

Investigating Moral Judgements in autistic children: integrating the observer's and the speaker's mind

Irene Garcia-Molina; Rosa-Ana Clemente-Estevan; Clara Andrés-Roqueta.

Irene Garcia-Molina: imolina@uji.es ; (+34) 964 387734

Rosa-Ana Clemente-Estevan: clemente@uji.es ; (+34) 964 729548

Clara Andrés-Roqueta: candres@uji.es ; (+34) 964 729559

Affiliation/address for all authors:

Universitat Jaume I de Castelló,

Departamento de Psicología Evolutiva, Educativa, Social y Metodología.

Av/ Vicent Sos Baynat s/n

12071 Castellón (Spain)

Acknowledgements

This research was submitted as partial fulfillment for the degree of PhD by the first author. The research was supported by studentship from Generalitat Valenciana (Val i+D) and the grant UJI-A2016-12 supported by Universitat Jaume I de Castellón. We would like to thank all the children, families, teachers and professors who kindly participated in this study.

Garcia-Molina, I., Clemente-Estevan, R. A., & Andrés-Roqueta, C. (2020). Investigating moral judgements in autistic children: integrating the observer's and the speaker's mind. *Journal of Cognitive Psychology*, 1-15. <https://doi.org/10.1080/20445911.2020.1856120>

Abstract

This study investigates the moral judgements that autistic children make in everyday situations. Moral Transgressions (MT) and *Faux Pas* (FP) stories were compared, in which stories the agent's morality and intention varied (MT: bad, FP: good), and were divided by the mediator (personal / material) as well. Sixty-two participants (30 autistic and 32 neurotypical children) answered forced-choice questions. The two groups did not differ significantly when responding to the different questions in either the MT or the FP stories. Related to the mediator of the action (hurt/damage), in between-group comparisons, the autistic group found difficulties in understanding the MT stories when the action directly affected another person (personal mediator). Comparisons between specific variables (agent's morality and intention) revealed that autistic children judged the morality of the agent in FP stories as severe as in the MT task, even when the agent's intention was understood. The subtle difficulties of autistic children could shed some light on how autistic individuals would judge social situations, from the lack of a robust ToM, to difficulties integrating some information and being socially flexible.

Keywords: Theory of Mind; Morality; Intention; Autism.

Investigating Moral Judgements in autistic children: integrating the observer's and the speaker's mind

Theory of Mind (ToM) is the capacity that allows us to predict and understand mental states, and therefore it is crucial for explaining others' behaviours based on mental reasoning, making us capable of judging actions and also the agents involved in a specific situation. Understanding intentions is a key aspect that allows us to attribute moral responsibility to an agent of an offensive or hurtful act (Baird & Astington, 2004; Cushman, 2015; Young & Saxe, 2009; Zalla & Leboyer, 2011). The shift in reasoning from outcome- to intent-based moral judgement has been a critical aspect of development since Piaget's (1932) seminal study. According to Piaget's Theory, young children's judgments are outcome-based, probably due to their lack of perspective taking, and they begin to make intent-based judgements between the ages of six and ten. Following this model, a growing number of studies investigated moral judgment in autistic individuals¹. Autism is a condition characterised by (i) deficits in social communication and interaction, and (ii) restricted and repetitive patterns of behaviours and interests (following DSM-5 criteria, APA, 2013). Social interaction issues such as problems making friends, bullying or mate crime can be explained by a deficit in ToM, because this capacity allows us to attribute intentions to others and, therefore, to evaluate their behaviours and judge them (see Margoni & Surian, 2016 for a review). Thus, autistic individuals could find difficulty in reasoning based on intent (e.g. Buon et al., 2013; Moran et al., 2011; Zucchelli, Nori, Gambetti, & Giusberti, 2018). Nevertheless, whether autistic children are less able to understand that the agent's evilness can be influenced by the understanding of his/her intention needs to be studied in depth.

¹ From here onwards the term autistic will be used, see Kenny et al. (2016). The acronym of Autism Spectrum Disorder – ASD will be employed only when groups are compared. See Vivanti (2019) for more information.

Moral judgements construction from the deconstruction of the intentional action

The folk concept of ‘intentional action’ explains how mental states can cause physical events (Cushman, 2015). The understanding of the elemental features that compose an intentional-action enables us to reconstruct its interrelation to moral judgements (see Figure 1).

In the moral literature in autism, aspects such as intentionality, responsibility, culpability or punishment have been widely studied (Buon et al., 2013; Killen Mulvey, Richardson, Jampol, & Woodward, 2011; Zalla & Leboyer). However, the aspect of agent’s morality (whether the agent was “good” or “bad”) has not been analysed yet in the autism field, as it has been recently introduced to the analysis of morality (Margoni & Surian, 2020).
INSERT FIGURE 1 NEAR HERE.

Figure 1. Intention of the action broken down into its elemental features and morality (adapted from Cushman, 2015).

Figure 1 shows the relationship between others’ mental states (non-observable categories such as desires, beliefs and intentions) and the actions and outcomes (observable categories – events). These actions can be the product of a deliberate action (plan under control) or an accidental action (non-controlled plan due to causes such as a mistakes, forgetfulness or lack of information). In short, people’s knowledge of others – or lack thereof – can give rise to beliefs that may or may not match reality, which can lead to ambiguity (Margoni & Surian, 2016). For example, in a moral transgression (MT), intention and outcome are unambiguous (a bad intent that causes a bad outcome, for an example see Appendix 1, the soup story in the current study). However, the faux pas is an ambiguous case, where intention and outcome are not with the same valence (good/neutral intent that causes a bad outcome, for an example see Appendix 1, the apple pie story).

Accidental action and the Faux Pas test.

The most common task used to explain the accidental situations is the Faux Pas (FP) test, and it has been used in some studies with autistic individuals to investigate the role of intentionality and mental states in the understanding of this particular social situation (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). A correct understanding of a FP would be: a) to identify the statement (someone said something that he/she should not have said); b) to comprehend that the speaker accidentally makes a comment (intentionality); c) that the comment hurts emotionally the listener (negative impact). What it is crucial in a FP is that the speaker usually has a false belief about reality (he/she does not know some important information), and consequently grasping his/her lack of intention to hurt the listener is a key aspect to correctly understand the situations. Nevertheless, the observer (the participant being assessed) can access to all the information and may be able to distinguish the intention of the speaker as good, regardless of whether the action and outcome are judged as 'bad'. In the next example, the two paths cited (speaker's and observer's perspective) can be analysed for better understanding: Will bought Thea a toy airplane for her birthday. A few months later, they were playing with it, and Will accidentally dropped it. "Don't worry," said Thea, "I never liked it anyway. Someone gave it to me for my birthday" (extracted from Baron-Cohen et al., 1999, and also used in the present study).

Path 1 (speaker's perspective): Thea wants to tell Will that it is fine, and she *believes* that another person bought the toy for her (False belief) – Thea plans the positive comment (Neutral-Good intention) – Thea says the comment as she planned (Deliberate 1st action) – Thea communicates her positive comment to Will so that he does not feel sad for dropping the toy (Good Outcome).

Path 2 (observer's perspective): Thea wants to say to Will that it is fine, but Thea does not know/remember that Will bought this toy (True belief: Will bought this toy a few

months ago) – Thea does not plan to hurt Will (Neutral-Good intention) – Thea makes the negative comment without controlling for certain aspects (Accidental 2nd action) – Thea’s negative comment hurts Will (Bad Outcome).

Therefore, the recognition of a FP depends specifically on the understanding of false beliefs and intentions, and these are ToM aspects that autistic people often misinterpret. In this sense, Zalla et al. (2009) observed that autistic adults could correctly identify the FP statement, but they failed to interpret the false belief and intentionality of the speaker. Moreover, they provided explanations in terms of malevolence by judging the speaker’s intention to humiliate or offend the listener as deliberate. In the same study, Zalla et al. (2009) suggested that inappropriate judgements of intention in autistic participants mostly occur when (a) the negative outcome of the agent’s action is produced accidentally, (b) the agent’s intention is inferred verbally, and (c) the judgements about intention are influenced by the moral valence of the outcome (i.e. good intention / bad outcome).

As far as we know, the morality of the agent who commits the FP (i.e. says an awkward comment) has not been analysed nor compared with tasks involving moral scenarios. Recently research is paying attention to the FP test, suggesting that even when autistic individuals attribute mental states correctly, intention is not fully used for moral reasoning (Buon et al., 2013; Zalla & Leboyer, 2011; Zalla, Barlassina, Buon, & Leboyer, 2011; Zalla, Sav, Stopin, Ahade, & Leboyer, 2009).

Intention- or outcome-based judgements and the complexity of the agent’s inferences

When there is an emotional impact (i.e. someone being hurt), autistic people tend to make judgements of the agent based on the consideration of the outcome, instead of the implicit intention (Garcia-Molina & Clemente, 2019a). This outcome-bias in judgements has been previously studied through stories of ambiguous valences (good intention / bad

outcome) which required a more substantial contribution of mental state reasoning. However, it seems that when clues are provided or lower verbal information is demanded (e.g. using pictures to present the stories), autistic adults can base their reasoning on intentions (Bellesi, Vyas, Jameel, & Channon, 2018; Fadda et al., 2016; Margoni & Surian, 2016; Young & Tsoi, 2013; Zalla & Leboyer, 2011).

Several studies have focused on experiments involving scenarios of harm based on the combination of intentions and outcomes (good / bad), in both autistic children (Fadda et al., 2016; Grant et al., 2005; Li, Zhu, & Chen, 2018; Li et al., 2019) and autistic adults (Buon et al., 2013; Rogé & Mullet, 2011).

We must highlight the study by Grant and colleagues (2005), who assessed autistic, NT and moderate learning difficulties children with moral stories. These moral stories were concerned with two factors: intentions (neutral / good / bad) and harmful outcomes (harm to a *person* or damage to *property*). Contrary to expectations, autistic children judged the agent's culpability on the basis of his/her intention, and they also stated that harm to a person was *worse* than damage to property.

In a similar study, Rogé and Mullet, (2011), tested if autistic participants (aged from 7 to 36 years old, distributed in three groups by age) were able to judge the perpetrator of the act in MT considering two factors: intention (good / bad) and severity (categorised as: not severe, medium and severe; i.e., severe consequences: hitting someone or breaking a radio).

Although autistic participants did not perform as well as their same chronological age and developmental comparison group, they were able to use intention consistently and to recognize and distinguish more serious consequences from medium or not severe ones. Other studies that compare object *versus* person's judgements are inconsistent with Grant et al. (2005) and Rogé and Mullet (2011). For example, Li et al. 's (2019) found that autistic children evaluated more negatively, more upset and their pupils were more dilated when a

character was damaging an object than harming a person. However, children from Li and colleagues' study were younger than ours and above-mentioned studies, who may be influenced by a great object obsession in an early age. In this line, none of the named studies explored whether there was a difference in the complexity of the inferences of mental states' characters: whether the victim is (emotionally) hurt directly – personal - or through damaging property – material mediator (i.e. differences between stealing an object and someone who hurts another). This is an aspect which would be interesting to explore in depth because stories classified as having a personal mediator would imply more complex mental state inferences than those in which the mediator is material.

Therefore, autistic individuals could recognise the severity of the harm, however, when the inferences of character's intentions are difficult to understand, bad outcomes would affect their judgements. For example, in the study by Li et al. (2018), autistic children were more severe with their punishments than NT children. This pattern of 'over-blaming' (the tendency to judge the intentions of the agent as worse than they actually are) would show how autistic individuals judged an agent according to the final outcome rather than the intention. Also, in the study by Buon et al. (2013), no differences were found between groups of autistic and NT adults in questions about the suffering of the agent or the recognition of the cause of suffering. However, the most interesting result of Buon and colleagues work was related to severity, since autistic adults judged the agent who harmed the victim accidentally more severely than the NT group, even though when the intention was correctly assigned, as already was mentioned by Moran et al. (2011). These results may show that, although the information about the mental states of the characters is inferred in the task, the autistic participants do have problems when it comes to using this information and generating a socially appropriate moral judgement (Zalla, Barlassina, Buon, and Leboyer, 2011; Zalla and Leboyer, 2011). As explained above, this difficulty can be more salient in a *faux pas*

situation, due to the important role of false beliefs and intentions (mental states) in its understanding. This pattern of ‘over-blaming’ could show how autistic individuals judged an agent according to (or with great consideration for) the final outcome rather than the intention. But this could also explain a difference between situations where the perpetrator’s action occurs directly to a person *versus* where the perpetrator’s action is directed towards an object, due to autistic people possibly finding it more difficult to infer mental states in both perpetrator and victim rather than perpetrator and object (no mental states).

To sum up, little research has been carried out on the analysis of the judgements in both MT and FP situations. MT and FP tasks have usually been studied separately in different autistic population and different samples (some classical studies as, MT in NT children: Baird & Astington, 2004; FP in autistic children: Baron-Cohen et al., 1999; MT in autistic children: Grant et al., 2005; FP in autistic adults: Zalla et al., 2009; and MT in autistic adults: Zalla et al., 2011). The present research aims to use FP and MT stories to determine whether the moral reasoning of autistic children differs depending on whether the situation is deliberate or accidental. In both tasks, it has a negative outcome for the listener / victim (feeling emotionally bad or sad), but on the one hand the intention and agent was classified as bad (MT stories) and on the other, they were good (FP stories). Moreover, a novelty of the present study is that the agent’s morality is also analysed, which has an important weight to determine whether autistic individuals’ judgements are affected by the outcome, instead of only considering the effect on the intention.

Hypotheses

Three hypotheses derived from three relevant aspects: (1) ambiguous and unambiguous cases; (2) complexity of the inferences; and (3) the judgement of the agent’s morality.

First hypothesis (1a) was that autistic participants will respond mostly based on the bad outcome in both tasks. Therefore, in the MT stories (bad intention / bad outcome) a similar performance of both groups will be expected (e.g., Grant et al., 2005 and Rogé & Mullet, 2011). Nevertheless, in the FP stories (good intention / bad outcome), there will be differences between groups, and autistic participants are expected to give more wrong answers, based mostly on the outcome, than NT participants. In relation to this (hypothesis 1b), FP stories (ambiguous situations) could be more difficult to understand than MT stories (unambiguous situations) for both groups, but the autistic group will give a greater number of wrong answers, due to they will base their judgements on the outcome (Margoni & Surian, 2016; Moran et al., 2011; Zalla & Leboyer, 2011; Zalla et al., 2009).

The second hypothesis was that, when the stories are divided by the mediator, autistic group will give more wrong answers than the NT group in those stories referring to a *personal* mediator, due to the complexity of mental states involved (MT example: getting revenge against someone; FP examples: insulting the mother's character). A similar performance of both groups is expected in those stories referring to a *material* mediator (Grant et al., 2005; Li et al., 2019; Rogé & Mullet, 2011; Zalla et al., 2011; Zalla & Leboyer, 2011).

The third hypothesis is focused on the evaluation of morality of the agent. The NT group will judge the agent based on his/her intention, independent of the outcome, thus, the judgements of the agent's morality in MT (judged as bad) will differ from the FP stories (judged as good). However, the autistic group will judge the agent more severely (as bad) in both types of stories, based on the final outcome. In this sense, in the autistic group, the responses of the agent's morality will be similar in MT and FP stories (e.g., Buon et al., 2013; Li et al., 2018; Moran et al., 2011; Zalla et al., 2011; Zalla & Leboyer, 2011).

Method

Participants

The sample size was determined by an a priori sample size calculation using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) for MANOVA analysis, for mixed ANOVA within-between groups, and for *t*-test for related samples. To detect effect sizes, it was indicated that a minimum of 58 participants would be adequate (*Cohen d* = 0.25, with power set at =.95 and at $\alpha = .05$). This sample is comparable to previous studies conducted in the autism and moral field (i.e. Grant et al., 2005, or Zalla et al., 2011). Sixty-two Spanish participants from ordinary schools took part in the present study. Their ages ranged between 84 and 145 months ($M = 113.68$ months, $SD = 17.85$); and the range of intelligence, based on participants' IQ, was between 80 and 130 ($M = 104.34$, $SD = 13.43$).

Autistic group: it was formed by thirty participants (5 girls and 25 boys) diagnosed with level 1 – requiring support - Autism Spectrum Disorder, according to DSM-5 (American Psychiatric Association, 2013). All the participants with ASD were fluent in Spanish, showed capacity for conversing with and understanding others, and they had an average or a high-average IQ. Details are shown below. Each child had been previously diagnosed with autism by a qualified paediatric neurologist or psychologist from a specialised centre at the time of the study, and they all met full criteria for autism based on the Autism Diagnostic Interview–revised ([ADI-R], Rutter, Bailey, & Lord, 2003) and the Autism Diagnostic Observation Schedule – Module 3 for verbally fluent and older children ([ADOS], Lord, Rutter, DiLavore, & Risi, 2001). All the children with ASD were attending ordinary schools and receiving specific intervention from a speech therapist at school while attending ordinary classrooms. Moreover, the research group ensured they all had a typical IQ with Sattler's (1992) short version of the Wechsler scale (WISC-III) (IQ range: 80 – 130), and they all passed a first-order false belief ToM task.

NT group: it was formed by thirty-two participants (7 girls and 25 boys) matched by their chronological age and intelligence (IQ range: 80 – 130). No significant differences were found between groups on their age ($M_{ASD} = 112.73 (18.01)$, $M_{NT} = 114.56 (17.96)$, $t(60) = .40$, $p = .690$, $d = .10$), and intelligence ($M_{ASD} = 102.83 (14.23)$, $M_{NT} = 105.75 (12.69)$, $t(60) = .85$, $p = .397$, $d = .22$).

Materials

Intelligence Quotient (IQ) tests. To assess intelligence, Sattler's short adaptation (1992) of the WISC-III intelligence scale was administered (Wechsler, 1991) which was constructed from the Block Design and Vocabulary subtests. The estimated full-scale IQ was computed by: a) adding the Vocabulary plus the Block Design scaled score, and b) following the classification according to age of the table 3.19 of Sattler (1992). The total scores on this scale appeared in the report of the autistic children written by the psychologist or specialised neurologist (with a maximum of two years since the last administration). Due to the high correlation in the ASD group between the total IQ on the scores in the report and the scores on the short form (as found in the classic studies: Ryan, 1981; Sattler, 1992), the short form was administered to the NT group as a reliable estimate of the group's intelligence quotient (acceptable range of reliability, between 0.80 and 0.89), with the main objective of ensuring comparable IQ levels in both groups.

Moral Transgressions and Faux Pas stories. A set of eight stories were presented by the research group using the E-prime 3.0 software (Psychology Software Tools, 2016) for computers with a touch screen for their forced-choice responses. The stories were classified as:

- *MT task*: Four stories of intentional situations extracted from Garcia-Molina, Clemente, Andrés-Roqueta and Rodríguez (2016);

- *FP task*: Four stories of accidental situations extracted from the ‘Faux Pas’ task by Baron-Cohen et al. (1999), Spanish translation by Garcia-Molina et al. (2016).

Each story consisted of two *vignettes* (illustrations) with dialogues and narration. Both situations (MT and FP) followed the same structure: (1) speaker / perpetrator agent with his/her beliefs and desires – (2) victim or emotionally affected agent (bad outcome). Situations differed in terms of intent: agents in MT scenarios had a bad or selfish desire and they acted deliberately; however, agents in FP scenarios had a false belief and they acted accidentally. The decoding features involved in the MT and FP stories can be seen in Table 1. In addition, each story was classified by two independent raters according to whether the victim would be (emotionally) hurt (a) through damaging property: material mediator (i.e., getting revenge by hurting a person; two MT stories: the balloons, the soup stories; two FP stories: the cook, the bathroom stories) or (b) directly: personal mediator (i.e., stealing the object; two MT stories: the yo-yo, the ice-cream stories; two FP stories: the plane, the apple pie stories). See Table 1 and appendix 1 for the four stories of *person* as the mediator, and Appendix 2 for the four stories of *material* as the mediator (based on Baron-Cohen et al.’s 1999 study).

[Table 1. NEAR HERE]

Scoring of MT and FP stories

Scoring (a) explained the scores of the recognition questions in general – Hypothesis 1 and 2. Scoring (b) was used only for Hypothesis 3, in order to focus the comparison between groups in the questions related only in relation to intention and morality.

a) Scoring in recognition questions.

At the end of each story, the same set of questions was asked for the MT and FP stories, following the questions in the study by Banerjee and Watling (2005). The answers were analysed through questions related to:

- (i) *detection of the bad action or sentence*: ‘Did someone do/say something that they should not have done/said?’ (yes);
- (ii) *recognition of the perpetrator / speaker agent*: ‘If so, who?’ (perpetrator/speaker in each story – however there were two incorrect options: victim and nobody);
- (iii) *intention*: ‘Do you think X wanted to make [the victim] feel like that?’ (MT: yes / bad; FP: no / good);
- (iv) *morality of the agent*: ‘Is [the agent] good or bad?’ (MT: bad, FP: good);
- (v) *morality of the action*: ‘Was what [the agent] did right or wrong?’ (bad or wrong);
- (vi) *morality of the outcome / emotion*: ‘How does [name of the victim] feel?’ (bad or sad).

The answers were scored as in Garcia-Molina, Clemente and Andrés-Roqueta’s (2019) study: 0 = wrong; 1 = correct, with a total of 6 correct points per story, which means a total of 24 points in MT stories and another 24 in FP. The correct answer to each question is indicated in brackets after each question for the reader (Scoring a).

b) Scoring for the comparison of intention and morality between MT and FP stories.

For comparing the relevant answers in MT and FP tasks only questions related to (iii) *intention*, (iv) *morality of the agent*, (v) *morality of the action*, and (vi) *morality of the outcome* were analysed. The questions (i) *detection of the bad action or sentence* and (ii) *recognition of the perpetrator / speaker agent* were excluded for analysis because both did not provide information about intention or morality.

The value for the FP stories in questions (iii) *intention* and (iv) *morality of the agent* was changed to compare the same valences between tasks. Each answer was scored as 0 or 1 point. Since the minimum score was 0 and the maximum score was 4 in each task (four stories), an overall score closes to 4 would be bad, while a score closes to 0 would be good.

Answers were scored and expected as follows:

(iii) *intention*: 1 point = Bad; 0 = Good (in an MT story it is expected to be 1; in an FP story is expected to be 0);

(iv) *morality of the agent*: 1 point = Bad; 0 = Good (in an MT story it is expected to be 1; in an FP story is expected to be 0);

(v) *morality of the action*: 1 point = Bad; 0 = Good (in both an MT and an FP story it is expected to be 1);

(vi) *morality of the outcome*: 1 point = Bad / Sad; 0 = Good / Happy (in both an MT and an FP story it is expected to be 1).

Procedure

The study protocol was granted ethical approval from the university and the Valencian Community Research Ethics Committee, and written informed consent was obtained from each child's parents or legal guardian. NT children were chosen from the same schools as the autistic children according to previously established criteria (IQ within the typical range and no diagnosis of comorbid psychiatric disorders, learning disabilities or injury involving the brain). In the following days, intelligence and ToM tasks were administered individually, in random order, in one or two sessions of approximately an hour and a half held in quiet rooms at the schools. All the tasks were administered in Spanish. In relation to the four MT and four FP stories, stimuli (*vignettes* and audio) were presented electronically using E-Prime 3.0 software (Psychology Software Tools, 2016), which randomised the order in which the stories appeared for each participant. Children were informed that they would watch and listen to a story and be asked questions at the end. They were told to watch and listen carefully and do their best. Six questions each story with forced-choice answers were asked. As the computer had a touch screen, participants' responses were recorded via E-prime software (Psychology Software Tools, 2016). Before the administration of the MT and FP

stories, children were trained with the touch screen computer and a similar task – which was not included in the current study.

Two out of sixty-four participants were discarded because neither of them was within the established parameters for their inclusion (having an IQ > 80 and passing a first-order false belief ToM task). For the coding data, although the E-Prime software recollected the data, the main researcher of the project collected the responses in a paper sheet for comparison. In addition, two colleagues reviewed the data in the SPSS sheet.

Data analysis

The data were analysed using the statistical package SPSS 24. The sample met the assumptions of the parametric analysis, according to the Kolmogorov-Smirnov homogeneity test. It was conducted a MANOVA for hypothesis 1(a) and a mixed ANOVA within- and between- groups for hypothesis 1(b) and 2. *T*-test for related samples was used to establish differences within groups in two variables (hypothesis 3). Independent Variables (IV) and Dependent Variables (DV) are specified for better understanding of the results: in hypothesis 1(a), IV was the Group (ASD – NT) and the DVs were the combination of the 6 variables ((i) detection; (ii) recognition, (iii) intention, (iv) morality of the agent, (v) morality of the action, (vi) morality of the outcome); in hypothesis 1(b), a 2 x 2 x 6 factorial design (IV was the Ambiguity x Group) was performed with each 6 DV specified before; in hypothesis 2, a 2 x 2 x 2 factorial design (IV was the Group x Complexity) was performed and the DV were the score of the stories classified as personal or material mediator; in hypothesis 3, a *t*-test for related samples comparing the DV: (iii) good/bad intention and (iv) good/bad agent's morality between MT and FP was performed in each group (NT and ASD). The observed data are compared against Δ_L and Δ_U in *t*-tests ($\Delta_L = -.3$ to $\Delta_U = .3$). When there are no

significant differences, we assume equal variances, based on Lakens, Scheel & Isager (2018). For all the analyses, the p value used to establish the statistical significance was .05.

Results

Differences in the MT and FP tasks between autistic and NT children.

In relation to hypothesis 1(a), which predicted that the autistic group will be as able as the NT group to respond to MT stories; though less able to respond correctly to FP stories than the NT group, there is no statistically significant difference in the MT and FP tasks based on which group participants were (ASD or NT). *MT*: $F(6, 55) = .79, p = .582$; Wilk's $\Lambda = .921$, partial $\eta^2 = .08$; *FP*: $F(6, 55) = .45, p = .840$; Wilk's $\Lambda = .953$, partial $\eta^2 = .05$.

Table 2 shows means and SD.

[TABLE 2 near here]

Regarding hypothesis 1(b), when performance in both groups of stories - divided by ambiguity (FP) / unambiguity (MT) - is compared within groups, the factor Ambiguity x Group interaction effect indicated that there was not a significant interaction: $F(6, 55) = .54, p = .777$; Wilk's $\Lambda = .945$. This effect tells us that the scoring of the two tasks did not differ between both groups. Taking a closer look at the different questions, there is not a significant Ambiguity x Group interaction in (i) detection: $F(1, 60) = 1.15, p = .288$; (ii) recognition: $F(1, 60) = 2.34, p = .131$; (iii) intention: $F(1, 60) = .62, p = .434$; (iv) agent's morality: $F(1, 60) = .09, p = .768$; (v) moral of the action: $F(1, 60) = .07, p = .788$; (vi) moral of the outcome: $F(1, 60) = .00, p = .983$.

It can be observed in Table 2 that ASD group performed similarly in both the MT and FP tasks, although the lowest score was found in the FP. Within the NT group, it can be seen that their performance was clearly better in the MT task.

Differences when the mediator is personal or material.

With regard to hypothesis 2, stories were divided into *personal* mediator and *material* mediator. Stories classified as personal were hypothesised to be more difficult to judge for the ASD group than the NT group. Significant differences between groups are only observed when the mediator was personal: $F(1, 60) = 5.36, p = .024$. More concretely, autistic group performed worse than NT group in the MT stories classified as *personal* mediator ($M_{ASD} = 9.23; M_{NT} = 10.37; t(29) = 2.66, p = .010, d = .67$). See Table 3 for more information between groups. The observed effect falls within the equivalence bounds and it is close enough to zero to be practically equivalent.

[Table 3. NEAR HERE]

Differences within groups in intention and agent's morality.

Related to hypothesis 3, intention and morality responses were compared in both groups for MT and FP tasks (see scores in section b of Material). Scores are understood from good to bad (from 0 to 4) in Figures 2 and 3.

Figures 2 and 3 show similar scores between groups when good / bad judgements are considered. Taking as a reference the fact that the higher they score, the worse (bad) the intentions / agents / actions / outcomes are, both figures show similar judgements between action and outcome in both groups. A more detailed analysis (paired *t*-test) focused on (iii) intention and (iv) morality of the agent revealed several significant differences between MT and FP stories in the NT group regarding different aspects: intentions ($M_{MT} = 2.13, M_{FP} = 1.03, r = .050; t(31) = 4.16, p < .001, d = .85$), morality of the agent ($M_{MT} = 2.84, M_{FP} = 2.03, r = .000, t(31) = 4.10, p < .001, d = .58$).

In the ASD group, significant differences were only found in intentions ($M_{MT} = 1.83, M_{FP} = .90, r = .004, t(29) = 4.26, p < .001, d = .77$). No significant differences were found in the variable morality of the agent between MT and FP stories ($M_{MT} = 2.70, M_{FP} = 2.37, r = .402,$

$t(29) = 1.11, p = .277, d = .22$). The observed effect falls within the equivalence bounds and it is close enough to zero to be practically equivalent.

The variables (iii) intention and (iv) morality of the agent related to the FP stories were always judged near the 0, as ‘less bad’, than those of the MT stories.

[Figure 2 and 3. Mean scores of the NT (2) and ASD (3) groups according to their moral judgements on: intention, agent, action and outcome in MT and FP tasks. * < .05; ** < .001.

INSERT NEAR HERE ONE NEXT TO THE OTHER]

Discussion

The present study was designed to investigate differences between groups (ASD – NT) according to their ability to interpret intentions and judge morality, in two types of deliberate and accidental scenarios (MT and FP), both based on social situations of everyday life. In this line, it is important to recognise whether autistic children – close to adolescence - can have greater difficulties making judgements based on intentions, due to it happening during this period where blunders and moral transgressions are often seen in school contexts and social relationships.

Our findings showed that the ASD and NT groups obtained very similar scores in both tasks, although with subtle difficulties in the ASD group regarding the judgements of MT when the mediator was personal, and also when they make judgements about the agent’s morality in the FP task. Several important aspects of this research are specified in more detail below.

Differences between stories and groups

Our first hypothesis was that autistic participants would be able to respond to the MT stories as well as their comparison group. However, in the FP stories autistic children could have more difficulty responding to the correct answer than NT children. Our findings showed that

autistic children responded in a similar way to their comparison group (NT), thereby confirming the results of other similar studies to the current work (Grant et al., 2005; Moran et al., 2001; Rogé & Mullet, 2011). Similar research explained that autistic individuals seem to perform best on tasks that are closest to their own experiences, or in which presentation it is used a mixed modality (with visual and verbal stimulus) – as in the current study (Bellesi et al., 2018; Garcia-Molina & Clemente, 2019b). In Grant et al.'s (2015) and Fadda et al.'s (2016) studies, the good scores of the autistic group were explained by the fact that they could learn by using their own experience or if they were explicitly taught, rather than applying complex ToM reasoning. Also, in the study by Garcia-Molina et al. (2019), ASD group were as able as NT group to respond to the questions about transgressions when the answers were forced-choice, as in the present study.

Following the hypothesis 1b, when MT and FP stories were compared, no significant differences were found between MT and FP stories, although both groups obtained the slightly lowest scores in FP stories. This finding is not in agreement with the established knowledge that accidental action is more complex to understand and appears later in child development (Cushman, Sheketoff, Wharton, & Carey, 2013), due to the need to understand the false belief and intentions of the agent involved in the stories (Baron-Cohen et al., 1999; Zalla et al., 2009). What could explain these similar results in MT and FP tasks is the importance of understanding intentions in all the stories, and not only in FP stories.

As in some previous studies (e.g., Buon et al., 2013; Zalla and Leboyer, 2011; Zalla et al., 2011), in the present work, most of the intentions of the FP stories were correctly identified as 'good'. As a noteworthy point, the (iii) *intention* question of the present study may be analysed here for better understanding. The question used in this work was linked to the plan/desire to emotionally hurt another person, as stated in previous studies (Banerjee and Watling, 2005). Thus, participants from both groups were aware of the listener's emotions

(‘Do you think X wanted to make [the victim] *feel* like that?’), which would be in line with the outcome: the agent is hurt. In other studies, this question was associated only with *desires* or *motives* (‘Did the agent have the intention to harm?’) (for more examples see Buon et al., 2013, or Zalla et al. 2009), which provide specific information of intentionality and harm, and could increase children’s tendency to make outcome-based judgements.

Personal and material mediator

With regard to hypothesis 2, it was predicted that autistic children would give more wrong answers in the *personal* mediator stories than NT children. Effectively, significant differences were found between groups, but only in the MT stories classified as *personal* mediator. As expected, the fact that the mediator through which the action occurs is directed towards another *person* (for example, having the desire to take revenge on a person) may be mentally more complex to interpret than when the action is directed towards an *object* (for example, having the desire to obtain an object belonging to someone else and to steal the object), or even when someone insults/despises another person (as in the cases of *personal* FP stories considered in the present study). In addition, in the case of the *person* as a mediator, the process usually involves more painful harm than in the case of an object, and it also seems to imply the concept of *reversibility* (objects can be replaced, but hurting someone cannot be undone; Grant et al., 2005). Recently, research is pointing to the problems that autistic children and adolescents can find in MT scenarios in real life (i.e., bullying) – to recognise it and to ask for help when they are bully victims – as a result of their difficulties in social understanding (Humphrey & Hebron, 2015).

The fact that the mediator is *personal* implies two understandings of mental states. First, there is the understanding of mental states of character X (the perpetrator of the action), with specific desires and beliefs. Second, there is the understanding that character Y, who is affected by the action, also has a mind and does not know the information about X’s desires.

Regardless whether the cases are ambiguous or unambiguous, to resolve them correctly, all this mental-states information has to be integrated, which shows the complexity for autistic individuals could find in both situations analysed (Moran et al., 2001; Zalla et al., 2011; Zalla and Leboyer, 2011). This explanation is confirmed by the fact that both groups responded very similarly to the stories classified as involving a *material* mediator, in which the inference of mental states is less complex.

Intention- or outcome- based judgements and agent's morality

Regarding the hypothesis 3, both groups were expected to judge the agent based on his/her intention, instead of the outcome. Both groups obtained very similar scores in the judgements of intention and morality, as can be observed visually in the graphs. Thus, autistic children in this study understood the intention (as happened in previous studies, i.e.; Grant et al., 2005). However, they rated the agent as 'bad' even when the intention was understood to be 'good' (Buon et al., 2013; Moran et al., 2011, Zalla et al., 2009). This fact is corroborated by the non-significant differences between the MT and FP stories regarding the morality of the agent in the ASD group. As expected, important differences did exist in the NT group, since the agent of the MT stories was judged more severely ('worse') than in the FP stories. However, in the ASD group, the morality of the agent tended towards 'bad' (even in FP), and such responses may be influenced by the bad outcome (character Y was hurt) even when the agent's intention was understood. This discrepancy suggests that autistic individuals could have problems with ToM capacity, since they show outcome-based judgements when cases were ambiguous: the *agent* is *bad* because the outcome was *bad* (Margoni & Surian, 2016). Important studies as Margoni and Surian (2016), Moran et al. (2011) or Zucchelli and colleagues (2018) suggest that autistic (and autistic personality traits) individuals can base their judgements on the outcomes or side effects in accidental harm, rather than the agent's intentions. Our results may add another step in this discussion, due to the outcome-based

judgement could be influencing not only the evaluation of the intention, but also the morality of the agent. Another possible explanation is that autistic individuals have difficulties when it comes to integrating both paths required in ‘faux pas’ (paths 1 and 2) or changing one path to another (see the explanation of the Introduction). On the one hand, there is the information and the mental states attributed to the *speaker*. On the other hand, there is the information that *we* as observers have. It is well established that autistic persons have problems in social flexibility and inhibition; namely, to inhibit the information from the first path in order to be flexible and integrate the information from the second path. For this reason, autistic individuals could have problems to make a moral decision based on, first, the false belief of the speaker committing a ‘faux pas’ – path 1 – or, second, the accidental harm – path 2. Autistic participants could remain on one of the two paths, understanding the main character’s false belief or the intention. Therefore, they could have difficulties processing all this information, making a correct decision regarding the agent’s morality. This explanation would be in line with a large body of evidence showing impairments in the domain of executive functions in autistic individuals (Hill & Bird, 2006; Margoni, Guglielmetti & Surian, 2019; Ozonoff, 1997; Zelazo et al., 2002). In the study by Moran et al. (2011), autistic individuals who successfully passed a false belief task showed impairments in integrating conflicting information about mental states (neutral intention) and bad outcomes in moral judgement. They concluded that what makes it difficult for autistic people to respond to ambiguous information about intention and outcome would be the lack of a robust and fully flexible ToM, and an executive dysfunction.

Limitations

Some shortcomings regarding this study need to be acknowledged here. The first issue concerns the interpretation of our results regarding our sample, due to participants in the autistic group of our study being only level 1 ASD. Therefore, our findings should be

carefully interpreted in order to be generalisable. Also, more details of the sample are necessary, such as the educational styles of the caregivers and teachers, as participants could have received different coaching and preparation on moral rules. The novelty of this work lies in the fact that, as far as we know, no other study has measured the morality of the agent between the main measures in MT and FP stories. However, moral dilemmas and 'faux pas' tasks have been widely administered. Although this work has focused on distinguishing intentions (good and bad), this distinction has not been made in the outcomes (good / bad), and there are no *good* outcomes in any of our tasks. For future studies, neutral or good results might be considered in order to see whether the results obtained are affected. Another possible limitation of the study might be that we did not include any measures of ToM abilities which do not involve moral judgements. Also, future research is needed using specific tasks focused on moral reasoning in different situations, combined with validated measures of different components of executive functions and ToM tasks. Finally, it would have been interesting to collect responses using Likert scales ('Who do you think was guiltier?', 'How guilty?') or collect verbal information related to more variables, e.g. punishment ('Would you punish X?').

Implications and Conclusions

In sum, the present findings show that autistic children do not have severe difficulties in reasoning adequately about moral judgements of everyday situations. This similarity between autistic and comparison groups reveals interesting knowledge about the moral reasoning of autistic individuals: when scenarios are based on dialogues, narration and illustrations (a mixed modality), and the questions are asked by forced-choice answers, autistic participants can perform as well as NT participants. The subtle differences found for beliefs or intentions when autistic responses were compared to the NT group can shed some light on the difficulty that autistic people can face in their real life. The interpretation of the intention in

interpersonal relationships is a key for the prevention of serious problems with peers, such as bullying. Actually, it is during childhood and adolescence when the importance of understanding peer relations is very important and belonging to a group of friends is crucial. For this reason, these kinds of tasks - which combine morality and mental states - can be a good way to show and find some clues to helping autistic people in their real relationships. Finally, it is relevant to mention the support of the explanation in Figure 1, adapted from previous works, which contributes in a visual way to break a 'faux pas' down into aspects that play an important role in its understanding (in terms of mental states and morality). This decoding would allow the professional to understand the subtle difficulties that autistic individuals encounter in similar tasks as well as in their interpersonal relationships. The importance of the tasks which follow the design of Figure 1 remains in the difficulty of integrating the observer's (the participant) and the characters' mind for understanding the whole situation. This kind of design is a good way to build scenarios more related to real life than prototypical tasks. In a broader sense, if we can improve the design of intervention tasks and tools for autistic individuals we may be able to improve their quality of life.

References

- Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological science*, *23*(8), 861-868. doi: 10.1177/0956797611434965
- Banerjee, R., & Watling, D. (2005). Children's understanding of faux pas. *Hellenic Journal of Psychology*, *2*, 27-45.
- Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders* *29*(5), 407–418.
- Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development*, *2*(103), 37–49. doi: 10.1002/cd.96
- Bellesi, G., Vyas, K., Jameel, L., & Channon, S. (2018). Moral reasoning about everyday situations in adults with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *52*, 1-11. doi: 10.1016/j.rasd.2018.04.009
- Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *Journal of autism and developmental disorders*, *43*(2), 458-470. doi: 10.1007/s10803-012-1588-7
- Cushman, F. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, *6*, 97-103. doi: 10.1016/j.copsyc.2015.06.003
- Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*(1), 6-21. doi: 10.1016/J.Cognition.2012.11.008.

- Fadda, R., Parisi, M., Ferretti, L., Saba, G., Foscoliano, M., Salvago, A., Doneddu, G. (2016). Exploring the Role of Theory of Mind in Moral Judgment: The Case of Children with Autism Spectrum Disorder, *Frontiers in Psychology*, 7, 523, 1-8. doi: 10.3389/fpsyg.2016.00523
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Garcia-Molina, I., & Clemente-Estevan, R. A. (2019a). Moral reasoning in autistic individuals: a theoretical review/El razonamiento moral en personas con Trastorno del Espectro Autista: una revisión teórica. *Estudios de Psicología*, 40(2), 312-337. doi: 10.1080/02109395.2019.1589083
- Garcia-Molina, I., & Clemente-Estevan, R. (2019b). Autism and Faux Pas. Influences of Presentation Modality and Working Memory. *The Spanish Journal of Psychology*, 22, E13. doi:10.1017/sjp.2019.13
- Garcia-Molina, I., Clemente-Estevan, R. A., & Andrés-Roqueta, C. (2019). Detailing the understanding of moral judgements in autism. A study with Spanish-speaking children. *Cognitive Development*, 52, 100815. doi: doi.org/10.1016/j.cogdev.2019.100815
- Garcia-Molina, I., Clemente-Estevan, R.A., Andrés-Roqueta, C., & Rodríguez, L. (2016). *Intervención mediante historias complejas de teoría de la mente. Meteduras de pata, juicios morales e ironías*. Psique, 17. Castelló de la Plana: Universitat Jaume I. doi: 10.6035/Psique.2016.17
- Grant, C.M., Boucher, J., Riggs, K.J., & Grayson, A. (2005). Moral understanding in children with autism. *Autism* 9, 317–331. doi: 10.1177/1362361305055418

- Humphrey, N., & Hebron, J. (2015) Bullying of children and adolescents with autism spectrum conditions: a 'state of the field' review. *International Journal of Inclusive Education, 19*(8), 845-862, doi: 10.1080/13603116.2014.981602
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442-462. doi: 10.1177/1362361315588200
- Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*(2), 197-215.
- Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: patterns of performance in a multiple case series. *Neuropsychologia, 44*(14), 2822–2835
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science, 1*(2), 259-269. doi: 10.1177/2515245918770963
- Li, T., Decety, J., Hu, X., Li, J., Lin, J., & Yi, L. (2019). Third-party sociomoral evaluations in children with autism spectrum disorder. *Child development, 90*(5), e584-e597. doi: 10.1111/cdev.13206
- Li, J., Zhu, L., & Chen, Z. (2018). The association between punishment and cooperation in children with high-functioning autism. *Journal of experimental child psychology, 171*, 1-13. doi: 10.1016/j.jecp.2018.01.015
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2008). *Autism diagnostic observation schedule: ADOS manual*. Los Angeles: Western Psychological Services.
- Margoni, F., Guglielmetti, G. & Surian, L. (2019). Brief Report: Young Children with Autism Can Generate Intent-Based Moral Judgments. *J Autism Dev Disord 49*, 5078–5085. doi: 10.1007/s10803-019-04212-9

- Margoni, F., & Surian, L. (2016). Mental State Understanding and Moral Judgment in children with Autistic Spectrum Disorder. *Frontiers in Psychology, 7*, 1478. doi: doi.org/10.3389/fpsyg.2016.01478
- Martin, J. W., & Cushman, F. (2016). Why we forgive what can't be controlled. *Cognition, 147*, 133-143. doi: 10.1016/j.cognition.2015.11.008
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological science, 19*(6), 549-557. doi: 10.1111/j.1467-9280.2008.02122.x
- Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences, 108*(7), 2688-2692. doi: 10.1073/pnas.1011734108
- Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 179–211). New York: Oxford University Press.
- Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from <https://www.pstnet.com>
- Rogé, B., & Mullet, E. (2011). Blame and forgiveness judgements among children, adolescents and adults with autism. *Autism, 15*(6), 702-712. doi: 10.1177/1362361310394219
- Rutter, M., Le Couteur, A., & Lord, C. (2003). ADI-R: The Autism Diagnostic Interview—Revised. *Los Angeles, CA: Western Psychological Services, 29*, 30.

- Ryan, J. J. (1981). Clinical utility of a WISC-R short form. *Journal of clinical psychology, 37*(2), 389-391.
- Sattler, J.M. (1992). *Assessment of children: WISC-III and WPPSI-R supplement*. San Diego: Jerome M. Sattler.
- Vivanti, G. (2019). Ask the Editor: What is the most appropriate way to talk about individuals with a diagnosis of autism? *J Autism Dev Disord.* doi:10.1007/s10803-019-04280-x
- Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children: Manual*. Psychological Corporation.
- Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47*, 2065–2072. doi: 10.1016/j.neuropsychologia.2009.03.020.
- Young, L., & Tsoi, L. (2013). When mental states matter, when they don't, and what that means for morality. *Social and Personality Psychology Compass, 7*(8), 585-604. doi: 10.1111/spc3.12044
- Zalla, T., Barlassina, L., Buon, M., & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition, 121*(1), 115-126.
- Zalla, T., & Leboyer, M. (2011). Judgment of intentionality and moral evaluation in individuals with high functioning autism. *Review of Philosophy and Psychology, 2*(4), 681-698. doi: 10.1016/j.cognition.2011.06.004
- Zalla, T., Sav, A. M., Stopin, A., Ahade, S., & Leboyer, M. (2009). Faux pas detection and intentional action in Asperger Syndrome. A replication on a French sample. *Journal*

of Autism and Developmental Disorders, 39(2), 373-382. doi: 10.1007/s10803-008-0634-y

Zelazo, P. D., Jacques, S., Burack, J., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant and Child Development*, 11(11), 171–195.

Zucchelli, M. M., Nori, R., Gambetti, E., & Giusberti, F. (2018). The influence of high autistic personality traits on the attribution of intentionality in typically developing individuals. *Journal of Cognitive Psychology*, 30(8), 840-853. doi: 10.1080/20445911.2018.1530241

Appendix 1. Personal mediator

The main character has avenged or harmed a *person* directly.

Example of a MT story (personal mediator): ‘The soup story’. See Figure 4.

"This is Diego, who wants to watch to his favorite cartoons, but his mother has to go to the doctor and asks him to prepare a soup for his little sister Luna. So, Diego cooks the soup. Luna loves soup and she wants to try it right away. Diego sees how Luna takes the spoon, but he keeps quiet and does not let her know that the soup is still burning, and his little sister burns her tongue".

INSERT FIGURE 4 NEAR HERE.

Figure 4. The soup story (personal mediator)

Example of a FP story (personal mediator): ‘The cubicles story’. See Figure 5.

“Irene was in one of the cubicles in the toilets at school. Teresa and Lara were at the sinks nearby. Teresa said «Do you know that new girl in the class?», «Yes, her name is Irene.

Doesn't she look really weird!» —replied Lara. Irene then came out of the cubicles. Teresa said «Oh, hello Irene, do you want to come and play rounders?».”

INSERT FIGURE 5 NEAR HERE.

Figure 5. The cubicles story (personal mediator)

Appendix 2. Material mediator

Example of a MT story (material mediator): ‘The yo-yo story’. See Figure 5.

The main character harmed a person through an object (i.e., he/she wanted the object).

“This is Luis and this is Iris. Luis has a yo-yo and he is playing with it because it is playtime. Iris is his classmate, and asks him to play a little. The bell rings, and all children have to go. Also, Luis. But Iris hides the yo-yo in her school bag to play later at home”.

INSERT FIGURE 6 NEAR HERE.

Figure 6. The yo-yo story (material mediator)

Example of a FP story (material mediator): ‘The apple pie story’ (adapted from Baron-Cohen et al. (1999). See Figure 6.

“Inés helped her mum make an apple pie for her uncle when he came to visit. She carried it out of the kitchen. “I made it just for you”, said Inés. “Mmm”, replied Uncle Tomás, “That looks lovely. I really love pies, except for apple, of course!”

INSERT FIGURE 7 NEAR HERE.

Figure 7. The apple pie story (material mediator)

Table 1. Decoding features involved in MT (moral transgression) and FP (faux pas) stories

Stories	Decoding					Mediator	
	Agent			Outcome		Complexity	
	Belief	Intention	Moral	Emotion	Moral	High	Low
MT	True	Bad	Bad	Sad	Bad	Personal	Material
FP	False	Good	Good	Sad	Bad	Personal	Material

MT = Moral transgression stories; FP = Faux Pas stories

Table 2. Means and SD of ASD and NT group of all the variables divided by MT and FP tasks.

	ASD (30)	NT (32)
MT	M (SD)	M (SD)
(i) detection	3.60 (.77)	3.81 (.47)
(ii) recognition	3.47 (.82)	3.81 (.47)
(iii) intention	1.83 (1.31)	2.13 (1.31)
(iv) moral agent	2.70 (1.44)	2.84 (1.46)
(v) moral action	3.83 (.38)	3.91 (.53)
(vi) moral outcome	3.50 (.68)	3.72 (.52)
FP		
(i)detection	3.43 (.97)	3.47 (.66)
(ii)recognition	3.23 (.82)	3.19 (.93)
(iii)intention	3.10 (1.09)	2.97 (1.28)
(iv)moral agent	1.63 (1.56)	1.97 (1.33)
(v) moral action	3.57 (.63)	3.59 (.76)
(vi)outcome	3.43 (.77)	3.66 (.70)

MT = Moral transgression task; FP = Faux pas task; M = mean; SD = Standard Deviation. Maximum score each variable = 4.

Table 3. Differences of means between groups (*t*-test) divided by *personal* or *material* mediators.

Mediator	ASD (30)		NT (32)		t	<i>p</i>	<i>d</i> ^a	SE(<i>r</i>) ^b
	M (SD)	Range	M (SD)	Range				
MT								
Personal	9.23 (1.74)	5-12	10.37 (1.64)	6-12	-2.66	.010	.67	.32
Material	9.70 (1.74)	6-12	9.84 (1.87)	3-12	-.313	.756	.07	.03
FP								
Personal	8.57 (1.89)	5-12	9.09 (1.98)	5-12	-1.07	.290	.27	.13
Material	9.83 (1.42)	7-12	9.75 (1.67)	6-12	.212	.833	.05	.02

MT = Moral transgression task; FP = Faux Pas task; M = mean; SD = Standard Deviation; *d*^a = Cohen's *d*; SE(*r*)^b = Size Effect













