'What's new for you?':

Interlocutor-specific perspective-taking and language interpretation in autistic and neuro-typical children

Kirsten Abbot-Smith[1],

David Williams[1],

Danielle Matthews[2]

[1]School of Psychology, University of Kent, UK

[2]Psychology, University of Sheffield, UK

Word count: 6196

Conflict of interest: The first and second authors received a small internal grant from the Faculty of Social Sciences at the University of Kent. Both declare that they have no conflict of interest. The third author received no funding and also declares no conflict of interest.

Contact:
Dr Kirsten Abbot-Smith
Kent Child Development Unit
School of Psychology,
University of Kent,
Keynes College
Canterbury CT2 7NP, UK
Email: K.Abbot-Smith@kent.ac.uk

Abstract

Background: Studies have found that children with Autism Spectrum Disorder (ASD) are more likely to make errors in appropriately *producing* referring expressions ('the dog' vs. 'the black dog') than are controls but *comprehend* them with equal facility. We tested whether this anomaly arises because comprehension studies have focused on manipulating perspective-taking at a 'generic speaker' level.

Method: We compared 24 autistic eight- to eleven-year-olds with 24 well-matched neuro-typical controls. Children interpreted requests (e.g. 'Can I have that ball?') in contexts which would be ambiguous (i.e. because the child can see two balls) if perspective-taking were not utilized. In the interlocutor-specific perspective-taking condition, the target was the particular object which was new for the speaker.
Children needed to take into account what the speaker had played with before and the fact that they were now expressing excitement about something new. In two control 'speaker-generic' conditions we tested children's ability to take the visual perspective of the speaker (where any speaker who stood behind a particular barrier would have the same perspective).

Results: The autistic group were significantly less likely to select the target and significantly more likely to request clarification in the 'interlocutor-specific' condition. Performance in the 'interlocutor-generic' (visual) perspective taking conditions did not differ between groups.

Conclusion: Autistic children, even those who are not intellectually-impaired, tend to have more difficulty than neuro-typical peers in comprehending referring expressions when this requires understanding that people comment on what is new for them.

Introduction

One of the two key symptom domains for Autism Spectrum Disorder (ASD) encompasses deficits in verbal social communication (DSM-5, APA, 2013). A significant hurdle to studying this atypicality is that, while we can readily notice it in everyday conversation, it is challenging to create experimental conditions that reliably elicit language which is appropriate (or not) for a given context. One type of language use that is so pervasive as to be open to experimental study is *reference*. Referring to something verbally requires selecting an expression (e.g., "the dog", or "it" or "the dog over there with the red collar") that is sufficiently, but not excessively, informative. Deficits in selecting appropriate referring expressions in *production* are characteristic of ASD (see Malkin, Abbot-Smith & Williams, 2018, for a systematic review). That is, autistic individuals frequently fail to tailor language for specific interlocutors (conversation partners) (Volden, Magill-Evans, Goulden & Clarke, 2007). In contrast, studies that have tested the *comprehension* of referring expressions have not observed significant ASD-specific difficulties with the ability to use the perspective of the interlocutor to interpret referring expressions. While this discrepancy is surprising, it is plausibly explained by the fact that studies of reference comprehension in ASD to date have only manipulated fairly simple forms of perspective-taking. It is therefore possible that autistic children experience difficulties in comprehending referring expressions when comprehension requires interlocutor-specific perspective-taking – i.e., consideration of mental content (knowledge, interests) relating to a specific individual and how this differs from the mental content of others. The purpose of the current study was to test this possibility with a reference interpretation task that required participants to consider what was 'new' to the interaction from the interlocutor's perspective.

To understand the importance of this manipulation it is necessary to first consider the socio-cognitive abilities that are likely to be required in everyday interaction when interpreting referring expressions used by another speaker. For example, if your partner asks you to 'pass the screwdriver', when you can see two screwdrivers, to successfully understand which screwdriver is meant, you need to consider your partner's perspective. In some sense, the listener needs to step into the speaker's shoes in order to interpret which referent (here: which screwdriver) is intended. Most theorists agree that some type of Theory of Mind or Mentalising process, at least in a broad sense, must be involved in interpreting referring expressions in these types of situations (Apperly & Butterfill, 2009; Clark & Marshall, 1981). There are many potential dimensions to Mentalising. One dimension concerns whether the child needs to consider another's affect (or emotional stance towards something) versus whether the child needs a cognitive understanding of the other individual (e.g. understanding what the individual does or does not know). Another dimension to Mentalising concerns the degree to which the child needs to consider whether various individuals might have differing perspectives (i.e. interlocutor-specific perspective-taking).

Individuals with ASD tend to perform significantly less well on Theory of Mind measures than do neuro-typical peers (e.g. White, Hill, Happé & Frith, 2011). Moreover, difficulties with a dynamic consideration of the perspective of others in real time in more naturalistic situations (Peterson, Garnett, Kelly & Attwood, 2009) are strong predictors of social functioning (Berenguer, Miranda, Colomer, Baixauli & Rosello, 2018; Jones et al., 2018). Yet, four out of the five studies of referring expression interpretation to date have found no significant differences between individuals with ASD and matched neuro-typical controls (TDs) (Begeer, Malle, Nieuwland, & Keysar, 2010; Malkin, Abbot-Smith, Williams & Ayling, 2018; Santiesteban, Shah, White, Bird, & Heyes, 2015; Volden, Mulcahy, & Holdgrafer, 1997)[1]. One possible reason for this is that most of these studies only

[1] The fifth study – by Schuh, Eigsti and Mirman (2016) - only found impairments relative to typical controls for their gaze fixations measure (and not for object selection accuracy). Crucially, even for the gaze fixations measure, the difference to typical controls was in fact subject to a three-way interaction between group, perspective-taking condition and working memory, whereby the Autistic group were more impaired than the neuro-typical group when required to use perspective-taking under high working memory load conditions. Thus, the impairment in the Autistic group did not appear to lie in perspective-taking *ability*, but rather they were more susceptible to the *performance* limitations

manipulated interlocutor-generic perspective-taking; that is, the perspective of the speaker would be the same, regardless of the identity of the specific speaker. A good example of why we consider these tasks interlocutor-generic is 'the director task', which is frequently used to assess how listeners interpret referring expressions (e.g. 'the duck' vs. 'the big duck'). In this task, the participant and speaker (director) sit on either side of a grid containing various objects, some of which are occluded from the director's view. Thus, when the director asks the participant to (for example) 'Pick up the duck', the participant needs to consider whether the director has visual access (or not) to both the ducks visible to the participant. Importantly, *any* interlocutor positioned in the director's seat would have the *same* perspective in this task - the participant is not required to consider the potential differences between individuals in terms of their past experiences of or affect towards certain objects.

In contrast, in many situations in everyday life, individuals need to consider the interlocutor-specific perspective of a speaker in order to successfully interpret referring expressions. For example, in a scenario in which you and your partner can both see two screwdrivers and she asks you to 'Pass me the screwdriver', the statement itself is underspecified (i.e. it would not, in itself, allow you to identify the referent). However, if you have both previously established that only one of the two screwdrivers is suitable for the flatpack furniture you are constructing, then determining the referent would be straightforward. This type of perspective-taking requires you to consider knowledge which is specific to your partner by virtue of their past experiences and the common ground you have built up with them; you would not draw the same conclusion regarding referential intent if your partner were somehow replaced by another individual with no prior experience of constructing flatpack furniture. Thus, to utilise interlocutor-specific perspective, a listener needs to consider individual differences in cognitive content (e.g. remembering what the partner does or does not know – either through

imposed by working memory.

verbal transmission or by tracking his / her experience) or individual differences in affective stance (i.e. what certain people like or are interested in) (see Moll & Kadipasaoglu, 2013, for a similar discussion).

To our knowledge, there is only one study to date that aimed to examine the role of *interlocutor-specific* perspective-taking in the interpretation of referring expressions by children with and without ASD (Malkin et al., 2018b: Study 1). In Malkin et al.'s (2018b) study, speaker perspective was manipulated in the following way. First, the participant partially constructed one toy (e.g. a woodpecker) with one experimenter (in the absence of the second experimenter), then the participant partially constructed a different toy (e.g. a telephone) with a second experimenter (in the absence of the first experimenter). Then, one of the two experimenters handed the participant a missing piece (e.g. string) that could complete either toy and told her 'Now you can do it'. Success in the task required the child to track the experience of *specific* interlocutors; the target referent for the ambiguous pronoun 'it' was the one which was 'old' or 'given' information for the speaker. Autistic children were just at good at this task as neuro-typical controls.

Malkin et al.'s (2018) task certainly required participants to track the interlocutor-specific experiences and to trace the pronoun that the interlocutor used back to the discourse that was specific to that particular interlocutor. This is interesting, since this anaphor resolution must be carried out in an interlocutor-specific manner. However, it is possible that children did not really need to consider how their own perspective differed from that of another individual in Malkin et al.'s task, which is still fairly simple compared to many of the real world perspective-taking situations that children regularly encounter. It is certainly the case that (amongst other factors) Malkin et al.'s task did not require the participants to consider the affective component of interlocutor-specific perspective-taking. This is important as affect interpretation plays such a frequent role in reference interpretation and perspective-taking more generally in everyday life. There are two routes by which a listener might be able to determine an individual affect in everyday situation. The first simply requires interpreting the individual's bodily cues to their affect. However, since individuals very often conceal (or attempt to conceal) these bodily cues in everyday life, another important route to determining likely affect is through simulation (e.g. Harris, 1992). That is, if I know dislike it when Fred calls me names, then I can fairly safely assume that Tom also dislikes it when Fred calls him names.

In the current study, we examined the ability to utilise interlocutor-specific perspective through this simulation route. That is, if I know that I tend to be more interested in toys that I have never seen before, then I can assume that the speaker is also more likely to be interested in a toy that she has never seen before. To investigate this, we adapted a paradigm developed by Tomasello and Haberl (2003) and Moll and Tomasello (2007). In the Moll and Tomasello (2007) paradigm, the child first jointly engages with the Requester (R) with one novel object and then a second novel object. At this point, R leaves the room and the child jointly engages with a second experimenter with a third novel object (the target). Finally, R returns to the room and says from the doorway '*Wow! Cool! Give it to me!*' whereby the child and R can see three novel objects. Typically-developing toddlers are above chance in interpreting the referring expression '*it*' as referring to the object that is new for the requesting experimenter (Requester).

In contrast to the study by Malkin et al. (2018b), in this 'excitement at the new' paradigm, the child needs to understand that R is interested in the object that is new for R and, thus, likely to comment on this (see Clark, Schreuder & Buttrick, 1983). The ability to interpret R's affect and the distinctive prosodic contour employed when commenting on new and exciting things would probably also play a role in this task. In sum, successful performance in this task may require the integration of several processes including an understanding of interlocutor-specific experience (and how this might result in one object thus appearing more salient), an understanding that others tend to comment on the new as well as the ability to interpret affect and prosody. The integration of these elements with the mechanisms of reference interpretation is a much more complex process than any previously manipulated in a reference interpretation task with individuals with ASD. However, these are the types of processes likely to be frequently required in interlocutor-specific perspective-taking in spontaneous real-life interaction (see Graham, San Juan & Khu, 2017).

In the current study, we tested autistic children and neuro-typical controls aged between eight and eleven years, matched on age, non-verbal IQ, receptive language, and gender. In our adaptation of Moll and Tomasello (2007), we told each child (C) that one experimenter (E2) had bought toys that the Requester (R) had not yet seen. For each trial, E2 passed one of these (e.g., pink ball) over to R, who discussed this with C. Then R left and E2 showed C another object of the same lexical type (e.g., yellow ball). When R returned, she and C could see both objects. R said "*Oh wow, I like that ball. Can you put that ball in my box?*". The key dependent variables were object choice (i.e. whether the child selected the object that was new for R) and number of clarification questions (as a measure of uncertainty). Participants also completed two interlocutor-generic perspective-taking control conditions. As for the experimental (interlocutor-specific) conditions, the two interlocutor-generic conditions also involved a choice between two objects of the same lexical type (e.g. two cars) for each trial but required the child to take the interlocutor's visual perspective into account. The first such interlocutor-generic condition aligned with previous studies in the literature in requiring the child to utilise information about *what* the speaker could perceive. The second interlocutor-generic condition required the child to utilise information about *how* the speaker perceived each object (which colour they perceived it to be).

Research questions

Research Question 1: *Are autistic children impaired relative to typical controls in their ability to take the 'excitement at the new' into account when interpreting referring expressions?*

We predicted that autistic children would select the correct object less often and would produce more clarification questions than would typical controls in the interlocutor-specific condition. In contrast, we predicted that there would be no significant between-groups differences on the interlocutor-generic control conditions for either dependent variable.

Research Question 2: *Are autistic difficulties with taking 'excitement at the new' into account related to impairments in affect recognition?*

Success in the Tomasello and Haberl (2003) / Moll and Tomasello (2007) paradigm could plausibly depend in part on the ability of children to recognise excitement on the part of the Requester. Autistic children have previously been found on average to have difficulty with affect recognition (e.g. Golan, Baron-Cohen & Golan, 2008). We therefore assessed all participants on a (standardised) non-verbal measure of affect recognition (NEPSY II, Korkman, Kirk & Kemp, 1998), which has been widely used in the ASD emotion recognition intervention literature (e.g. Williams, Gray & Tonge, 2012) and on which autistic children have been found to score on average lower than their typical peers (e.g. Loukusa, Maekinen, Kuusikko-Gauffin, Ebeling & Moilanen, 2014). We investigate, first, whether there are significant between-groups differences on the NEPSY II affect recognition measure and, second, whether there is a relationship between performance on this measure and our experimental task.

Method

Participants

We tested 60 children aged between eight and eleven years, of which 32 were neuro-typical and 28 were autistic. From this sample, we excluded eight neuro-typical children because their IQ scores were too high to allow them to be matched and one child because his parent did not return the SRS questionnaire. We also commenced testing with but excluded four autistic children, two because they scored outside the typical range on either non-verbal IQ or receptive language and two because they became too distressed / non-compliant for testing.

Our final sample thus consisted of 48 children with 24 in each diagnostic group. Of the neuro-typical group 15 were recruited through the Kent Child Development Unit database and the rest via three mainstream primary schools. Of the autistic group, eight attended specialist provisions for autistic children, nine were recruited via a local county autism support service, six were recruited via the Kent Child Development Unit database and one was recruited via a mainstream primary school. Children in the autistic group had been diagnosed within the British National Health Service either by a clinical psychologist or a paediatrician. Parents of children in both groups completed the Social Responsiveness Scale (SRS, Constantino & Gruber, 2007). T-scores of 76 or higher are in the severe autism range.

Autistic and neuro-typical children were matched on chronological age, core language and non-verbal IQ. To assess core language, we carried out a receptive language test, namely the *Following Directions* sub-test, from the Clinical Evaluation of Language Fundamentals® - Fifth Edition (CELF®-5, Wiig, Semel & Secord, 2013). To assess non-verbal IQ, we carried out the *Matrix Reasoning* subtest from the Wechsler Abbreviated Scale of Intelligence II (Wechsler, 2011). The cognitive profiles of the final sample are shown in Table 1 below. Written consent was obtained from a parent of each individual child. Each individual child also gave his or her verbal assent.

Table 1. Cognitive profile of the children who were included.

	Autistic (n = 24; 18 males)	**Neuro-typical (n = 24; 18 males)**		
	Mean (SD)	Mean (SD)	*p*	*d*
Chronological Age in Months	120.04 (13.87)	118.13 (14.91)	.65	0.13
Receptive language scaled score[1]	10.13 (2.88)	10.71 (2.70)	.47	0.21
Non-verbal IQ T-score[2]	46.58 (8.92)	49.42 (6.87)	.22	0.36
Social Responsiveness Scale T-score	88.13 (4.22)	45.92 (7.02)	< .001	7.40
Theory of Mind Composite [3]	3.52 (1.13)	4.27 (.93)	.02	0.72

1. *Following Directions* sub-test of CELF®-5 (Wiig, Semel & Secord, 2013)
2. *Matrix Reasoning* sub-test of the WASI II (Wechsler, 2011)
3. This consisted of one point for the first order change of location task (Baron-Cohen, Leslie & Frith, 1985), three points for the 'Birthday Surprise' test of second order false belief understanding (Sullivan, Zaitchik & Tager-Flusberg, 1994) and two points for the 'Kittens' (white lie) vignette from Happé's (1994) 'Strange Stories' (max. = 6 points). See Appendix B for details.

Procedure and measures

Overall Procedure

The Requester (R) wore dark sunglasses during the study. On first meeting the child, R explained (showing the child a large box in which there were already toys) that she was busy organising a play event and would therefore have to often leave the room. The experimental trials were presented as 'breaks' from the standardised (non-verbal IQ, core language, NEPSY affect recognition) and Theory of Mind tests (Baron-Cohen et al., 1985; Happ**é,** 1994; Sullivan et al., 1994), and were interspersed between them. The order of administration of the standardized measures was: Receptive language, Non-verbal IQ, NEPSY II affect recognition and finally the Theory of Mind tests, which are outlined in the Appendix. After the test of receptive language, we counterbalanced across participants within each group whether children experienced an experimental (interlocutor-specific) trial first or a control (interlocutor-generic) trial first.

Affect recognition:

To assess affect recognition, we used the NEPSY II standardized measure (Korkman et al., 1998). Initially, the participant is shown for each item a picture at the top of a page of a child with the facial affect of a basic emotion (happiness, anger, sadness, neutral, fear or disgust). The participant is asked to select from a series of faces at the bottom of the page the child (different child) 'who feels the same way as the child at the top'. For some items, the participant has to select two target faces. For other items, the participant has to retain the affect in short term memory while selecting targets.

Experimental task: interlocutor-specific perspective-taking 'What is new for you?'

The child sat behind a table facing an open door. E2 sat at the short end of the table to the left of the child. R sat facing the child. In the interlocutor-specific perspective-taking condition, R first stated that E2 had just bought some toys that even R had never seen. E2 then pulled out a box and took out the first item (e.g. a Spiderman whose head turns around) and put it in the centre of the table whilst ensuring that neither R nor the child could see the remaining items in the box. R made a least two comments about the object to the child. These two comments were semi-scripted (e.g. R might say '*Oh, a spiderman! He's very bendy. Oh look – his head turns around like an owl*') but if the child made spontaneous comments about the object, R followed in on the child's comments. Comments were always about the specific object and not about that type of object in general. R allowed the child to handle the object and they engaged in joint attention for around 10-20 seconds. E2 then put the first object back into the box and retrieved a second object of the same lexical type (e.g. a Spiderman holding a web shooter). R and the child again engaged in joint attention for 10-20 seconds focussed on the second object, again making two to three semi-scripted comments. For example, R might say "*This spiderman seems to be holding a web shooter. Oh look! If you squeeze his legs like this, you can make the web shooter turn round*". During the handling of both objects, E2 did not comment and R was interested but not excited.

R then made an excuse (e.g. '*oh, I forgot I need to quickly speak to Mr X. Back in a minute*') and left the room, whereupon E2 pushed the second object to a corner of the table and brought out a third object of the same lexical type (e.g. a Spiderman with transparent legs and power boosters). Child and E2 engaged in joint attention with the third object for 10-20 seconds. This was again partially scripted, as for R's comments about the other two objects. For example, E2 might say '*This looks like a water-filled Spiderman. He has power boosters on his feet. Would you like to hold him?*". E2 then ensured that the two objects (foil vs. target) were each in one corner of the table, to the left and right of the child, and coughed to signal that R could re-enter the room. R's affect on appearing in the doorway was surprised and delighted. Neither object was occluded from R's view. From the doorway she said "*Wow! Look at THAT [TOY NAME]! Can you put that [TOY NAME e.g. ball] in my box?".* Since R stood equidistant from the two potential referents and was wearing dark glasses, this utterance was always ambiguous if the child did not take the prior shared experience with R into account.

The target was the object that R had not previously seen (i.e. was 'new' for R) and the foil was the object with which the child had jointly engaged in attending to with R. There were three trials per participant. Every trial where the child selected the target object was given a score of one. If the child selected the foil object (i.e. the object with which the child and R had previously jointly engaged), this was given a score of zero. If the child picked up both objects and put them both into the box, this was also given a score of zero. If the child was already touching an object when R walked into the room, this trial was coded as 'unscoreable' and hence missing data. This occurred on 0.7% of trials. For this reason proportion scores were used for the object-selection dependent variable. Children sometimes also made clarification requests (e.g. '*Which ball?*'). These were transcribed and the frequency with which they were made was taken as a measure of uncertainty.

Each child experienced three experimental trials. Target object location was counterbalanced so it was on the left-hand side of the child two out of three times for half of each group (and one out of three times for the other half of each group). We also counterbalanced across participants which object of a pair (e.g. flashing ball vs. spiky ball) was the target. There were six pairs of possible objects. The particular three pairs used for a particular participant was counterbalanced across participants. All objects were familiar to this age group in the UK at the time of testing, and were found in pilot testing to usually be of interest to autistic children.

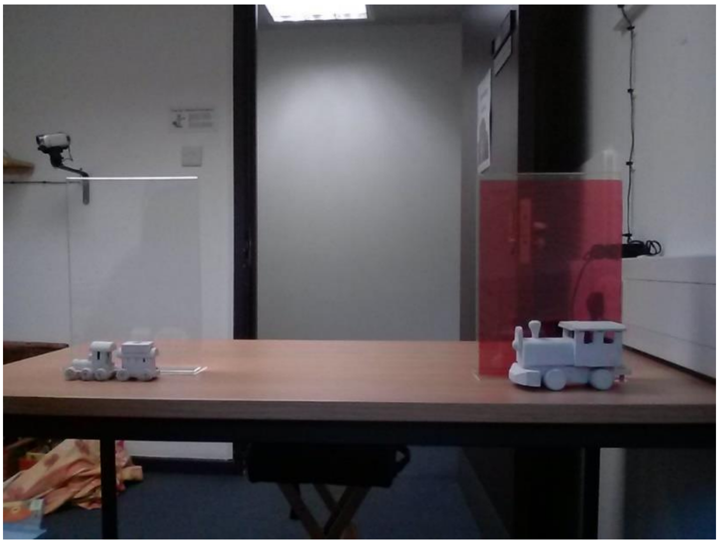
Interlocutor-generic perspective-taking conditions

Prior to each interlocutor-generic trial, R made an excuse (e.g. '*Oh, I think I've forgotten my keys in the cloakroom. Back in a second.*') and left the room. In the interlocutor-generic 'what-perceived' condition, E2 then placed two objects in front of the child. An opaque cardboard occluder was placed in front of one object so that it could not be seen from R's perspective. When R reappeared in the doorway, she was initially looking in her box of toys. She then looked up and said in an offhand manner '*Ah, I've been looking for that [TOY NAME, e.g. ball]*. Since R stood equidistant from the two potential referents and was wearing dark glasses, this utterance was always ambiguous if the child did not take into account that one object was hidden from R's perspective. The target was the non-occluded object. The particular three object pairs assigned to a participant was counterbalanced across participants. Some of the object pairs overlapped with the experimental condition and some overlapped with the other interlocutor-generic perspective-taking condition outlined below.

In the second interlocutor-generic perspective-taking condition, the child needed to take into account *how* the interlocutor perceived the objects that, in reality, were always white. In these trials, E2 placed two objects in front of the child, whereby a transparent colour filter occluder (along the lines of that used in Moll and Meltzoff, 2011) was placed in front of one object so that it appeared to be a different colour (red, yellow or blue) from R's perspective. An example trial is shown in Fig 1 below. When R reappeared in the doorway, she was initially looking in her box of toys. She then looked up in a distracted manner and said in an offhand manner '*Ah, I've been looking for that ADJECTIVE [TOY NAME, e.g. blue balloon]*. For one of the three trials per participant, the colour adjective used was 'white' and thus the non-filtered object was the target. For the other two interlocutor-generic 'how-perceived' trials per participant, the colour adjective specified the filtered object. Since R stood equidistant from the two potential referents and was wearing dark glasses, this utterance was always ambiguous if the child did not take into account that one object was not white from R's perspective. For all interlocutor-generic 'how-perceived' trials, the target was the object that matched the colour from R's perspective.

Across the six control trials (three for interlocutor-generic-'what' and three for interlocutor-generic-'how' perspective-taking) the target object side was counterbalanced within and across participants. The specific object of a pair that was a target was also counterbalanced across participants. As for the experimental condition, every control trial where the child selected the target object was given a score of one. Similarly, if the child selected the foil object or both objects and put them both into the box, this was also given a score of zero. As for the experimental condition, if the child was already touching an object when R walked into the room, this trial was coded as ‘unscoreable’. Two additional situations occasionally occurred in the control conditions, resulting in unscoreable trials: 1) if the child pulled both objects out from behind the screens before the test question was asked or 2) if the child indicated that he or she was revealing information to R (e.g. by moving the object from behind the opaque barrier and saying, for example, ‘*there’s this one as well*’). Over the interlocutor-generic trials, 3% of trials were unscoreable for object choice and thus we used proportion scores. In addition to object choice, - as for the experimental condition - we also scored the usage of clarification questions in the control conditions.

Fig 1: View from the perspective of the participant (here: in an interlocutor-generic condition)



Post-test compliance control task

Because we were worried that autistic children might differ from neuro-typical controls at the group level in terms of their desire to comply with experimenter requests per se, at the end of the entire testing session, each participant participated in three ‘compliance control’ trials. In this trials, they were asked to put one of a pair of objects (different lexical types e.g. robot vs. helicopter) into the box. All participants scored 100% on compliance.

Reliability

The first author coded all the trials for object choice. A second coder scored object choice for 77% of the dataset. Agreement was very good between the two coders (Cohen's $k = .88$). The first author also transcribed and counted all clarification requests (e.g. '*Which one*?') and a second coder coded eight of the 48 children (i.e. 17% of data), with excellent inter-rater agreement (ICC = .952, $p < .001$).

Results

Research Question 1: *Are autistic children impaired relative to neuro-typical controls in taking the 'excitement at the new' into account when interpreting referring expressions?*

Control interlocutor-generic perspective-taking measures

Table 2 shows the performance of autistic and neuro-typical participants in the two interlocutor-generic perspective-taking control tasks. There were no between-groups differences for either object choice or frequency of clarification questions for either the 'what' or 'how' interlocutor-generic perspective-taking tasks. Table 2 shows that all effect sizes for between-groups differences were small in magnitude. Both groups of participants selected the correct object at levels significantly above chance, with effects that were large in magnitude, all ts > 4.99, all ps < .001, all ds > 1.02. Success on these tasks among autistic participants fits with previous literature examining how listeners utilise information about what the speaker sees when interpreting referring expressions (Begeer et al., 2010; Santiesteban et al., 2015; Volden et al., 1997). In the context of the current study, it suggests that any significant between-groups difference in our experimental condition cannot be explained by demands of engaging with a referential communication task.

Table 2: Between-groups comparisons for the interlocutor-generic perspective-taking control measures

	Autistic	**Neuro-typical**		
	Mean (SD)	Mean (SD)	*p*	*d*
WHAT: object choice (proportion correct)	.79 (.26)	.87 (.26)	.27	0.13
HOW: object choice (proportion correct)	.79 (.28)	.90 (.17)	.10	0.47
WHAT: clarification requests (raw frequency)	.83 (1.09)	.58 (1.21)	.46	0.22
HOW: clarification requests (raw frequency)	1.29 (2.79)	.88 (1.94)	.55	0.17

Experimental 'excited about the new' condition

Autistic participants (M = .63, SD = .39) selected the correct object (i.e., the one that was new for the Requester) significantly less frequently than did neuro-typical participants (M = .81, SD = .23), ($t(23) = 2.06$, $p < .05$, $d = 0.57$). The neuro-typical children selected the target object at well above chance levels, associated with a large effect size ($t(23) = 6.77$, $p < .001$, $d = 1.38$), which shows that they assumed that R was excited about and commenting on the object that was *new* for her. In contrast, the performance of the autistic group did not differ from chance, $t(23) = 1.59$, $p = .13$, $d = 0.32$), indicating that as a group autistic children did not necessarily make this assumption. The same results were found for our clarification request measure. The autistic group asked significantly more clarification questions in our 'excited about the new' experimental task (M = 1.67, SD = 1.50) than did the neuro-typical group (M = 0.79, SD = 1.10), $t(23) = 2.31$, $p < .05$, $d = 0.67$, indicating that they showed greater uncertainty as to R's referential intent. In sum, we find support for our hypothesis that autistic children struggle, relative to neuro-typical controls, to consider what is new for the interlocutor when interpreting referring expressions.

Table 3: Between-groups comparison for the experimental ‘excitement at the new’ condition

	Autistic	*Neuro-typical*		
	Mean (SD)	**Mean (SD)**	***p***	***d***
Object choice (proportions)	.63 (.39)	.81 (.23)	.045	0.57
Clarification requests (raw frequency)	1.67 (1.50)	.79 (1.10)	.026	0.67

Research Question 2*: *Are autistic difficulties in taking ‘excitement at the new’ into account related to difficulties with affect recognition?

One plausible reason why the autistic children differed significantly from well-matched controls on our experimental task is that success may have required the ability to interpret R’s excited affect. The between-groups difference in affect recognition on the NEPSY II was not significant ($t(46) = 1.40$, $p = .17$, $d = 0.40$). The mean scaled scores for both groups surrounded the population mean of 10 (autistic M = 9.79, SD = 2.52; neuro-typical M = 10.83; SD = 2.67). More importantly, we examined whether performance on the NEPSY task correlated with performance in the experimental (interlocutor-specific perspective-taking) condition. This was not the case for the autistic group, nor for the neuro-typical group, nor for the groups conflated (see Table 5 below for effect sizes). Thus, it is unlikely that affect recognition explains the group differences in the comprehension of reference.

Table 4. Performance on affect recognition

	Autistic (n = 24; 18 males)	**Neuro-typical (n = 24; 18 males)**		
	Mean (SD)	Mean (SD)	*p*	*d*
NEPSY II Affect Recognition	9.79 (2.52)	10.83 (2.67)	.17	0.40

Secondary analyses

Performance in the experimental (interlocutor-specific perspective-taking) task was also not correlated for the autistic group with either age, non-verbal IQ, receptive language, affect recognition, autistic symptoms or Theory of Mind). The same is true for the neuro-typical group.[2] See Table 5 below for all relationships with the experimental measure dependent variables, by group.

[2] The same pattern of results was found if tau was used rather than Pearson's r.

Table 5: Correlations with the ‘excitement at the new’ dependent variables, by group

	Autistic		*Neuro-typical*		*Groups Combined*	
	Object choice	**Clarification Questions**	**Object choice**	**Clarification Questions**	**Object choice**	**Clarification Questions**
NEPSY Affect Recognition raw	-.06	-.04	-.12	.29	-.05	.06
Theory of Mind composite	.02	.00	.03	.21	.12	-.04
Core language (CELF) raw	-.14	.04	.07	.20	-.04	.06
IQ (Matrices) raw score	.03	-.28	.11	.15	.10	-.18
SRS questionnaire raw score	.09	.31	.16	-.05	-.24$^{\forall}$	.36*
Age in months	-.05	-.20	.15	-.06	.00	-.10

$^{\forall}$ = p < .1 * = p < .05 ** = p < .01 *** = p < .001

General Discussion

The current study was the first to test whether autistic children have difficulty interpreting referring expressions in a tasks that goes beyond requiring basic, interlocutor-generic perspective taking. In the current study, the participants had to consider which element was new to the discourse from the speaker's perspective (given shared past experience) in order to correctly interpret otherwise ambiguous use of referring expressions (i.e. 'that ball' when both speaker and listener could see two balls). We compared 24 eight- to eleven-year-olds autistic children with 24 well-matched neuro-typical controls. For both dependent variables (object choice and clarification questions), we found significant differences between the groups, indicating that while neuro-typical children assumed that the Requesting Experimenter was excited about and referring to the object that was new for her, the autistic group was less likely to do so. In contrast, for the interlocutor-generic perspective-taking control tasks, there were no significant between group differences for either dependent variable. This aligns with the previous four reference interpretation studies with autistic individuals (Begeer et al., 2010; Santiesteban et al., 2015; Schuh et al., 2016; Volden et al., 1997).

Accounting for autistic difficulties with interlocutor-specific perspective-taking

This raises the question of what exactly constitutes the difficulty that autistic individuals had in our 'excitement about the new' experimental task. A number of non-mutually exclusive possibilities warrant further exploration. To succeed on our task, one needs to 1) retrieve a memory about a *specific interlocutor's* experience, 2) understand the *emotional tone* of the test question and emotional affect of the speaker and 3) understand that people *comment on new things*. Malkin et al. (2018b) already established autistic children can retrieve a memory about a specific interlocutor's experience. Indeed, all bar one of our participants passed a First Order False Belief test, which is accepted as being developmentally subsequent to Knowledge-Access (Pratt & Bryant, 1990). However, either of the other points may be challenging for a child on the autism spectrum, as might juggling all these constraints at once.

We think it unlikely that the difficulties of the autistic group can be reduced to basic facial affect recognition. First, There were no group-level differences in non-verbal affect interpretation as assessed by the standardised NEPSY task. Second, there was no relationship between the autistic group's performance in the interlocutor-specific condition and their NEPSY affect recognition scores. However, while there is no clear-cut evidence that autistic individuals are necessarily impaired in behavioural measures of affect recognition, review papers tend to find much clearer evidence of affect recognition impairments when using electrophysiological, eye-tracking or brain imaging measures (e.g. Harms et al., 2010). Thus, one possibility which needs further exploration is whether autistic children find it more difficult than do neuro-typical children to integrate an interpretation of emotion with the interpretation of reference. A second possibility for future investigation is whether difficulties interpreting the emotional tone of the speaker's request might have contributed to diminished performance in our autistic group (although a number of studies suggest comparable performance for autistic children and neuro-typical controls when using prosody to interpret pragmatic intent, see e.g. Wang, Lee, Sigman & Dapretto, 2016).

One aspect of our interlocutor-specific perspective-taking task which probably did pose difficulties for autistic children is the need to be sensitive to the fact that the requester was more likely to comment on the object that was new (for her). That is, in Clark et al.'s (1983) terms, the object which is new to the discourse is assumed to be more salient. This is taken for granted by neuro-typical children in the earliest stages of language acquisition (e.g., Akthar, Carpenter & Tomasello, 1996; Bannard, Rosner & Matthews, 2017; Tomasello &

Akthar, 1995) and even pre-linguistically (Moll & Tomasello, 2007; O'Neill & Happé, 2000). Indeed, the tendency of neuro-typical language users to comment on elements which are new to the discourse is so strong that certain patterns are typologically marked (e.g. DuBois, 1987, see also Chafe, 1976). However, it is not something that is demonstrated by preschool autistic children (O'Neill & Happé, 2000).

One possibility here is that if an autistic individual is not him or herself necessarily more interested in elements which are new to the discourse, then he or she should find it more difficult than a neuro-typical children to simulate that their interlocutor finds the new object more salient. Another possibility is that some autistic children do themselves find new objects more interesting but do not necessarily realise that people tend to comment on new rather than given elements in the discourse. Either of these possibilities would have profound consequences for spontaneous conversation and might in fact be related to impairments in conversation skills more generally, including a tendency to monologue on favourite topics (e.g. Nadig, Lee, Singh, Bosshart & Ozonoff, 2011). It is therefore critical that future studies attempt to tease apart the source of this difficulty.

Limitations and considerations for future research

We did not find any correlations between our Theory of Mind measures, on the one hand, and with the ability to interpret referring expressions on the experimental task, on the other hand; not even in the experimental condition requiring interlocutor-specific perspective-taking, which arguably requires mentalising of some sort. This might because each Theory of Mind task and each of our experimental perspective-taking tasks assesses different aspects of mentalising. That is, it may be that children master each aspect of social cognition in a fairly piecemeal way. An alternative test, such as Reading the Mind in the Eyes (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) for example, might correlate with interlocutor-specific perspective-taking.

One limitation of the current study is that we were unable to control the prosody of the test questions absolutely, because they were uttered ‘live’ by the Requester. A second limitation is that due to our sample size of 48, we were only able to compare the autistic and neuro-typical children at the group level.

Summary and conclusions

In sum, the current study is the first to demonstrate empirically that autistic children are have greater difficulties than do neuro-typicals in the interpretation of referring expressions. We argue that the reason this is the first such demonstration is that the aspect of perspective-taking manipulated in our task requires a deeper level of consideration of the mental states than the perspectives and degree of requisite cue interpretation manipulated in previous studies. Future research is needed to investigate which particular components of interlocutor-specific perspective-taking pose difficulties for autistic individuals.

Acknowledgements

The authors are grateful to the families, schools and children that participated, including those recruited via our developmental unit database and via a local autism charity. Thank-you to Annabel Carter, Sophie Darroch, Emma Freeman, Jolie Keemink, Lauren Jenner, Lucy Pettifor and Nicola Vince for assistance with data collection. Portions of this work were presented at the International Society for Autism Research Annual Conference, Rotterdam, Netherlands, 9th – 12th May 2018 and at 14th International Congress for the Study of Child Language, Lyon, France, 2017. This study was partially funded internally by the University of the first author via a small grant from the Faculty for Social Sciences.

Ethical Standards

This study was partially funded internally by the University of the first author via a small grant from the Faculty for Social Sciences. There was no external funding and no grant number.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study (see 'participants' section for details).

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*(2), 635-645 doi: 10.1111/j.1467-8624.1996.tb01756.x

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?. *Psychological Review*, *116*(4), 953. doi 10.1037/a0016923

Bannard, C., Rosner, M., & Matthews, D. (2017). What's Worth Talking About? Information Theory Reveals How Children Balance Informativeness and Ease of Production. *Psychological science, 28*(7), 954-966. doi: 10.1177/0956797617699848

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21(1), 37-46. doi.10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*, 241–251.

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning ASD and typically developing controls. *European Journal of Developmental Psychology, 7*(1), 104-122. doi:10.1080/17405620903024263

Berenguer, C., Miranda, A., Colomer, C., Baixauli, I. & Rosello, B. (2018). Contribution of theory of mind, executive functioning, and pragmatics to socialization behaviors of children with high-functioning Autism. *Journal of Autism and Developmental*

Disorders, 48(2), 430–441. doi: 10.1007/s10803-017-3349-0
Chafe,W. (1976) Giveness, contrastiveness, definiteness, subjects, topics and points of view. In Li, C.N. (Ed.). Subject and topic. P. 25-55. New York: Academic Press.
Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webher, & I. A. Sag (Eds.), *Elements of discourse understanding*. Cambridge: Cambridge University Press.
Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground at the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*(2), 245-258.
Charman, T., Jones, C. R., Pickles, A., Simonoff, E., Baird, G., & Happé, F. (2011). Defining the cognitive phenotype of autism. *Brain research, 1380*, 10-21. doi. 10.1016/j.brainres.2010.10.075
Constantino, J. N., & Gruber, C. P. (2007). *Social responsiveness scale (SRS)* Western Psychological Services Los Angeles, CA.
Dahlgren, S., & Dahlgren Sandberg, A. D. (2008). Referential communication in children with autism spectrum disorder. *Autism: The International Journal of Research and Practice, 12*(4), 335-348. doi:10.1177/1362361308091648
DuBois, J. (1987). The discourse basis of ergativity. Language, 63: 805-855.Frascarelli, M. (2000). The syntax-phonology interface in focus and topic constructions in Italian. Dordrecht: Kluwer Academic Publishers
Fukumura, K. (2015). Development of audience design in children with and without ASD. *Developmental Psychology*, doi:10.1037/dev0000064
Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'Reading the Mind in Films' Task [Child Version]: Complex Emotion and Mental State Recognition in Children with and without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders, 38*: 1534-1541. DOI 10.1007/s10803-007-0533-7
Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental disorders*, *24*(2), 129-154. doi:10.1007/BF02172093
Harms, M., Martin, A., Wallace, G. (2010). Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. *Neuropsychological Review, 20,* 290-322. doi: 10.1007/s11065-010-9138-6 ·
Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind & Language*, *7*(1-2), 120-144. doi:10.1111/j.1468-0017.1992.tb00201.x
Jones, C. R., Simonoff, E., Baird, G., Pickles, A., Marsden, A. J., Tregay, J., ... & Charman, T. (2018). The association between theory of mind, executive function, and the symptoms of autism spectrum disorder. *Autism Research*, *11*(1), 95-109.
Khu, M., Chambers, C. & Graham, S. (2018). When you're happy and I know it: *Child Development DOI: 10.1111/cdev.12855*
Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment.* Chicago: Psychological Corporation.
Landa, R. (2000). Social language use in Asperger Syndrome and high-functioning autism. In Klin, A., Volkmar, F. & Sparrow, S. (Eds.). *Asperger Syndrome*. pp 125- 155. New York / London: The Guildford Press
Loukusa, S., Maekinen, L., Kuusikko-Gauffin, S., Ebeling, H., & Moilanen, I. (2014). Theory of mind and emotion recognition skills in children with specific language impairment, autism spectrum disorder and typical development: group differences and connection to knowledge of grammatical morphology, word-finding abilities and working memory. *International Journal of Language and Communication Disorders,*

49(4): 498-507. doi: 10.1111/1460-6984.12091

Malkin, L., Abbot-Smith, K. & Williams, D. (2018a). Is referring expressions impaired in autism spectrum disorder? A systematic review. *Autism and Developmental Language Impairments, 3:* 1-24. doi: 10.1177/2396941518763166

Malkin, L., Abbot-Smith, K., Williams, D., & Ayling, J. (2018b) When do children with Autism Spectrum Disorder take common ground into account during communication? *Autism Research* 10.1002/aur.2007

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience*, 7, 558. doi:10.3389/fnhum.2013.00558

Moll, H. & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology, 43*(2): 309-317. doi: 10.1037/0012-1649.43.2.309

Moll, H. & Meltzoff A. (2011). How Does It Look? Level 2 Perspective-Taking at 36 Months of Age. *Child Development, 82*(2), 661-673. http://dx.doi.org/10.1111/j.1467-8624.2010.01571.x

Nadig, A., Seth, S., & Sasson, M. (2015). Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology, 6.* doi:10.3389/fpsyg.2015.01888

Nadig, A., Vivanti, G., & Ozonoff, S. (2009). Adaptation of object descriptions to a partner under increasing communicative demands: A comparison of children with and without ASD. *ASD Research,* (2), 334-347. doi:10.1002/aur.102

Nadig, A., Lee, I., Singh, L., Bosshart, K. & Ozonoff, S. (2011). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia,* 48(9), 2730-2739. doi: 10.1016/j.neuropsychologia.2010.05.020

O'Neill, D.K. & Happé, F. G. (2000). Noticing and commenting on what's new: differences and similarities among 22-month-old typically developing children, children with Down syndrome and children with autism. *Developmental Science, 3*(4), 457-478. doi:10.1111/1467-7687.00139

Pearson, A., Ropar, D. & Hamilton, A. (2013). A review of visual perspective-taking in autism spectrum disorder. *Frontiers in Human Neuroscience, 7*, article 652, doi: 10.3389/fnhum.2013.00652

Pearson, A., Marsh, L., Ropar, D., & Hamilton, A. (2015). Cognitive Mechanisms underlying visual perspective-taking in typical and ASC children. *Autism Research, 9*(1): 121-130. doi: 10.1002/aur.1501

Peterson, C.C., Garnett, M., Kelly, A., & Attwood, T. (2009). Everyday social and conversation applications of theory-of-mind understanding by children with autism-spectrum disorders or typical development. *European Child and Adolescent Psychiatry, 18*, 105–115

Pratt, C. & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973-982 doi: 10.1111/j.1467-8624.1990.tb02835.x

Santiesteban, I., Shah, P., White, S., Bird, G., & Heyes, C. (2015). Mentalizing or submentalizing in a communication task? Evidence from ASD and a camera control. *Psychonomic Bulletin & Review, 22*(3), 844-849. doi:10.3758/s13423-014-0716-0

Schuh, J. M., Eigsti, I. M., & Mirman, D. (2016). Discourse comprehension in autism spectrum disorder: Effects of working memory load and common ground. *Autism Research, 9*(12), 1340-1352. doi. 10.1002/aur.1632

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*(3), 395-402. doi:10.1037/0012-1649.30.3.395.

Tomasello, M. & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development, 10*(2), 201-224. doi:10.1016/0885-2014(95)90009-8

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, *39*(5), 906-912. doi:10.1037/0012-1649.39.5.906

Volden, J., Magill-Evans, J., Goulden, K. & Clarke, M. (2007). Varying language register according to listener needs in speakers in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*(2): 1139-1154. doi: 10.1007/s10803-006-0256-1

Volden, J., Mulcahy, R. F., & Holdgrafer, G. (1997). Pragmatic language disorder and perspective-taking in autistic speakers. *Applied Psycholinguistics, 18*(2), 181-98. doi:10.1017/S0142716400009966.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain*, 129, 932–943. doi: 10.1093/brain/awl032

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II).* San Antonio, TX: NCS Pearson.

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the Strange Stories: Revealing Mentalizing Impairments in Autism. *Child Development*, *80*(4), 1097-1117. doi:10.1111/j.1467-8624.2009.01319.x

Wiig, E.H., Semel, E., & Secord, W. (2013) *Clinical Evaluation of Language Fundamentals – Fifth Edition (CELF-5)*. San Antonio, TX: NCS Pearson Inc.

Williams, B., Gray, K. & Tonge, B. (2012). Teaching emotion recognition skills to young children with autism: a randomized controlled trial of an emotion training programme. *Journal of Child Psychology and Psychiatry, 53*(12): 1268-1276. doi: :10.1111/j.1469-7610.2012.02593.x

Appendix

Theory of Mind Composite

Our Theory of Mind composite was administered in a stepwise manner (see Charman et al., 2011, for a similar procedure). We first administered a test of second order Theory of Mind developed by Sullivan, Zaitchik & Tager-Flusberg (1994), accompanied by pictures, in which a mother has a false belief about the mental state of the son (i.e. she doesn't know that he knows that she has bought him a puppy for this birthday). At the end of the story the tester asks the child three questions about what the mother will tell another person, if she is asked whether her son knows what he is getting for his birthday. If the child failed all of the above second-order false belief test questions, the test administrator administered a first order 'change of location' false belief tasks (Baron-Cohen, Leslie & Frith, 1985). If, however, the child passed two out of three second order false belief questions, he or she was credited with passing first order false belief. Finally, we administered one item ('Kittens') from Happé (1994), in which the child is told a story and then asked why a protagonist lied. The answer is scored on a three-point scale in terms of the child's ability to explain that the protagonist wishes to persuade another person to do something. We also used Happé's (1994) physical inferencing burglar and mouse story as a control for inferencing ability per se. Only one typically-developing child and three autistic children failed this measure.