Moral reasoning about everyday situations in adults with autism spectrum disorder

Giulia Bellesi*, Karishma Vyas, Leila Jameel, Shelley Channon
Department of Experimental Psychology University College London, London, United Kingdom

ARTICLE INFO

Keywords:
Autism spectrum disorders
Social cognition
Social behaviour
Empathy
Moral reasoning
Everyday functioning

ABSTRACT

Little work has explored moral reasoning in adults with autism spectrum disorder (ASD). The present research used two novel scenario-based tasks to investigate how adults with and without ASD reason about moral rules and apply them in everyday interactions. The Rule Transgressions task examined decision-making in situations where breaking a rule might lead to a personal advantage. People with ASD did not differ from control participants in how frequently they chose to carry out the transgressions or their ratings of acceptability. However, they gave fewer sophisticated rationales to explain why the transgressions were wrong. The Social Intentionality task examined how participants judged and reacted to intentional and accidental transgressions performed by a story character. Both groups judged and responded more harshly to intentional transgressions than unintentional ones, but participants with ASD gave harsher responses overall. The findings suggest some preservation of the understanding of moral rules in ASD, but difficulties in applying and reasoning about these.

1. Introduction

Moral rules provide people with information about how they “ought to act”; their application and reinforcement helps to minimise conflict and promote cooperation within complex social groups (Staub, 2003). There is evidence that consideration of what is “right” or “wrong” relies on a combination of both reasoned and more automatic socio-emotional processes; cognitive and emotional empathy in particular appear to have an important role in supporting moral reasoning (Blair & Fowler, 2008; Cushman, Young, & Greene, 2010). Through reliance on cognitive empathy, the ability to understand others’ perspectives, people can simulate in their minds what they would or would not like to happen if they were in the same situation (Gibbs et al., 2007). Emotional empathy, on the other hand, allows people to ‘feel’ for others (Pizarro, 2000). Witnessing someone being harmed elicits congruent affective responses in the observers, which can facilitate their ability to reason about and implement morally prescribed rules (Eisenberg & Strayer, 1987; Singer et al., 2004).

Although it is unclear whether emotional empathy is preserved in individuals with ASD, there is good evidence to suggest that cognitive empathy is impaired (Blair, 2008). Given the role of this resource for facilitating moral reasoning (Gibbs et al., 2007), people with ASD might be expected to differ from neurotypical individuals in this. The evidence, however, is mixed, and relates predominantly to children and adolescents with ASD. Initial studies examining the ability to differentiate between conventional versus moral transgressions or anti- versus pro-social acts have shown no group differences (James & Blair, 1996; Leslie, Mallon, & DiCorcia, 2006). Notably, though, there is some evidence that the ability of children and adolescents with ASD to reason about moral problems may not be as sophisticated (Grant, Boucher, Riggs, & Grayson, 2005). For instance, in a study by Shulman et al. (2012),
they generated fewer sophisticated justifications than neurotypical participants to explain why different transgressions were wrong, containing more nonspecific condemnations (e.g., “that’s bad”) or citing basic rules (“it is forbidden to hit others”).

Some studies within moral reasoning have also explored whether people with ASD show similar sensitivity to intentionality in their judgments. Neurotypical people typically judge others more harshly for pursuing a detrimental course of action if they carried it out on purpose rather than accidentally (Armsby, 1971; Cushman, 2008; Cushman, Shekoff, Wharton, & Carey, 2013; Lagnado & Channon, 2008). Previous tasks where the agent’s intentions were not clearly spelt out have found that people with ASD often over-attribute intent to unintentional actions, and judge these more harshly than neurotypical participants (Buon et al., 2013; Koster-Hale, Saxe, Dungan, & Young, 2013; Moran et al., 2011; Rogé & Mullet, 2011; Salvano-Paridieu et al., 2016). On the other hand, in studies where the difference in the agent’s intent was more explicit, they correctly differentiated between intentional and unintentional actions (Grant et al., 2005). Indeed, in a study by Channon et al. (2011) participants with ASD not only differentiated between intentional and unintentional transgressions, but did so even to a greater extent than neurotypical individuals, since they judged intentional wrong-doings more harshly and accidental ones more leniently. It is possible that when demands on mentalistic abilities are kept low by spelling out and contrasting intentions, people with ASD can use this information in similar ways to neurotypical people to guide their moral judgements.

A main limitation of previous research is that studies have been predominantly conducted with children and adolescents. Moral reasoning is thought to follow a developmental trajectory, characterised by the succession of different, hierarchical stages (Kohlberg, 1973). Initially, children rely on an utilitarian approach, focusing on the practical consequences of their actions; however, as they grow older and become adults, they increasingly show a more sophisticated ability to grasp and apply universally applicable concepts such as fairness, justice and human rights (Smetana, 2006; Turiet, 2006). Moral development has been argued to continue throughout lifespan (Kohlberg, Levine, & Hewer, 1983). Since little work has been conducted with adults with ASD, it is currently unclear whether they follow the same trajectory as neurotypical individuals in their moral reasoning. Notably, moral development has been associated with maturation and increased activity in the prefrontal cortices of the brain, particularly the ventromedial and orbitofrontal regions (Fang et al., 2017; Pascual, Gallardo-Pujol, & Rodrigues, 2013). Both structural and functional abnormalities in such areas have been implicated in both children and adults with ASD (e.g., Amaral, Schumann, & Nordahl, 2008; Ha, Sohn, Kim, Sim, & Cheon, 2015). An interesting question is therefore whether moral reasoning remains impaired in adults with ASD, or whether they instead follow an alternative, perhaps slower trajectory in their moral development.

Another weakness of previous tasks is that they typically involve the use of serious and unfamiliar transgressions (e.g., an agent murdering a victim). These tasks are unlikely to reflect the rich and varied types of moral issues that people need to resolve in their everyday interactions; performance on them is thus unlikely to be representative of how people respond to moral problems in commonplace situations and contexts.

The present study aimed to address these limitations in the literature, by presenting adults with and without ASD with two novel scenario-based tasks: the Rule Transgressions task and the Social Intentionality task. Both the Rule Transgressions and the Social Intentionality tasks were developed especially for the present research; however, similar scenario-based tasks have been used in the field of social cognition (e.g., Callenmark, Kjellin, Rönqvist, & Bölte, 2014; Zalla, Sav, Stopin, Ahade, & Leboyer, 2009). In contrast with many previous moral tasks, both the Rule Transgressions and Social Intentionality tasks presented participants with real-life types of situations and problems (e.g., participants had to decide where they would say a small lie, or how they would react to a friend betraying their trust); use of everyday-type of social situations can help to enhance ecological validity and thus our understanding of how ASD may actually affect people’s behaviours in real-world scenarios. Tasks such as the present ones also provide with opportunities to gain a more detailed insight into participants’ thought and reasoning processes and shed light with respect to the more subtle differences found between neurotypical individuals and people on the high-functioning end of the autism spectrum. The tasks were designed and administered concurrently as they both examined how people evaluate different but related aspects of moral reasoning.

The Rule Transgressions task explored how people evaluate transgressions that may lead them to gain a personal advantage. Self-interest is a powerful motivator for unethical behaviour: people are more likely to break a norm if they think that this could help them to obtain something that they value (Bersoff, 1999), and also to view such transgressions as more permissible (Bocian & Wojciszke, 2014). People with ASD are often described as self-focused (e.g., De Vignemont, 2007), and have been found to prioritise their own benefits over those of others in social situations (Bellesi, Jameel, Vyas, Crawford, & Channon, 2016; Jameel, Vyas, Bellesi, Cassell, & Channon, 2015). Thus, one interesting question is how they perceive transgressions that may be advantageous to their perpetrators. The Rule Transgressions task examined this by presenting participants with situations where a character could achieve something instead follow an alternative, perhaps slower trajectory in their moral development.

The Social Intentionality task aimed to extend previous work examining sensitivity to intentionality in adults with ASD, and also explore their responses to harmful acts. Participants read different variants of the same written scenarios, where a fictional character engaged in either an intentional or unintentional wrongdoing. As reviewed above, when the distinction between intentional and unintentional actions is explicit, people with ASD seem able to differentiate between those in their judgments (Channon, Lagnado, Fitzpatrick, Drury, & Taylor, 2011; Grant et al., 2005). Hypotheses for the present study were made in accordance with the findings by Channon et al. (2011), since a similar methodology and population were used. It was hypothesised that both groups would differentiate between intentional and unintentional actions by judging and responding more harshly to intentional transgressions versus unintentional ones. Moreover, on the basis of previous evidence that adults with ASD may differentiate even more strongly
between intentional and unintentional actions (Channon et al., 2011), it was also hypothesised that, compared to control participants, they would judge and react more harshly to the transgressions in the intentional condition, and less harshly to those in the unintentional condition.

2. Methods

2.1. Participants and procedure

A-priori power analyses based on the commonly used conventions provided by Cohen (1992) indicated that a sample size of 20 participants per group would be adequate to detect large effect sizes, with power set at 80% and alpha at 10%. This sample size is also comparable to previous studies of ASD and social cognition (e.g., Bellesi et al., 2016; Shulman et al., 2012). Twenty university students with ASD were recruited by advertising for people who had been given a diagnosis of Asperger’s syndrome or autism by a clinician. Twenty control student participants were also recruited by advertisement at the same university. The diagnosis of ASD was confirmed through the use of a semi-structured interview to assess whether they met DSM-5 criteria for ASD. The interview was devised by the research team, including two clinicians with extensive experience in the field, who trained the first author to conduct the interviews. The interview lasted approximately 60 min and related closely to current DSM-5 criteria. It enquired in detail about the presence of qualitative difficulties in social interactions, non-social impairments (e.g., repetitive and stereotyped patterns of behaviour), and impact of symptoms on everyday functioning.

Inclusion criteria were fluency in English, age between 18 and 30, and a full-scale IQ (FSIQ) score of 85 or above on the Wechsler test of adult reading (WTAR; Wechsler, 2001). The WTAR is a widely used neuropsychological assessment tool with good validity and reliability in estimating IQ (e.g., Spreen & Strauss, 2006). Participants with history of learning disability or illness or injury involving the brain were excluded.

Participants’ demographic characteristics and their WTAR FSIQ mean scores and standard deviations (SD) are shown in Table 1. There was no significant group difference in age, t(38) = .87, p = .381, or WTAR FSIQ, t(36) = 1.55, p = .063. All participants provided informed consent before taking part in the research and were paid £20 at the end. Both tasks were administered in counterbalanced order by the first author within the same testing session, which lasted approximately three hours in total. The study was approved by the UCL Research Ethics Committee and carried out in accordance with the Declaration of Helsinki as revised in 2000.

2.2. The rule transgressions task

This task assessed how people reasoned about situations where violating a rule might help them to gain a personal advantage (Fig. 1). Participants read 10 descriptions of everyday situations. In each, a main character wanted to achieve or obtain something, but could do so only by breaching a rule (e.g., by lying). Two different orders of presentation of the scenarios were used within each participant group, to reduce order effects. The main character’s gender was male in half of the scenarios, and a female in the other half, and the transgressions occurred in a variety of different contexts (e.g., family, social, or workplace). After reading the task instructions, participants read and responded to an example scenario and questions; no feedback was provided in order to keep the task administration as standardised as possible and not to influence performance. Each scenario was shown one at a time in printed form on a separate page of the same booklet, and remained on display when the question was presented to avoid potential confounding effects from memory load. Participants provided all their answers verbally.

Decisions to carry out the transgressions. For each scenario, participants were asked to say whether they would personally choose to carry out the transgression if they were the story character. The number of times each participant said that they would carry out the transgressions was averaged across scenarios and converted into a percentage.

Acceptability ratings. Participants rated on a scale from 1 to 10 how acceptable it would be to carry out the transgression. Ratings were averaged across scenarios and converted into percentages.

Rationales. Participants were asked to explain why people might think that it is wrong for someone to carry out the transgressions. Responses were classified into two categories, ‘simple’ or ‘sophisticated’. Rationales were categorised as simple when participants reiterated elements of the scenarios, provided nonspecific condemnations of the transgressions, or cited simple rules that prohibited them without further explanation. They were classified as sophisticated when they showed a higher-level understanding of the transgressions, for example by referring to abstract principles or recognising the possible damaging effects on others (Fig. 1). Responses could score for only one of the two categories; if both were met, the best answer was taken. All scenarios were double-scored by a rater who was not blind to group membership, and by a blind independent rater. The inter-rater agreement was found to be

Table 1

<table>
<thead>
<tr>
<th>Group with ASD</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mean (SD)</strong></td>
</tr>
<tr>
<td>Age (years)</td>
<td>22.65 (3.56)</td>
</tr>
<tr>
<td>No. of male participants</td>
<td>12 out of 20</td>
</tr>
<tr>
<td>FSIQ (points)</td>
<td>110.75 (5.32)</td>
</tr>
</tbody>
</table>
Example scenario: Alan is unhappy with his current job, and finds an opening to work for the company of his dreams. He meets all the job criteria except for one, since he does not have much knowledge about computer programming. He has the opportunity to lie about his skills in the application form.

Question 1: If you were Alan, would you lie about your skills in this situation? Yes vs. No

Question 2: On a scale from 1 to 10, where 1 represents ‘not at all acceptable’ and 10 represents ‘very acceptable’, how acceptable is it to do that in this situation?

Question 3: Why might people think that it is wrong for someone to do that in this situation?

Examples of rationales:

- Simple: “*Alan does not have experience with computer programming*” / “*It’s lying*”

- Sophisticated: “*He is taking advantage of a system based on supplying truthful information*” / “*It is unfair towards other candidates, who may be better qualified for the job*”

Fig. 1. Example scenario, questions and rationales from the Rule Transgressions task.

substantial, with Cohen’s Kappa = .716 (Landis & Koch, 1977). All disagreements were resolved through discussion. Scores for each strategy were averaged across scenarios and converted into percentages.

2.3. The social intentionality task

This task assessed how people reasoned about everyday situations where someone acted to their detriment, either intentionally or accidentally (Fig. 2). Participants read 8 descriptions of social interactions. There were two different variants for each scenario, so that the character acted to the participants’ detriment either intentionally or unintentionally. After reading the task instructions, participants worked through an example scenario comprising both an intentional and unintentional variant.

Response strategies. For each scenario variant, participants were asked what they would do in response. Their responses were scored depending on how harsh these were towards the character. Responses were assigned a score of 1 when participants said they would not do or say anything; a score of 2 when they stated that they would raise the issues with the character (e.g., discuss the situation with them, or rely on them less); a score of 3 when participants chose particularly harsh courses of actions such as responding in kind, or ending their friendship with the character (Fig. 2). All scenarios were double-scored; the inter-rater agreement was found to be substantial, with Cohen’s Kappa = .802 (Landis & Koch, 1977). Scores were averaged up across scenarios, so that each participant received a total score ranging between 8 and 24. These were then converted into percentages.

Acceptability ratings. Participants rated on a scale from 1 to 10 how acceptable the character’s action was. Ratings were averaged across scenarios and converted into percentages.

Rank-ordering. After giving a response to the scenario and rating acceptability, participants were told that the character apologised for what they did, and were presented with four different courses of action. These varied in how harsh they were, and involved: a) remaining friends with the character; b) remaining friends with the character, but acting slightly differently towards them (e.g., putting less trust in them, or relying on them less); c) responding in kind; d) ending their friendship with the character. Participants were asked to rank-order the four options to say how they would behave towards the character, from the most to the least likely (Fig. 2).

The response options ranked by participants as the ‘most likely’ were considered. Participants received a score of 1 when they chose a) as their most likely option, 2 when they chose b), and 3 when they chose c) or d). Options c) and d) scored the same because they were the harshest responses, and because people may have different perceptions concerning which option between them is harsher. Scores were averaged across scenarios and converted into percentages.
**Example scenario:** You tell your course mate in confidence about an embarrassing thing that happened to you. Later on, a mutual friend asks you about it.

**Intentional variant:** You find out that your course mate revealed it to them, thinking that it was too funny a story not to share it.

**Unintentional variant:** You find out that your course mate let it slip after having too much to drink.

**Question 1:** What would you do in this situation?

Examples of strategies:

- No response: “I would accept the situation, and move on” / “I wouldn’t do anything at all”
- Moderate: “I would approach the course mate, and tell them I was unhappy with what they did” / “I would be more cautious in revealing personal information to them”
- Extreme: “I would also reveal one of my course mate’s secrets to someone else” / “I would stop being friends with the course mate”

**Question 2:** On a scale from 1 to 10, how acceptable is what your course mate did?

**Question 3:** You mentioned it to your course mate, who apologised. Please rank-order the following options to say how you would behave towards your course mate from the MOST likely to the LEAST likely:

a) Continue being friendly to her as usual
b) Continue being friendly to her, but do not confide in her anymore
c) Spread a rumour about her too
d) Stop being friendly to her

Fig. 2. Example scenario, questions, and strategies from the Social Intentionality task.

3. Results

3.1. Data analysis

Means and standard deviations (SD) for each task are shown in Table 2 and 3 respectively. A significance level of $p < .05$ was adopted throughout. All assumptions for the analyses described below were met. All analyses were also repeated adding IQ as a covariate; inclusion of this did not alter the pattern of findings.

<table>
<thead>
<tr>
<th>Table 2 Mean scores, standard deviations, and significance for the Rule Transgressions task.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Decisions to carry out the transgressions (%)</td>
</tr>
<tr>
<td>Acceptability ratings (%)</td>
</tr>
<tr>
<td>Sophisticated rationales (%)</td>
</tr>
</tbody>
</table>

** $p < .01$. **
Table 3
Mean scores, standard deviations, and significance for the Social Intentionality task.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Group with ASD</th>
<th>Control group</th>
<th>Group main effect</th>
<th>Effect size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harshness of free response strategies (%)</td>
<td>Condition</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
</tr>
<tr>
<td>Intentional</td>
<td>69.17 (9.79)</td>
<td>63.96 (5.78)</td>
<td>-</td>
<td>0.65</td>
</tr>
<tr>
<td>Unintentional</td>
<td>55.63 (4.93)</td>
<td>52.71 (4.33)</td>
<td>-</td>
<td>0.63</td>
</tr>
<tr>
<td>Acceptability of characters’ actions (%)</td>
<td>Condition</td>
<td>24.81 (7.61)</td>
<td>34.13 (11.03)</td>
<td></td>
</tr>
<tr>
<td>Intentional</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Unintentional</td>
<td>56.00 (12.15)</td>
<td>68.19 (13.27)</td>
<td>-</td>
<td>0.96</td>
</tr>
<tr>
<td>Harshness of preferred strategy in rank-ordering (%)</td>
<td>Condition</td>
<td>52.03 (13.15) 45.00</td>
<td>68.19 (13.27)</td>
<td></td>
</tr>
<tr>
<td>Intentional</td>
<td>(7.75)</td>
<td></td>
<td>-</td>
<td>0.65</td>
</tr>
<tr>
<td>Unintentional</td>
<td>35.47 (6.35) 30.16 (3.83)</td>
<td></td>
<td>-</td>
<td>1.01</td>
</tr>
</tbody>
</table>

* p < .05.
** p < .01.

3.2. The rule transgressions task

Decisions to carry out the transgressions. The groups were compared on the number of times that they said they would personally choose to carry out the transgressions. There was no significant difference between groups, t(38) = 1.42, p = .164.

Acceptability ratings. The groups were compared on their acceptability ratings. There was no significant difference between groups, t(38) = 1.29, p = .204.

Rationales. The groups were compared on the types of rationales they generated to explain why carrying out the transgressions was wrong. There was a significant difference between groups in the proportion of sophisticated rationales used, t(38) = 3.62, p = .001. Inspection of the means revealed that whilst participants with ASD generated sophisticated rationales less than half the time, the control group relied on these for more than half of their responses.

Further analyses were carried out to explore whether there might be a relationship between IQ and generation of sophisticated rationales. There was no significant correlation between IQ scores and percentage of sophisticated rationales for either the group with ASD, r = .342, p = .140, or the control group, r = .118, p = .621. A one-way ANCOVA comparing groups on the number of sophisticated rationales and including IQ as a covariate showed no significant main effect of this, F(1, 37) = 1.61, p = .213.

3.3. The social intentionality task

Harshness of free response strategies. The groups were compared on the degree of harshness displayed in responding to the scenarios. A 2 × 2 ANOVA (group by condition: intentional or unintentional) found a significant main effect of intentionality, F(1,38) = 92.18, p < .001, and a significant main effect of group, F(1,38) = 6.26, p = .017. There was no significant intentionality by group interaction, F(1,38) = .79, p = .380. Inspection of the means revealed that all participants had higher harshness scores for the intentional condition compared to the unintentional condition. Participants with ASD had higher harshness scores compared to control participants overall.

Acceptability ratings of character’s actions. The groups were compared on their ratings of the acceptability of the character’s actions. A 2 × 2 ANOVA showed a significant main effect of intentionality, F(1,38) = 294.94, p < .001, and a significant main effect of group, F(1,38) = 12.88, p = .001. There was no significant intentionality by group interaction, F(1,38) = .573, p = .454. Inspection of the means revealed that all participants rated the character’s actions as less acceptable in the intentional condition compared to the unintentional condition. Participants with ASD gave lower acceptability ratings compared to control participants overall.

Harshness of preferred strategy in rank-ordering. The groups were compared on their composite scores reflecting how harsh their preferred response option was after an apology from the character. A 2 × 2 ANOVA showed a significant main effect of intentionality, F(1,38) = 124.63, p < .001, and a significant main effect of group, F(1,38) = 7.29, p = .010. There was no significant intentionality by group interaction, F(1,38) = .373, p = .545. Inspection of the means revealed that all participants had higher scores for the intentional condition compared to the unintentional condition. Participants with ASD had higher scores compared to control participants overall.

4. Discussion

The present research examined how adults with ASD reason about moral rules using two novel scenario-based measures, the Rule Transgressions and Social Intentionality tasks. Both tasks presented participants with familiar and commonplace type of situations and moral problems, to gain a more detailed understanding of how ASD may affect people’s behaviour in real-life scenarios, and illuminate the presence of potential subtle differences in the reasoning and thought processes shown by neurotypical adults versus those on the high-functioning end of the autism spectrum. The Rule Transgressions task presented participants with situations where transgressing a rule might help them to gain a personal advantage. It was hypothesised that participants with ASD would choose to carry out transgressions more often than control participants, and rate these as more acceptable. Neither prediction was confirmed, since the groups did not differ in their performance on either question. However, the prediction that people with ASD would generate significantly fewer sophisticated rationales than neurotypical participants to explain why the transgressions were wrong was
confirmed.

People tend to view transgressions that may help them to gain something positive as more acceptable than when they do not (Bocian & Wojciszke, 2014; Epley & Caruso, 2004). Since individuals with ASD have been previously shown to be more likely to prioritise their own preferences in social interactions (Bellesi et al., 2016; Jameel et al., 2015), it seemed logical to expect that they would be even more likely to choose courses of actions that were advantageous to them and rate these are more acceptable in the Rule Transgressions task, compared to control participants. The finding that, in fact, groups did not differ on either aspect of performance is therefore somewhat surprising and notable in the context of previous literature on social performance in ASD. It indicates that, in these types of circumstances, adults with ASD may be able to put aside self-interest and make similar moral judgements to neurotypical people.

Notably, however, participants with ASD still generated fewer sophisticated rationales than control participants when explaining why the transgressions were wrong. For example, they often reiterated information (e.g., “he lied”) or mentioned simple rules (e.g., “lying is not correct”), whereas control participants’ rationales often referred to abstract principles or the effect of rule-breaking on others (e.g., “it would compromise the opportunities of other, more prepared candidates”). There is already some previous evidence that children and adolescents with ASD show less sophisticated understanding of moral rule violations in their rationales (e.g., Grant et al., 2005; Shulman et al., 2012). Moral development is thought to comprise a series of stages, whereby, as individuals grow up, their reasoning increasingly relies on the application of abstract and universal ethical principles (Smetana, 2006; Turiel, 2006). Since most previous studies have been conducted with children and adolescents, to date it is unclear whether people with ASD follow the same developmental trajectory in their moral development, i.e. whether their moral reasoning continues to remain impaired throughout the lifespan, or whether this simply follows a different, perhaps slower, pattern. The present findings add to existing literature by indicating that, although people with ASD may still be able to recognise and rate transgressions in similar ways to neurotypical people, their understanding of these remains less refined and more rule-based.

Moving now to the findings of the Social Intentionality task, in this, participants read scenarios where a fictional character did something detrimental to them, either intentionally or accidentally. Firstly, it was expected that both groups would respond more harshly to intentional actions than to unintentional ones. This was supported: compared to unintentional actions, all participants rated intentional actions as less acceptable, generated harsher response strategies, and preferred harsher options when rank-ordering different courses of action that they might take after the characters had apologised. Secondly, it was hypothesised that participants with ASD would show heightened sensitivity to intentionality, generating and preferring harsher response strategies, and rating the characters’ actions as less acceptable in the intentional condition; the opposite pattern of performance was expected in the unintentional condition. This hypothesis was not supported, since there was no interaction between condition and group. Interestingly, however, a main effect of group was found, since participants with ASD rated the characters’ actions as less acceptable, generated harsher strategies, and preferred harsher options in their rank-ordering overall.

The finding that people with ASD could differentiate between intentional and accidental actions is consistent with previous evidence that they can do so when the agent’s intent is made explicit (e.g., Grant et al., 2005). Of course, it is possible that participants might have struggled to differentiate between intentional and unintentional actions had they been required to work out the character’s intentions from scratch (Buon et al., 2013; Moran et al., 2011). It might be interesting for future research to use a modified version of the present task, where the character’s intentions are less clearly spelt out, to examine their performance under these, more demanding circumstances.

A perhaps more key finding from the Social Intentionality Task was that participants with ASD were harsher in their responses than control participants, regardless of the agent’s intent. Participants with ASD judged the character’s actions as less acceptable, and preferred harsher strategies to deal with the situations. For instance, people with ASD were more likely to choose to end their friendship with the character or to respond in kind (e.g., to reveal one of the character’s personal secrets as well). This finding has interesting implications in that it suggests that high-functioning adults with ASD may present with a tendency in real-life scenarios to respond more harshly than neurotypical individuals to unpleasant social interactions, regardless of whether the person they were interacting with harmed them intentionally or by accident. Since people with ASD often struggle maintaining friendships and positive social relationships (e.g., Bauminger & Shulman, 2003; Locke, Ishijima, Kasari, & London, 2010), it is possible that such qualitative differences in how they approach challenging social interactions might contribute to maintain and exacerbate some of their difficulties in real-life exchanges. This might also explain why our findings slightly deviate from those of a previous study by Channon et al. (2011). In this, adults with ASD differentiated more strongly in their judgements of blame between intentional and unintentional actions, so that they evaluated intentional actions more harshly and unintentional ones more leniently compared to neurotypical participants. However, the scenarios in the Channon et al. (2011) study described interactions between two strangers, whereas in the present experiment participants were always the victim of the agent’s actions. It is therefore likely that individuals with ASD judge and react more harshly to transgressions that have a detrimental effect on them personally than neurotypical individuals. Future studies may explore this by manipulating the identity of the transgressions’ victims in a systematic way.

It should perhaps be noted that the outcomes of the transgressions presented in either task were not described, since in real life, this information is not typically available in advance, especially in the context of relatively mild violations. This leaves unanswered the question of whether other group differences would have emerged, had participants been informed about the transgression outcomes. It has been suggested that people with ASD might be over-reliant on information about action outcomes compared to other factors (such as the agent’s mental states) when judging moral transgressions (Margoni & Surian, 2016). It may be interesting for further research to explore whether this is also the case when judging mild violations by manipulating the action consequences more explicitly.
4.1. Theoretical models of ASD

Various explanations might account for the current findings, including reliance on social knowledge and impaired empathy. Moral reasoning is thought to rely on both learnt, explicit principles, and socio-emotional processes including cognitive empathy (Haidt, 2003). Since there is robust evidence suggesting that this resource is impaired in ASD (Frith, 2001), it would be implausible to explain intact aspects of performance in relation to preserved cognitive empathy. One more plausible explanation is that people with ASD may have developed a set of learnt rules that allow them to make similar moral judgments to neurotypical individuals. In the Rule Transgressions task, reliance on rules might explain why the group with ASD did not differ from neurotypical participants in their decisions to carry out the transgressions and acceptability ratings. Reliance on learnt knowledge might have also been sufficient to differentiate similarly to control participants between intentional and unintentional transgressions in the Social Intentionality task.

Another (although not mutually exclusive) explanation for the lack of group differences in certain aspects of performance might relate to participants’ levels of intellectual abilities. It can be noted that group differences with respect to FSIQ score approached statistical significance, with the group with ASD appearing slightly more able. It has been previously suggested that high-functioning people with ASD might be able to “hack out” solutions to social cognition tasks through reliance on intact intellectual skills and logical thinking (e.g., Frith & Happé, 1999; Hill & Frith, 2003). Higher verbal IQ in particular has been proposed to potentially help those with ASD “bootstrapping” their impaired perspective-taking abilities (Durrleman & Franck, 2015; Livingston & Happe, 2017). The precise mechanisms linking intact or above-average intellectual abilities with compensation in ASD need clarification ( Livingston & Happé, 2017). It has been proposed that individuals with preserved intellectual functioning might be able, through time and experience, to form associations between behaviours and outcomes, leading to gradual acquisition of mental state concepts and enhanced social performance in both experimental tasks and real-life situations (Frith, 2001). It is thus possible that average (or enhanced intellectual skills) in ASD might indeed lead to greater ability to acquire, rely on and apply logically social knowledge and rules, despite still lacking intuitive empathic skills in everyday social interactions (Bowler, 1992). However, the present participants were still impaired in some aspects of performance, including poorer verbal rationales, and inclusion of IQ as a covariate in the analyses was not found to alter the pattern the findings, suggesting that this did not influence performance significantly. Generation of sophisticated rationales, which is likely to have been the most demanding aspect of the study, was also found not to be related to intellectual abilities. Taken together, the findings would seem to suggest that intellectual abilities and/or social knowledge alone are unlikely to explain the pattern of findings, and thus the available resources might be adequate only for lower-demand types of tasks.

Deficits in cognitive empathy might represent a plausible explanation for impaired aspects of performance in ASD. In the Rule Transgressions task, it is likely that the requirement to generate rationales drew more heavily on mentalistic abilities, since all transgressions were going to have unpleasant consequences on others in the long-term. For example, in the job application scenario, lying on the application form would reduce other candidates’ chances of being considered for the job; if participants were hired, their lack of experience might damage the company in the long-run; their employer might eventually find out about their lie, and feel disappointed. Difficulties taking others’ perspectives may have hindered the ability of people with ASD to consider or anticipate these possible implications, and thus reach a more sophisticated understanding of the scenarios. By contrast, neurotypical participants may have been more skilled at appreciating the potential consequences of the transgressions with respect to the opportunities, rights, or welfare of others, and also societal living in a wider sense. Similarly, in the Social Intentionality task, the responses by participants with ASD might have been less tempered by socio-emotional processes such as empathy. It is possible that, when presented with transgressions having a detrimental effect on them, individuals with ASD might rely even more rigidly on learnt rules, and be less tolerant of one’s actions regardless of intent, leading to a more ‘black-and-white’ and stricter pattern of responses. Overall, these findings suggest that, although reliance on norms and deliberate reasoning may allow individuals to make correct moral judgements, cognitive empathy is still likely to be an essential skill for higher-level understanding of everyday moral problems.

The suggestion that differences in performance by people with ASD may stem from difficulties appreciating others’ mental states is further corroborated by previous evidence exploring their responses to social versus non-social problems. People with ASD have been shown to struggle explaining social events (e.g., sarcastic remarks), whilst they can perform relatively similarly to neurotypical participants for non-social items (Channon, Crawford, Orlowska, Parikh, & Thoma, 2014; Spek, Schatorjé, Scholte, & van Berckelaer-Onnes, 2009). Evidence of group differences in moral reasoning about social versus non-social events has been hinted recently by Vyas, Jameel, Bellesi, Crawford, and Channon, 2017. Students scoring high versus low on self-report measures of autistic traits read adapted versions of the classic trolley and footbridge dilemmas, where an agent had to make a decision that would either favour one character at the expense of a group, or vice versa. The scenarios outcome involved either physical (e.g., death) or social (e.g., emotional distress) harm. Participants explained why agents might choose to favour either one individual or the group. Whereas people scoring low in number of autistic traits were more likely to consider the characters’ perspectives in situations involving social harm (and less in situations involving physical harm), people with high scores showed the opposite pattern. Considering the characters’ perspectives in social situations, where consequences for them were not as obvious (e.g., distress) might have required to rely on more sophisticated reasoning skills and thus been more demanding for those scoring high on number of autistic traits. On the other hand, it may have been easier for them to refer to characters for physical scenarios, where consequences were more clear-cut (e.g., death). Taken together, these findings suggest that problems relying more heavily on reasoning about others’ mental states may be harder for people with ASD; on the other hand, stimuli for which it is sufficient to rely on analytical considerations may be less demanding, leading to fewer group differences. Of course future research might explore this further by for example replicating the present tasks using scenarios with more and less marked social versus non-social components.
4.2. Limitations and future research

It is important to consider some limitations of the studies. A first possible issue concerns the external validity of scenario-based tasks. Although these have many strengths, it could be argued that they are still less complex and interpersonally stimulating than naturally-occurring situations, and thus may not reflect how people would actually perform in everyday life.

Both scorers of the qualitative responses were aware of the tasks’ hypotheses; it cannot be excluded that this may have skewed the direction of their ratings. However, one of the scorers was blind to group membership, which should have significantly reduced the risk of bias. The interrater reliability coefficients were also high for both tasks, suggesting that participants’ responses could be reliably discriminated. Finally, diagnosis of ASD was confirmed via a semi-structured interview, rather than standardised instruments such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989). However, it is worth noting that the ADOS does not relate directly to DSM-5 criteria, and all participants had already been diagnosed prior to participation; moreover, our previous experience using this instrument suggested it to be rather limited and somewhat inappropriate with high-functioning adults.

Finally, although previous moral reasoning experiments have also used similar number of participants (e.g., Shulman et al., 2012), it is worth noting that our samples were also relatively modest in size. Sample sizes were determined on the basis of Cohen’s (1992) criteria using an alpha level of 10%, which is often considered appropriate when conducting exploratory research (see e.g., Cohen, 1992, Schumm, Pratt, Hartenstein, Jenkins, & Johnson, 2013; Warner, 2013). Large effect sizes were expected on the basis of previous findings comparing people with ASD and neurotypical individuals on scenario-based tasks assessing everyday functioning and problem-solving; for instance, in a study assessing people’s responses to everyday awkward scenarios Bellesi et al. (2016) found all statistically significant group differences to have a Cohen’s d effect size > 1. Shulman et al. (2012) in their investigation with children and adolescents also found mostly large effect sizes. Nevertheless, the possibility remains that some group differences might have gone undetected or have been weakened in size as a result of reduced statistical power resulting from small sample sizes.

4.3. Implications

Current intervention tools have been predominantly designed for lower-functioning children or adolescents (Bishop-Fitzpatrick, Minshew, & Eack, 2012), and typically focus on teaching specific cognitive skills (e.g., taking others’ perspectives; Fisher & Happé, 2005), or broad pragmatic abilities (e.g., maintaining eye contact; Rao, Beidel, & Murray, 2008). Although these may prove effective within the laboratory context, the skills acquired within training sessions may be difficult to generalise to real-life situations, leading to no actual practical improvements (Begeer et al., 2011; Locke, Rotheram-Fuller, Xie, Harker, & Mandell, 2014). The current findings suggest that everyday life-type tasks such as those described here might be sensitive tools for identifying both impaired and preserved abilities in adults with ASD in solving common social problems.

The findings of this research also indicate that measures such as the Rule Transgressions and the Social Intentionality tasks might be instrumental in identifying and targeting potential discrepancies between knowledge and understanding of rules. Previous evidence using similar types of vignettes as the present study has shown that reliance on rules alone might help individuals with ASD to disguise their social impairments to an extent, but is unlikely to be a sufficient compensatory tool on its own, leading to inflexible and rigid patterns of behaviour. Improving people’s awareness of rules in combination with their ability to reason about these in a more sophisticated manner might support the generation of more flexible and socially appropriate responses, especially in complex and unpredictable social situations (Channon, Charman, Heap, Crawford, & Rios, 2001; Howlin, Goode, Hutton, & Rutter, 2004).

5. Conclusions

The present studies examined moral reasoning about everyday transgressions in adults with ASD. When presented with transgressions that may help them to gain a personal advantage, they did not differ from control participants in the number of transgressions they chose to carry out or ratings of acceptability. However, they generated fewer sophisticated rationales for their judgments. In a second task, when presented with intentional and unintentional transgressions with a detrimental effect on them, people with ASD were harsher in their responses overall than control participants. These findings advance our current understanding of moral reasoning in high-functioning adults with ASD and support the presence of subtle qualitative differences in how they reason about and solve common moral problems; these might underpin or exacerbate some of their real-life impairments in everyday social interactions.

Funding

The study was conducted as part of a PhD funded by a UCL Impact Scholarship.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgements

We thank Simone Freschi for double-scoring the data, and all the participants who took part in the study for their time and effort.
References


References


