Preferably, you would like to have a girl, who is slim, blond and healthy. Genetic testing would allow you to control for these criteria.</td>
<td>You are currently in a professional probation period. You make a potentially job-endangering mistake, though it was not seen by anyone. It could have happened to anyone in the office. 'I do not intervene when a co-worker is blamed.'</td>
<td>You come home tired. Since your physician recommended more sporting activity to compensate for your office job, you applied for an aerobic course. However, today you are not in the mood. 'I go despite being tired.'</td>
</tr>
</tbody>
</table>

Functional data covered the whole brain and were obtained with an echo-planar imaging (EPI) sequence sensitive to the blood oxygenation level dependent (BOLD) effect (T2*, TR/TE = 2000/30 ms, FoV = 200 × 200 mm2, matrix = 64 × 64, 36 slices, alpha = 76°, voxel size: 3.1 × 3.1 × 3.1 mm3, slice thickness = 3.1 mm, gap = 15%). For each participant, 609 volumes were acquired. This resulted in a scanning time of 25 min for the moral task and structural measurement. The first three images were discarded to account for T1 stabilization effects.

Data analysis

Functional data were processed and analyzed using statistical parametric mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab 11 (The Mathworks, Inc., Natick, MA, USA). The 606 volumes of each participant were aligned to the mean image. Realigned images were normalized to the standard anatomical Montreal Neurological Institute (MNI) coordinate space resulting in a voxel size of 3 × 3 × 3 mm3 and smoothed with a Gaussian kernel of 8 mm full-width-at-half-maximum. Subsequently, a 1/128 Hz high-pass filter removed low frequency noise.

For each subject, we determined the three trial types of the SE, the IND and the BL condition for the whole duration of each trial. These contrast effects were taken to the second level. Applying the general linear model (GLM), a flexible factorial analysis was performed contrasting SE vs BL, IND vs BL and both moral conditions pooled vs (weighted) BL contrasts for each group separately as well as comparing the ASD and the HC group.

In order to prevent confounds with alexithymia (Bird et al., 2010), we correlated brain activation in patients and HCs during the Moral vs BL contrast with the TAS20 score. Moreover, due to the inter-individual variability regarding the ADOS-G Module 4 criteria for ASD, we additionally correlated brain activation with the overall ADOS score in order to control for misinterpretations due to diagnostic differences.

One-sample analyses were family wise error (FWE) corrected (P < 0.05) for multiple comparisons with a threshold extent of 20 voxels. For group comparisons (representing a less robust difference of a difference) and multiple regression analyses calculated separately for patients and non-autistic subjects, a more liberal Monte-Carlo corrected threshold was chosen. Monte-Carlo simulations were computed using AlphaSim by Ward (2000) implemented in AFNI 2011 (Cox, 2012).

Assuming a per voxel probability threshold of P = 0.001, after 1,000 simulations a cluster size of 18 contiguous resampled voxels was indicated to correct for multiple comparisons at P < 0.05.

As this is the very first study on moral reasoning in ASD, there is no prior data to perform adequate regions of interest analyses (Poldrack and Mumford, 2009).

Anatomical localization was performed using the WFU Pick Atlas implemented as a tool in SPM 8.

Analysis of the behavioral data

Behavioral data were analyzed using the IBM® Statistical Package for Social Sciences (SPSS) 20. As Kolmogorov-Smirnov tests demonstrated normal distribution of the reaction times during the fMRI task (SE condition: Z = 0.76, P = 0.62; IND condition: Z = 1.10, P = 0.18; BL condition: Z = 0.51, P = 0.96), a 2 × 3 repeated measures ANOVA was performed with the between-subject factor group (ASD and healthy subjects) and the within-subject repeated measures factor condition (SE, IND, BL). Non-parametric χ2 tests were applied to analyze the decisions made. For further group comparisons, we used t-tests for independent samples (P < 0.05), if not otherwise specified.

Image acquisition

Images were acquired on a 3T Siemens Trio MR scanner at the Department of Psychiatry, Psychotherapy and Psychosomatic Medicine, RWTH Aachen University. A 4 min magnetization-prepared rapid acquisition gradient echo image (MP-RAGE) T1-weighted sequence was used to acquire structural images (TR = 1900 ms, TE = 2.52 ms, TI = 900 ms, matrix = 256 × 256, 176 slices, FoV: 250 × 250 mm2, alpha = 9°, voxel size = 1 × 1 × 1 mm3).

Stimuli were standardized according to the following criteria:

1. Sentence structure of each dilemma: main clause beginning with ‘You’ and a verb (e.g. ‘You want to have a child.’)—main clause plus subordinate clause (e.g. ‘Preferably, you would like to have a girl who is slim, blond, and healthy.’)—main clause (e.g. ‘Genetic testing would allow you to control for these criteria.’)—proposed solution: main clause (e.g. ‘I do not decide on genetic testing.’);
2. Number of words: 35–40, among them 5–7 for the potential solution;
3. Solution sentence beginning with ‘I’ and a verb (for examples of each category—translated from German into English—see Table 2).

Stimuli were validated in a preceding pilot study including 31 healthy participants (for further information see Supplementary Figure S1). The two moral and the high-level BL (control) conditions did not differ significantly (Kruskal–Wallis tests) concerning realism [χ2(2) = 1.96, P = 0.38], the intensity of the perceived dilemma [χ2(2) = 4.17, P = 0.12] and how easy it was to put oneself into each situation [χ2(2) = 1.91, P = 0.39], respectively. There was only a significant difference for emotional intensity due to a weaker emotional involvement during the high-level BL dilemmas [χ2(2) = 26.89, P < 0.01].

For the target answer, we used Kohlberg’s (Kohlberg et al., 1983) ‘moral pyramid’ for orientation, and according to age and education of our participants set the target level to level 4 of the pyramid. This level specifies a sophisticated ability to ponder one’s own interests and the interests of other people and society.

Stimuli were presented in an event-related design in a pseudo-randomized way using Presentation® 9.0 software (Neurobehavioral Systems Inc., San Francisco, CA, USA). Each dilemma text was presented for 13.5 s. After 9 s, a potential solution of the dilemma was presented underneath the text for 4.5 s and participants were asked to agree or disagree with the solution via left vs right button press using their right index or middle finger. Each dilemma was followed by a fixation cross with a jittered length of 5–7 s. Subjects also performed an empathy task described elsewhere.

The first three images were discarded to account for T1 stabilization effects.
RESULTS

Behavioral data

The 2 × 3 repeated measures ANOVA analyzing reaction times (Figure 1) yielded a significant main effect of condition $F(2,108) = 6.02$, $P < 0.01$. The main effect of group marginally failed to reach significance $F(2,108) = 3.25$, $P = 0.08$. No significant group × condition interaction was found $F(2,108) = 0.98$, $P = 0.38$.

Post hoc paired t-tests (P < 0.02, Bonferroni-corrected) on mean reaction times revealed a significant difference between the SE and BL condition ($t(55) = 3.21$, $P < 0.01$) as well as between the IND and the BL condition ($t(55) = 2.97$, $P < 0.01$), with longer reaction times for the moral conditions, respectively, but no differences between the SE and IND condition ($t(55) = 0.67$, $P = 0.51$).

There was no significant difference between the two groups with regard to the answering pattern in terms of the Kohlberg-targets (SE and IND conditions as compared to the BL). Group differences marginally failed to reach significance. SE and the IND conditions as compared to the BL. Group differences marginally failed to reach significance. SE = social–ethical dilemmas; IND = individual gain vs altruistic losses; BL = non-moral daily dilemma high-level baseline; RT = Reaction times; s = seconds.

Imaging data

One-sample analyses

Pooling both moral conditions revealed activation patterns in the HC group bilaterally in the TPJ region, the medial frontal gyrus and the precuneus (Table 3 and Figure 2).

In ASD patients, the same contrast yielded activation in the medial frontal gyr, the middle frontal gyrus and the precuneus bilaterally, in the left TPJ, the left inferior temporal gyrus and in the right supramarginal gyrus (Table 3 and Figure 2).

Comparing the SE dilemmas with the IND condition revealed a significant signal change in the posterior cingulate gyrus for the HC and the ASD group (Table 3).

The opposite contrast of IND vs SE yielded no significant signal change in the ASD group, while the HC group showed stronger activation in the left TPJ region, the left cerebellar declive, the right inferior frontal gyrus (IFG), the right middle temporal gyrus and the right lingual gyrus (Table 3).

Group comparisons

For the pooled Moral vs BL contrast, increased brain activation in ASD patients compared to non-autistic individuals was found in the posterior cingulate cortex bilaterally and in the right ventral-part of the ACC (Table 4).

For the separate analysis of dilemma processing and solution finding see Supplementary Table S5 and Supplementary Figure S2.

The group comparison on the SE vs BL contrast yielded hyperactivation in ASD patients compared to the HC group in the bilateral ACC, the left posterior cingulate gyrus, the right precuneus and in the right supramarginal gyrus (Table 4 and Figure 3A, B).

Table 3 Brain activation for all moral contrasts—one-sample analyses

<table>
<thead>
<tr>
<th>Group(s)</th>
<th>Region</th>
<th>BA</th>
<th>Side</th>
<th>k</th>
<th>t</th>
<th>MNI-coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral vs BL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS D vs BL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC vs BL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC vs AS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IND vs BL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC vs IND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IND vs SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comparing the HC and the ASD group regarding the IND vs BL contrast revealed reduced activation in the ASD group in the left amygdala and in the left IFG (Table 4 and Figure 3C, D).

The IND vs SE contrast revealed decreased activation for ASD patients compared to healthy subjects in the left amygdala, in the left insula and in the left posterior cingulate cortex (Table 4 and Figure 3E, F).

The SE vs IND contrast yielded no significant BOLD signal change (Table 4).

Group comparisons without medicated patients are found in Supplementary Table S1.

**Correlation and covariate analyses**

The correlation with the TAS20 and overall ADOS-G (Module 4) score yielded no significant results.

**DISCUSSION**

The aim of this study was to investigate the capacity of moral reasoning and its neural correlates in patients with ASD compared to non-autistic subjects by applying two different moral categories in order to determine potential specific deficits in moral decision making.

Despite a lack of behavioral differences between ASD patients and HC subjects, and despite comparable brain activation patterns for the single-group analyses, patients in comparison to HC subjects, as hypothesized, exhibited significant hypoactivations in the left amygdala, the left insula and the left IFG, for the individual gain vs collective losses dilemmas in particular (Figure 3C, D). On the other hand, we found increased activation in ASD patients in posterior regions, including the precuneus and particularly the posterior cingulate gyrus (Figure 3A, B), but also in the ACC, when comparing the ASD and the control group for the SE and pooled moral dilemmas.

These results of specific hypo- and hyperactivations in patients compared to healthy subjects might be interpreted in the light of different underlying phenomena.

**Morality and empathy deficits in patients with ASD—differential activation during IND and SE dilemmas**

For the IND vs BL and IND vs SE contrast (Figure 3C, F), we found decreased activation in ASD patients compared to healthy subjects in emotion- and empathy-related areas (Lamm et al., 2007; Schulte-Rüther et al., 2010) including areas of the human mirror neuron system (Schulte-Rüther et al., 2010), namely in the IFG, and of the limbic system, especially the amygdala, which is distinctively depicted by the parameter estimates (Figure 3D, F).

---

**Table 4** Brain activation for all moral contrasts—group comparisons

<table>
<thead>
<tr>
<th>Group(s)</th>
<th>Region</th>
<th>BA</th>
<th>Side</th>
<th>k</th>
<th>t</th>
<th>MNI-coordinates (x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral vs BL</td>
<td>HC &gt; ASD</td>
<td>N.S.</td>
<td>Anterior cingulate cortex</td>
<td>R</td>
<td>23</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>ASD &gt; HC</td>
<td></td>
<td>Posterior cingulate gyrus</td>
<td>BA31</td>
<td>R</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Posterior cingulate gyrus</td>
<td>L</td>
<td>21</td>
<td>3.65</td>
</tr>
<tr>
<td>SE vs BL</td>
<td>HC &gt; ASD</td>
<td>N.S.</td>
<td>Anterior cingulate cortex</td>
<td>BA32</td>
<td>L</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>ASD &gt; HC</td>
<td></td>
<td>Anterior cingulate cortex</td>
<td>L + R</td>
<td>22</td>
<td>3.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Posterior cingulate gyrus</td>
<td>L</td>
<td>39</td>
<td>4.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Precuneus</td>
<td>BA7</td>
<td>R</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Supramarginal gyrus</td>
<td>L</td>
<td>23</td>
<td>3.87</td>
</tr>
<tr>
<td>IND vs BL</td>
<td>HC &gt; ASD</td>
<td></td>
<td>Amygdala</td>
<td>L</td>
<td>22</td>
<td>4.51</td>
</tr>
<tr>
<td></td>
<td>ASD &gt; HC</td>
<td>N.S.</td>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>26</td>
<td>4.47</td>
</tr>
<tr>
<td>SE vs IND</td>
<td>HC &gt; ASD</td>
<td>N.S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IND vs SE</td>
<td>HC &gt; ASD</td>
<td></td>
<td>Amygdala</td>
<td>L</td>
<td>30</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Insula</td>
<td>L</td>
<td>27</td>
<td>4.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Posterior cingulate cortex</td>
<td>L</td>
<td>22</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Whole-brain analysis; flexible-factorial ANOVA; threshold = P < 0.05 Monte-Carlo corrected; Moral = combined SE and IND dilemmas; SE = social-ethical dilemmas; IND = individual gain vs collective losses dilemmas; BL = (weighted) high-level baseline; ASD = patients with autism spectrum disorder; HC = Healthy controls; L = left; R = right; BA = Brodmann's Area; N.S. = not significant.

---
Our two-person-situations pictured by the IND dilemmas (Figure 3C, D) in contrary to the SE condition (Figure 3A, B) not only ask for a certain ability to put oneself into context with another individual, but also to feel with the acting persons. Though general empathy deficits in ASD are questioned (Silani et al., 2008; Bird et al., 2010), the hypoactivations in ASD patients compared to healthy subjects particularly during the IND condition of our study might hence point to specific alterations in underlying networks of empathy. This interpretation is in line with the findings of Zalla et al. (2011) in patients with ASD suggesting difficulties during solving those moral dilemmas which require knowledge about intentions of the agent and about the affective outcome of an action.

Also, in contrary to Bird et al. (2010), who suggested empathy deficits in ASD patients as an effect of alexithymia rather than ASD itself, our results were not affected by the score of the alexithymia scale TAS20.

Hypoactivations of empathy-related regions in patients compared to healthy subjects during our study were restricted to the left hemisphere. Unlike the right hemisphere, which predominantly seems to cover basal aspects of empathic emotions such as pain empathy, the left hemisphere is proposed to be involved in the emotional evaluation of more sophisticated social-cognitive input (Bzdok et al., 2012).

Hence, while there might not be a general empathic deficit based on ASD during basal emotion processing, our findings might indicate ASD-specific impairments in emotional involvement during processing of higher-level social situations, as depicted by our moral (IND) dilemmas.

**Morbidity and executive dysfunctions**

Interestingly, the network of brain regions, which we found comparably activated for both groups in the one-sample analyses of all moral contrasts (Figure 2), is also part of the ‘default mode network’, e.g. of regions activated during the so-called ‘brain resting state’.

For the SE condition and pooled moral contrast, group comparisons revealed abnormal hyperactivations for ASD patients in the posterior cingulate as a crucial part of this default network.

A recent study (Kennedy et al., 2006) found that, unlike healthy subjects, ASD patients did not show reduced activation in default network regions comprising the VMPFC, the posterior cingulate gyrus and the precuneus during cognitively demanding tasks (see...
also Wirth et al., 2011). Likewise, an abnormal functional connectivity within default network regions was reported (Assaf et al., 2010), especially between the medial PFC/ACC and the precuneus in ASD patients.

The ‘default mode network’ is considerably overlapping with regions associated with self-referential (Buckner and Carroll, 2006; see also below) and emotional (Vogt et al., 1992) processing. Particularly, the posterior cingulate cortex also seems to subserve evaluative functions including self-reflection (Vogt et al., 1992). Thus, alterations during ‘resting state’ might also indicate a putatively modified way of processing and evaluating (social) input, and consequently were suggested to be the underlying reason for social impairments of ASD patients (Kennedy et al., 2006).

The SE condition representing complex societal issues particularly required integrating prior personal knowledge order to adequately solve our moral task, and hence, our findings might point to altered processing of social information in ASD patients.

Behavioral studies revealed impairments of ASD patients in different aspects of information processing (Happé and Frith, 2006; O’Connor and Kirk, 2008), as well as executive dysfunctions (Hill, 2004), including impaired decision making (Channon et al., 2001; Schmitz et al., 2007). With it, difficulties in the processing of social information were proposed as the underlying factor of the social impairments in ASD (O’Connor and Kirk, 2008).

Though our sample of ASD patients did not significantly differ regarding the response patterns during moral reasoning, as compared to healthy participants, they tended to exhibit longer reaction times, especially during the SE condition, which might point to patients’ altered social information processing and/or to executive dysfunctions including decision making.

Remarkably, we also found abnormal hyperactivation in our patients compared to non-autistic subjects in the ACC (area BA32 in particular), during the Moral vs BL and SE vs BL (Figure 3A, B) contrast. Hyperactivations of the ACC as a crucial area for conflict and error monitoring and decision making (Bush et al., 2000), and of its adjacent regions (Dichter et al., 2009), have been interpreted as compensatory mechanisms of ASD patients.

The idea of a ‘failure to deactivate’ during cognitively demanding tasks in ASD patients compared to control subjects in context with ACC hyperactivation might thus be a hint for an altered and putatively less efficient way of solving social situations. These deficits in information processing might aggravate social impairments and also be related to the increased ACC activation indicating the need for increased effort to solve complex social situations.

However, the distinct hypoactivations of empathy-related regions, during the IND vs BL and IND vs SE contrasts in particular (Figure 3C–F), described earlier, suggest that it is unlikely that the differential brain activation is only linked to cognitive deficits as executive dysfunctions or impaired reading comprehension.

Additionally, in contrary to the widely suggested deficits on the level of executive functions including verbal fluency in autistic patients (Turner, 1999), and despite patients’ tendency to longer reaction times during our moral task, our autistic participants did not differ significantly from HCs (Table 1) on the respective neuropsychological tests, namely the Trail Making Test (TMT-A/-B) and the test for verbal fluency (RWT). Adequate performance of ASD patients during these tests is also in line with Spek et al. (2009), who found verbal fluency impairments in patients with Asperger’s syndrome only in one specific category of the task, namely the putatively most ‘social’ subscale ‘professions’, while all of the others were preserved.

Likewise, Leinhardt et al. (2011) found no significant performance differences in high-functioning autistic adults during the TMT, and interpreted this finding with the heterogeneity of the autistic spectrum. Not least, our high-functioning ASD group revealed a comparatively high educational and intellectual level, which might explain the rather preserved performance.

**Moral decision making and altered self-reflection**

A striking aspect of the one-sample contrasts (Figure 2) is the involvement of the particular network of prefrontal and posterior cortical midline structures also active during tasks investigating self-reflection (Johnson et al., 2006; for a review: Buckner and Carroll, 2006) or self-other distinction (Uddin et al., 2006). During group comparisons, we found abnormal hyperactivations in ASD patients especially in the posterior cingulate gyrus for the Moral vs BL and the SE vs BL contrast, also highlighted by the parameter estimates (Figure 3B for the SE vs BL contrast).

Several authors propose a close linkage between social or moral judgments and the understanding of others’ and one’s own mind (di Martino and Castellanos, 2003; Farley et al., 2010; Zahavi, 2010). Consequently, tasks testing for social cognition and self-reflection result in similar brain activation patterns comprising the VMPFC, and posterior regions, such as the precuneus (Young et al., 2007). Hence, social judgments, as reflected in our moral dilemmas, need a mature understanding of one’s own inner life.

At the same time, behavioral data suggest impairments of ASD patients in different aspects of self-awareness, particularly with regard to agency (Farley et al., 2010), and thus relate ASD to a disorder of the ‘self’ (Dritschel et al., 2010; Zahavi, 2010).

The posterior cingulate cortex is suggested to be crucial for experimental self-reflection and here for the processing of outward-directed, i.e. social and contextual, elements of an event (Johnson et al., 2006). Given a ‘weak central coherence’ in ASD patients (Happe and Frith, 2006) and our above suggestions of altered executive functions, they might need to put more effort in solving tasks which require combination of complex information. As the SEs in particular ask for integrating several aspects, i.e. on political decisions, and hence are related to the processing of the ‘outward-directed’ elements of an event, this might find its neural correlate in the most pronounced hyperactivation in ASD patients compared to non-autistic subjects (Figure 3A, B).

**LIMITATIONS AND CONCLUSIONS**

Our study has some limitations: Our stimuli were not designed to detect potential differences between ASD patients and HCs with regard to underlying motivations and beliefs behind a chosen solution during moral decision making, e.g. attempted vs intentional harms (Moran et al., 2011; Zalla et al., 2011; Shulman et al., 2012).

Neither did we include a test for reading abilities and reading comprehension. However, we matched both groups according their educational level, and assessing verbal fluency suggested rather similar verbal abilities.

Despite comparable behavioral results and similar brain activation patterns for the single-group analyses, distinct group differences between ASD patients and controls were seen for the SE, IND and pooled moral dilemma conditions, in which ASD patients showed significantly reduced activation of empathy-related regions as compared to the HC group, particularly during the IND condition, and likewise putatively compensatory abnormal hyperactivations of regions associated with the ‘default mode network’ and also self-projective cognition.

Though according to our study ASD patients do not exhibit fundamental deficits in moral reasoning, these specific brain activation patterns propose an altered way of (meta-)cognitive processing of social information and indicate a neural endophenotype for ASD reflecting dysfunctions in the underlying network.


