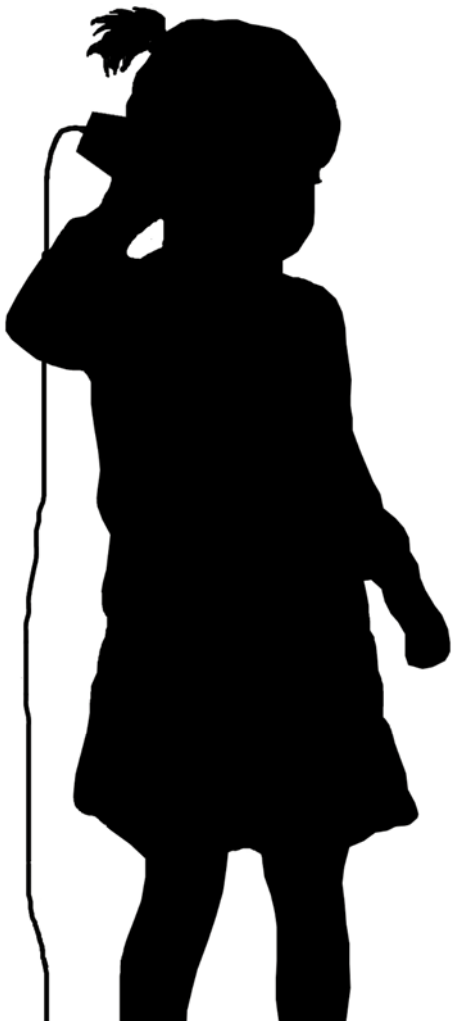
Communication abilities of children with ASD and ADHD

Kuijper, Sanne Joanna Maria

5

Pronoun processing in children with ASD and ADHD: the role of Theory of Mind, inhibition, and working memory

This chapter has been submitted as:
Kuijper, S. J. M., Hartman, C. A., & Hendriks, P. (2015). Pronoun processing in children with ASD and ADHD: the role of Theory of Mind, inhibition, and working memory.



Abstract

In language acquisition, the interpretation of object pronouns (e.g., *him*) is typically delayed compared to the interpretation of reflexives (e.g., *himself*). The current study provides insights into the mechanisms underlying this delay in pronoun interpretation. Children between 6 and 12 years old with autism spectrum disorder (ASD; n=47), attention-deficit/hyperactivity disorder (ADHD; n=36) or typically developing (TD; n=38) were tested on their interpretation and production of pronouns and reflexives and on Theory of Mind, response inhibition, and working memory. It was found that both Theory of Mind and inhibition were associated with performance on pronoun interpretation. Furthermore, children with ASD, children with ADHD and typically developing children demonstrated a similar delay in pronoun interpretation. These findings support an explanation of pronoun interpretation in terms of perspective taking: a hearer needs to take into account the grammatical perspective of the speaker for the correct interpretation of pronouns and therefore needs sufficient Theory of Mind and inhibition skills. The findings further suggest that children with ASD and children with ADHD do not have more problems than TD children in taking into account the grammatical perspective of a generic speaker, despite their known difficulties in perspective taking with specific conversational partners in specific situations.

Introduction

A fundamental aspect of children's language acquisition is learning what the linguistic expressions in their language refer to. Proper names (e.g., *John*) generally have a fixed reference. In contrast, personal pronouns (e.g., *he, she, him, her*) and reflexives (*e.g., himself, herself*) depend on other words in the sentence or the discourse for their interpretation. For instance, in the example "Bruce got upset when John accidentally hit him" the pronoun *him* refers back to the subject of the previous sentence, Bruce. The fact that *him* cannot refer back to the subject of the same sentence, John, indicates that not only the linguistic discourse, but also grammatical principles play a role (Chomsky's binding principles, 1981): pronouns in direct object position cannot refer back to the subject of the same clause, whereas reflexives in the same position must refer back to the subject (e.g., *himself* must refer back to the subject *John* in "Bruce laughed when John accidentally hit himself").

Because of their dependence on grammatical factors as well as the linguistic discourse, pronouns are expected to be relatively hard to acquire. However, this does not seem to hold for children's *production* of these forms: in typically developing children, pronouns already emerge between the age of 2;0 and 3;0 (Rozendaal, 2008 for Dutch; Bloom, Barss, Nicol, & Conway, 1994 for English). Reflexives appear somewhat later (not before age 3;3; Rozendaal, 2008). This developmental pattern is consistent with their frequency of use: children use pronouns much more frequently (in 2-5 % of the utterances; Bloom et al., 1994) than reflexives (in less than 0.5 % of the utterances; Bloom et al., 1994). Although children produce pronouns earlier and more frequently than reflexives, the order of acquisition in *interpretation* is the other way around: children's interpretation of pronouns in languages such as English and Dutch seems to be delayed in comparison to their interpretation of reflexives. For example, in a situation in which two referents are present in the discourse, 4- and 5-year-old children incorrectly allow the pronoun in sentences like (1) to have the same reference as the subject (coreferential interpretation). At the same time, they correctly allow a coreferential interpretation for reflexives such as in (2) (Chien & Wexler, 1990; Koster, 1993; Philip & Coopmans, 1996; Spenader, Smits, & Hendriks, 2009; Van Rij, Van Rijn, & Hendriks, 2010). Only around the age of 10 or 11 years, children approach an adult-like level of pronoun comprehension (Başkent, van Rij, Ng, Free, & Hendriks, 2013; Koster, 1993). This phenomenon is known as the Pronoun Interpretation Problem.

(1) The elephant is hitting him.
(2) The elephant is hitting himself.

In the linguistic literature, various explanations have been put forward for the Pronoun Interpretation Problem. Chien and Wexler (1990) argue that children possess the relevant grammatical knowledge required for a mature interpretation of pronouns and reflexives, but still lack the pragmatic skills to determine in which contexts an exception to the grammatical principles of pronoun usage is allowed. Chien and Wexler (1990) refer to accidental coreference as a source of confusion for children. Accidental coreference occurs when the object pronoun and the subject of the sentence accidentally refer to the same individual (as *he* and *him* do in "That must be John. At least he looks like him", (Chien & Wexler, 1990, p.256). This is only possible in certain (rare) contexts. Chien and Wexler (1990) argue that children have pragmatic difficulties to distinguish between contexts in which accidental coreference is permitted and contexts in which it is not. However, under the view that children's errors with pronouns are due to their confusion about accidental coreference, children should also show problems in pronoun production and use pronouns to express a coreferential meaning in all contexts, so also in contexts in which accidental coreference is not allowed. However, children's production of pronouns in an experimental setting is found to be adult-like (De Villiers, Cahillane, & Altreuter, 2006; Matthews, Lieven, Theakston, & Tomasello, 2009; Spenader et al., 2009). This makes an explanation in terms of lack of pragmatic skills highly unlikely.

A more recent explanation of the Pronoun Interpretation Problem attributes this problem to children's insufficient working memory capacity (Reinhart, 2011). According to Reinhart, pronouns can receive their interpretation in two different ways: by syntactic binding and by pragmatic coreference. If the grammar allows two interpretational possibilities, the process of reference-set computation is required (Reinhart, 2006, 2011). Reference-set computation compares the different structures and their interpretations, and discards an interpretation if there is a more economical way to obtain that interpretation. Adults use reference-set computation to block pragmatic coreference between an object pronoun and the local subject. Reinhart (2011) claims that children have insufficient working memory to perform this costly computation and therefore resort to guessing in their interpretation of pronouns. Reference-set computation does not apply in production, since the speaker already knows which meaning he or she wants to express. Therefore, children's production of pronouns is predicted to be adult-like (Reinhart, 2006).

Chien and Wexler's pragmatic account and Reinhart's processing account explain the Pronoun Interpretation Problem as the result of children's problems with peripheral aspects of the grammar. In contrast, Hendriks and Spenader (2006) argue that the Pronoun Interpretation Problem is caused by core properties of the grammar itself. They argue that the grammar consists of direction-sensitive constraints, which may yield a different outcome in comprehension than in production. To achieve successful communication in spite of potentially different outcomes in production

and comprehension, perspective taking is required. When encountering a pronoun, adult listeners block coreference between an object pronoun and the subject because they take into account the perspective of a generic speaker: if a speaker would have wanted to express a coreferential interpretation, the speaker would have used a reflexive instead of a pronoun. Children up to the age of 5 are not yet able to take into account the perspective of the speaker in their interpretation and hence show the Pronoun Interpretation Problem. Grammatical perspective taking may require Theory of Mind (Hendriks, 2014). Furthermore, it may require inhibition, since the listener must block the coreferential meaning in order to select the correct non-coreferential meaning for the pronoun. Inhibition has previously been linked to second language acquisition (e.g., Bialystok, Martin, & Viswanathan, 2005; Green, 2013; Yoshida, Tran, Benitez, & Kuwabara, 2011), but whether inhibition also plays a role in first language acquisition is yet unclear.

In the present study, we investigate these three possible underlying mechanisms of pronoun comprehension and production, i.e., working memory, Theory of Mind, and inhibition. We designed our study in such a way that we maximized the variation in mechanisms as well as outcome measures by including children with autism spectrum disorder (ASD), children with attention-deficit/hyperactivity disorder (ADHD) and a group of typically developing (TD) children in our sample. Problems in Theory of Mind have been frequently reported in children with ASD (e.g., Baron-Cohen, Tager-Flusberg, & Cohen, 1993) and sometimes in children with ADHD (Buitelaar, Van der Wees, Swaab-Barneveld, & Van der Gaag, 1999; Kuijper et al., 2015). Furthermore, working memory problems and problems in inhibition have been reported in children with ADHD and children with ASD (e.g., Geurts, Verté, Oosterlaan, Roeyers, & Sergeant, 2004; Hill, 2004; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005; Nydén, Billstedt, Hjelmquist, & Gillberg, 2001).

Besides difficulties with Theory of Mind, working memory, and inhibition, both children with ASD and children with ADHD show problems with language. Pragmatic problems are among the core deficits of ASD (American Psychiatric Association, 2000). Compared to pragmatic deficits, less is known about grammatical problems in ASD and ADHD. Some studies did not find grammatical impairments in children with ASD (Bartolucci, Pierce, Streiner, & Eppel, 1976; Tager-Flusberg, 1981). In contrast, other studies found evidence for grammatical impairments in (some) children with ASD (Eigsti, Bennetto, & Dadlani, 2007; Kjelgaard & Tager-Flusberg, 2001). These results indicate that there is heterogeneity in grammatical impairments in ASD (for an overview, see Boucher, 2012). In ADHD, language deficits are not part of the diagnosis. However, pragmatic use of language is often found to be impaired (for an overview, see Green, Johnson, & Bretherton, 2013). Most studies that investigated grammatical impairment in ADHD did not find impairments in children with ADHD (e.g., Geurts & Embrechts, 2008; Geurts et al., 2004; Helland, Biringer, Helland, & Heimann, 2012;

Kim & Kaiser, 2000). In contrast, other studies found evidence for grammatical impairments in children with ADHD (Oram, Fine, Okamoto, & Tannock, 1999; Papaeliou, Maniadaki, & Kakouros, 2015).

Although the findings on grammatical impairments of children with ASD and ADHD are equivocal, it may well be that children with ASD or ADHD experience a greater delay in pronoun interpretation than typically developing children. Recent findings of Perovic, Modyanova, and Wexler (2013) showed that high-functioning children with ASD and TD children demonstrated similar difficulties in their comprehension of pronouns. A recent study investigating referential perspective taking showed that, compared to TD children, children with ADHD had no extra difficulties in taking into account the speaker's perspective when they had to act upon instructions like "take the big duck" in situations in which the speaker did not have access to the same visual information as the child (Nilsen, Mangal & Macdonald, 2013). Pronoun interpretation has not been investigated in children with ADHD yet. Thus, in addition to our main aim of investigating possible mechanisms underlying the Pronoun Interpretation Problem, our study will also yield further insight into possible difficulties in grammatical perspective taking in both children with ASD and ADHD.

In our study we focus on children in the age range of 6 to 12 years, as in this age range in TD children the Pronoun Interpretation Problem gradually decreases (Başkent et al., 2013). Therefore, we expect most variation in pronoun interpretation performance in this age range. To investigate possible mechanisms underlying the interpretation of pronouns, we administer Theory of Mind tasks, a working memory task, and an inhibition task. Following Hendriks and Spenader (2006), pronoun interpretation is expected to be associated with Theory of Mind and inhibition. Alternatively, following Reinhart's (2011) account, pronoun interpretation is hypothesized to be associated with working memory.

Methods

Participants

In total 126 Dutch-speaking children were tested (51 with ASD, 37 with ADHD, and 38 TD children), ranging in age from 6;1 to 12;10 (M=9;1, SD=1;9).

ASD

Children in the ASD group were diagnosed with Autistic Disorder (N=10), PDD-NOS (N=34) or Asperger's Disorder (N=7) by independent clinicians on the basis of the DSM-IV-TR criteria (American Psychiatric Association, 2000). Additional inclusion criteria were that the children had an FSIQ above 75 and verbal communication skills. Furthermore, both the Autism Diagnostic Interview Revised (ADI-R) (Rutter, Le Couteur, & Lord, 2003) and the Autism Diagnostic Observation Schema (ADOS) (Lord,

Rutter, DiLavore, & Risi, 1999) were administered by trained psychologists. Children in this study were included in the ASD group if they met ADOS criteria for autism or ASD and/or ADI-R criteria for autism or ASD (cf. Risi et al.'s ASD2 criteria, 2006). Three children from the ASD group were excluded from further analysis because they did not meet these criteria, leaving 48 children in the ASD group. To document the extent to which ADHD symptoms were present, the Parent Interview for Child Symptoms (PICS) (Ickowicz et al., 2006) was administered. Seven children in the ASD group scored above the ADHD cut-offs on the PICS (see Table 1). In line with their clinical ASD diagnosis, we included these children in the ASD group.

ADHD

Children in the ADHD group were diagnosed with Combined type (N=19), Predominantly Hyperactive-Impulsive type (N=12) or Predominantly Inattentive type (N=6) by independent clinicians on the basis of the DSM-IV-TR criteria (American Psychiatric Association, 2000).. Furthermore, both the PICS (Ickowicz et al., 2006) and the TTI-IV (Tannock et al., 2002) were administered by trained psychologists. Six children with ADHD lacked TTI information. Four of them already scored above the cut-off for ADHD based on parent information alone. The remaining two children scored 1 point below the cut-off for ADHD. Since these children scored comparable on the PICS to the other children in the ADHD group (for whom TTI scores combined with their PICS scores reached the cut-off), we included them in the analyses. Seven children in the ADHD group scored within ASD criteria on the ADOS or ADI-R (see Table 1). In line with their clinical diagnosis, we included these children in the ADHD group.

TD

Children in the TD group had not been diagnosed with ASD or ADHD. The ADOS, ADI-R and PICS were administered by trained clinicians in this group as well. None of the children scored above the cut-offs for ASD or ADHD described above.

Materials

Background variables

IQ was assessed by two subtests (Vocabulary and Block Design) of the Dutch Wechsler Intelligence Scale for Children (WISC-III NL) (Kort et al., 2002). Verbal ability was assessed by the Dutch version of the Peabody Picture Vocabulary Test-III (PPVT) (Schlichting, 2005). Group means and standard deviations for age, IQ, PPVT, and clinical interviews, can be found in Table 1.

Comprehension

To test the comprehension of pronouns and reflexives, we carried out a Picture Verification Task. Children saw one picture at a time. The picture showed two animals

Table 1 Mean Scores (Standard Deviations) of age, clinical interviews, WISC-III, PPVT, Theory of Mind task, n-back task, Stop task per Participant Group

	ASD (*n* = 47)		**ADHD** (*n* = 36)		**TD** (*n* = 38)		Group differences (Bonferroni corrected post hoc analyses)
	M	(SD)	M	(SD)	M	(SD)	
% Male	87		83		66		
Age	9;3	(1;10)	8;9	(1;7)	9;0	(1;9)	n.s.
ADI-R[1]							
Social Interaction	16.40	(6.06)	4.58	(4.10)	1.82	(3.09)	ASD***>ADHD>TD*
Communication	12.62	(4.38)	4.03	(2.68)	1.34	(1.55)	ASD***>ADHD>TD**
Stereotyped Behavior	4.40	(2.59)	1.42	(1.56)	0.32	(0.66)	ASD***>ADHD>TD*
Behavior < 3 yr	3.00	(0.98)	1.47	(1.54)	0.13	(0.41)	ASD***>ADHD>TD***
ADOS module 3[2]							
Communication	2.67	(1.43)	1.09	(0.92)	0.53	(0.76)	ASD***>ADHD,TD
Social interaction	7.26	(3.12)	2.57	(1.96)	1.50	(1.72)	ASD***>ADHD,TD
Com+Soc	9.93	(4.17)	3.66	(2.57)	2.03	(1.99)	ASD***>ADHD,TD
RRB	1.13	(1.24)	0.29	(0.57)	0.16	(0.44)	ASD***>ADHD,TD
Social Affect	8.89	(4.19)	2.83	(2.36)	1.74	(2.02)	ASD***>ADHD,TD
SA+RRB	10.02	(4.68)	3.11	(2.37)	1.89	(2.15)	ASD***>ADHD,TD
PICS[3]							
Inattention	2.26	(2.07)	3.61	(2.18)	0.11	(0.39)	ADHD**>ASD>TD***
hyperactivity/impulsivity	1.98	(1.97)	5.22	(2.45)	0.29	(0.57)	ADHD***>ASD>TD***

	ASD (*n* = 47)		**ADHD** (*n* = 36)		**TD** (*n* = 38)		Group differences (Bonferroni corrected post hoc analyses)
	M	(SD)	M	(SD)	M	(SD)	
WISC-III							
Block Design	9.87	(3.57)	8.33	(3.02)	11.16	(3.23)	ADHD<TD**
Vocabulary	8.81	(3.18)	9.44	(2.10)	11.82	(2.51)	ASD***, ADHD**<TD
Estimated Full scale IQ	96.19	(17.47)	93.26	(12.80)	109.02	(13.64)	ASD, ADHD<TD***
PPVT							
WBQ	104.85	(14.33)	99.97	(12.57)	108.84	(10.72)	ADHD<TD*
Theory of Mind Task							
Proportion correct FB1	0.89	(0.19)	0.88	(0.14)	0.94	(0.11)	n.s.
Proportion correct FB2	0.56	(0.40)	0.55	(0.34)	0.78	(0.29)	ASD, ADHD<TD*
N-Back Task Number correct 2back	39.02	(7.95)	38.19	(7.45)	41.77	(5.28)	n.s.
Stop Task SSRT	257.39	(96.51)	254.84	(94.25)	256.74	(77.59)	n.s.

* = $p < .05$; ** = $p < .01$; *** = $p < .001$; n.s = non-significant

[1] Five children in the ADHD group scored on the ADI-R above the cut-off for ASD (on the basis of Risi et al.'s (2006) criteria)

[2] Two children in the ADHD group scored above the ADOS criteria for ASD

[3] Seven children in the ASD group scored within our criteria for ADHD on the PICS (above or one point below the cut-off on the PICS)

engaged in an other-oriented action (Figure 1) or a self-oriented action (Figure 2). At the same time, the child heard an introductory sentence, followed by a test sentence with either a pronoun or a reflexive (see example (3) and (4)).

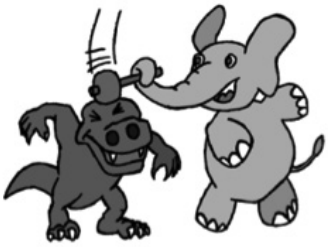


Figure 1 Other-oriented action

Figure 2 Self-oriented action

(3) Introductory sentence:
Een krokodil en een olifant zijn op de stoep.
'An alligator and an elephant are on the sidewalk.'

(4) Test sentence:
De olifant slaat hem/zichzelf.
'The elephant is hitting him/himself.'

The materials were based on the materials of Spenader et al. (2009) and Van Rij et al. (2010). The transitive verbs that were used in the test sentences were the Dutch translations of *to tickle*, *to hit*, *to bite*, *to point to*, *to draw*, *to paint*, *to tie*, *to make up*, and *to dress*. The child was asked whether the prerecorded sentence matched the picture or not. Children had to respond by pressing the "yes" key when the sentence matched the picture; if not, they had to press the "no" key. On trials for which the child decided that the sentence did not match the picture, he or she was always asked to explain why. A second tester noted these justifications.

The task started with two practice items to determine whether the children understood the task. The comprehension task consisted of 34 items: 2 practice items, 16 test items (8 reflexive items and 8 pronoun items), and 16 control items without a pronoun or reflexive, which were included to measure children's general understanding of the task. It is expected that children do not make errors on reflexive items if they are not impaired in their grammatical abilities. In half of the test items the sentence matched the picture (match condition). In the other half of the items the sentence and the picture did not match (mismatch condition); mismatch items contained either a picture of an other-oriented action in combination with a sentence

with a reflexive or a picture of a self-oriented action in combination with a sentence with a pronoun.

Production

To check whether children's production of pronouns and reflexive was adult-like, we used a sentence elicitation task. This task was based on the materials of Spenader et al. (2009). Pictures that were used in the production task were similar to those in the comprehension task. When a picture with an other-oriented action was used in the comprehension task, the corresponding picture with the self-directed action was used in the production task and vice versa. In this way, no picture was shown in both comprehension and production, to avoid possible priming effects. The production task consisted of 16 items in total: two practice items and 14 test items. No filler items were used. Half of the items displayed an other-oriented action, the other half a self-oriented action.

Children saw one picture at a time. They were instructed to first introduce both animals and then to describe the action, leading to sentences like "I see an elephant and a crocodile. The elephant is hitting himself". The production task started with two practice items to determine whether the children understood the task, before they were presented with the test items.

Theory of Mind

To test Theory of Mind, we used a False Belief task adopted from Hollebrandse, Van Hout, and Hendriks (2014) (for an example, see Appendix 1). False Belief tasks involve the understanding that another person has his or her own beliefs and that these can be different from one's own beliefs (e.g., Baron-Cohen, Leslie, & Frith, 1985). The task we used measured both first-order False Belief (FB) (in which the belief of one other person is involved) and second-order FB (which is the understanding of the beliefs of another person about a third person). The task consisted of eight stories, each of which was accompanied by four pictures that were presented one by one on a computer screen. The task was divided in two blocks with a short break in between. The order of stories was counterbalanced across participants. Each story contained one second-order FB question and two first-order FB questions. The first first-order FB question was asked in the middle of the story. The second first-order FB question was asked at the end of the story, preceded by the second-order FB question. The second first-order FB question was presented in order to check whether children had difficulties with the length and complexity of the story.

One item was removed from further analysis since item analysis showed that the response on this item differed from the other seven items: on the second first-order FB question, mean accuracy on this item was only .48, while mean accuracy on other items varied between .79 and .92. Additionally, on this item, mean accuracy on the

second-order FB question was higher (.80) than on the easier first-order FB question (.48). Inspection of this item revealed that its content differed from the other items in that an extra belief had inadvertently been introduced, which made the first-order FB answer less plausible.

Two dependent measures were calculated: mean accuracy on the first first-order FB question (FB1) and mean accuracy on the second-order FB question (FB2).

Response inhibition

Response inhibition was tested with a Stop Task adopted from Van den Wildenberg and Christoffels (2010). In this task, simple drawings of a tree and a door were presented on the computer screen. During go-trials, participants were asked to press the button corresponding with the picture on a two-button box. In 30% of the trials, a visual stop-signal was presented: a red square frame surrounding the picture border. When confronted with the stop-signal, participants had to inhibit the go-response by not pressing the button. The interval between the onset of the go-picture and the onset of the stop-signal (stop-signal delay) was set at 200 ms on the first stop-trial. An online tracking algorithm adjusted stop-signal delay as a function of individual stopping performance (Levitt, 1971). If the participant was able to stop, the stop-signal delay increased by 50 ms, thereby decreasing the chances of successful inhibition on the next stop-trial. After a failed-inhibition trial, the stop-signal delay decreased by 50 ms. This adaptive algorithm ensured successful inhibition on about 50% of the stop-trials, a procedure that yields reliable estimates of the Stop Signal Reaction Time (SSRT) (Band, van der Molen, & Logan, 2003). SSRT was calculated as a measure for response inhibition.

Working Memory

Working memory was tested with an n-back task (Owen, Mcmillan, & Laird, 2005) including three experimental conditions: 0-back (baseline), 1-back, and 2-back. In each condition, pictures were presented randomly on a computer screen with a stimulus duration of 1000 milliseconds, followed by an interstimulus interval of 1500 milliseconds. In the 0-back condition, participants were instructed to press the yes-button when they saw a picture of a car, and to press the no-button when another picture appeared. In the 1-back condition, participants had to press the yes-button when the picture matched the picture immediately preceding it, and otherwise press the no-button. In the 2-back condition, participants had to press the yes-button when the picture matched the picture that appeared two pictures back. Participants started with a practice session of 15 trials per condition (0-, 1- and 2-back), followed by a test session consisting of four sequences of 15 trials per condition (resulting in a total of 60 trials per condition). The total numbers correct on the 2-back condition was calculated as a measure of working memory (WM).

Procedure

Children and their parents were recruited by brochures at schools and in outpatient clinics for child and adolescent psychiatry in Groningen. They took part in a larger study on communication in ASD and ADHD. Children and parents came to the lab together. All parents of participating children signed for informed consent. Children were tested individually on a single day in a quiet testing room with two experimenters present. After every task children had a short break.

Two participants were excluded from further analysis, one (ASD) because he finished neither the pronoun comprehension task nor the pronoun production task, and another (ADHD) because he scored below .75 on the control items in the comprehension task. Furthermore, one child (ASD) conducted only half of the Theory of Mind task and was removed from further analyses that included this task. One child (ASD) did not finish the n-back task and was removed from analyses including the n-back task. Another child (ADHD) did not complete the Stop task and consequently was excluded from analyses involving this task. Finally, one child (ADHD) did not finish the n-back nor the Stop task and was excluded from analyses involving these tasks.

Coding of production data

Children's answers on the production task were voice-recorded. In the production task, more answers are acceptable than just pronouns or reflexives. For pictures showing an other-oriented action, the use of full noun phrases ("the elephant is hitting the crocodile") is also a correct way of describing such actions. Both the use of pronouns ("the elephant is hitting him") and the use of full noun phrases were coded as correct responses in this condition. For pictures showing a self-oriented action, only the use of a reflexive ("the elephant is hitting himself") was treated as accurate. Only active sentences that contained a subject and an object referring to one of the two animals were included in analyses (93.1% of all items). All items were scored independently by two coders, who were blind to the participant's diagnosis. The coders scored the grammatical form of the object (pronoun, reflexive, full noun phrase or other). Inter-scorer agreement was high (Cohen's kappa = .95).

Data analysis

The data were analyzed using Generalized Linear Mixed Models. A logit link was used to accommodate the repeatedly measured binary outcome variable (i.e., accuracy of pronoun interpretation, denoted below as Accuracy) (Heck, Thomas, & Tabata, 2012; Jaeger, 2008). Compound symmetry was used as covariance matrix. First we tested for differences between groups in pronoun comprehension. Contrasts between diagnostic groups and controls (ASD vs. TD and ADHD vs. TD) were dummy-coded and included as fixed factors in the analysis. Whether the sentence matched the picture (coded as 0) or not (coded as 1) was additionally included as a fixed factor. This last

factor was included because previous studies showed clear differences between match and mismatch conditions, probably caused by a yes-bias (see also Chien & Wexler, 1990; Van Rij et al., 2010). In addition to these three main effects (denoted as ASD, ADHD, and Match) we included two two-way interactions (ASD*Match, ADHD*Match) in the model. A two-way interaction or main effect that had no effect on Accuracy (p>.05) was removed from the model.

Next we examined possible mechanisms underlying pronoun comprehension by including the relevant parameters derived from the Theory of Mind task (FB1 and FB2), the n-back task (WM), and the Stop task (SSRT), respectively. All four were mean-centered around a value of zero and were included, in four separate analyses, as fixed factors in the aforementioned model. Interactions that had no effect on Accuracy (p > .05) were removed from the model.

Finally, we tested whether found associations held up when all main and interaction effects with a significance value of p ≤ .05 were examined simultaneously in a multiple GLMM analysis.

Results

Comprehension Task

Table 2 Mean proportions correct responses and standard deviations per group and per condition in the comprehension task

	Pronoun		**Pronoun**		**Reflexive**		**Reflexive**	
	Match		**Mismatch**		**Match**		**Mismatch**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TD	.98	.07	.87	.29	.99	.04	1	.00
ASD	.98	.07	.84	.28	.99	.05	.99	.05
ADHD	.94	.12	.75	.34	.97	.08	.99	.04

In line with our expectations, neither the reflexive match nor the reflexive mismatch condition yielded a substantial number of errors (see Table 2). Therefore we did not statistically test for differences in reflexive interpretation between the groups. Below, our focus is on the two pronoun conditions.

Clinical groups

As expected (Chien & Wexler, 1990; Van Rij et al., 2010), significant effect of Match was found (see Table 3), indicating that more errors were made in the Pronoun Mismatch condition than in the Pronoun Match condition. Interactions of ASD or ADHD with Match did not contribute significantly to participants' scores on the comprehension task (all p-values > .05) showing that this effect held for all groups. In addition, the main effects of ASD and ADHD did not significantly contribute to Accuracy. With no differences among the groups, we conclude that errors in pronoun comprehension are not explained by the presence of ASD or ADHD. In subsequent analyses, main and interaction effects related to diagnostic group were removed, leaving a model that included two main effects (Mechanism and Match) and one interaction effect (Mechanism*Match).

With no differences among the groups, we conclude that errors in pronoun comprehension are not explained by the presence of ASD or ADHD. In subsequent analyses, main and interaction effects related to diagnostic group were removed, leaving a model that included two main effects (Mechanism and Match) and one interaction effect (Mechanism*Match).

Table 3 Estimated effects for Clinical group and Match on accuracy in pronoun interpretation

Predictor	Estimate	SE
Match	-1.94***	.28
ASD vs TD	-.26	.50
ADHD vs TD	-.79	.50

* p<.05; **p< .01; ***p<.001

Mechanisms

No interaction effect of Match with any of the mechanisms was found (all p-values > .05). Therefore, in the final model only the main effects of each of the mechanisms and Match were included, first separately, and next in the multiple GLMM. We found a main effect of FB2 (See Table 4). Lower scores on second-order False Belief questions were associated with lower Accuracy scores in both the Pronoun Match and the Pronoun Mismatch condition. We also found a significant main effect of SSRT. Higher SSRT scores (indicating lower inhibition) were associated with more errors in the pronoun conditions. No significant effects of FB1 or working memory were found. In all four analyses, the main effect of Match remained significant: more errors were made in the Pronoun Mismatch condition than in the Pronoun Match condition.

Table 4 Estimated effects of Mechanism and Match on accuracy in pronoun interpretation

	FB1		FB2		SSRT		WM	
Predictor	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Match	-2.00***	.30	-2.02***	.30	-2.05***	.30	-1.99***	.29
Mechanism	1.71	1.00	1.23*	.51	-.006**	.002	.031	.02

* p<.05; **p< .01; ***p<.001

FB2, SSRT and Match were included in a multiple GLMM (Table 5). All aforementioned associations remained significant. Thus, when adjusted for the effect of SSRT, lower scores on FB2 questions were still associated with lower Accuracy scores in the pronoun conditions. Vice versa, when adjusted for the effect of FB2, higher SSRT scores were still associated with lower Accuracy scores in the pronoun conditions. Furthermore, a main effect for Match remained: adjusted for the effects of FB2 and SSRT, children still performed worse in the pronoun mismatch condition than in the pronoun match condition.

Table 5 Estimated multiple mechanisms model of accuracy in pronouns interpretation

	Estimate	SE
Intercept	3.72***	.31
Match	-2.07***	.31
FB2	1.13*	.50
SSRT	-.005*	.002

Note: only main effects and interactions p<.05 in univariate analyses were included in multiple analyses; * p<.05; **p< .01; ***p<.001

Production Task

In production, consistent with our expectations, children hardly made any mistakes (see Table 6). With all three groups performing at ceiling, we did not test for group differences in production accuracy.

Recall that, for the other-oriented action, both the use of a full noun phrase and the use of a pronoun were scored as correct responses. To check whether the absence

of group differences was due to this scoring procedure, we tested, post-hoc, if children with ASD or ADHD differed from TD children in their use of full noun phrases and pronouns. A GLMM was performed on all items in the other-oriented condition, with full noun phrase (yes or no) as binary dependent variable and two dummy-coded contrasts between diagnostic groups and controls (ASD vs. TD and ADHD vs. TD) as fixed factors. No significant differences between the groups were found (all p-values > .05): children with ASD used a full noun phrase in 96% of the cases, children with ADHD in 95% of the cases and TD children in 94% of the cases. This indicates that children with ASD and children with ADHD use the same linguistic forms as TD children to express other-oriented and self-oriented actions.

Table 6 Mean proportions correct responses per group and per condition in the production task

	Other-oriented action		**Self-oriented action**	
	Mean	SD	Mean	SD
TD	1.0	.00	1.0	.06
ASD	.99	.11	.99	.08
ADHD	1.0	.00	.99	.12

Discussion

The first aim of this study was to provide insights in the possible mechanisms underlying the Pronoun Interpretation Problem. We found that both second-order False Belief performance and SSRT were associated with task performance on the pronoun interpretation task. These results suggest that perspective taking and inhibition are necessary in pronoun interpretation. This is compatible with the constraint-based account of the Pronoun Interpretation Problem by Hendriks and Spenader (2006). We did not find a relation between working memory and performance on pronoun interpretation and thus found no support for Reinhart's claim that sufficient working memory is necessary for reference-set computation. Furthermore, we did not find support for Chien and Wexler's (1990) pragmatic explanation of the Pronoun Interpretation Problem: we found correct production of pronouns, corroborating previous findings with typically developing children that children produce pronouns in an adult-like way (De Villiers et al., 2006; Matthews et al., 2009; Spenader et al., 2009). In addition, our results show that also children with

ASD or ADHD produce pronouns in an adult-like way, at least in the age group between 6 and 12 years old that we investigated. We found an effect of second-order False Belief understanding, and no effect of first-order False Belief understanding. The absence of an association between first-order False Belief understanding and pronoun interpretation is probably due to the ceiling effect in first order False Belief performance (see Table 1). That is, first-order False Belief understanding is generally mastered before the age at which children start to interpret pronouns correctly. This is the main reason why we included second-order False Belief understanding. Second-order False Belief understanding was found to relate to pronoun interpretation, which indicates that taking into account another person's perspective is important in comprehending ambiguous pronouns.

A second aim of this study was to elucidate the processes of pronoun and reflexive interpretation and production in children with ASD and ADHD. We did not find any differences in pronoun interpretation between children with ASD, children with ADHD, and TD children: all three groups showed worse performance on the pronoun mismatch condition than on the pronoun match condition. This difference between match and mismatch conditions corresponds with previous findings (Chien & Wexler, 1990; Van Rij et al., 2010). Children exhibiting the Pronoun Interpretation Problem generally accept both interpretations. Therefore they will correctly accept the disjoint interpretation in the match condition, but also incorrectly accept the coreferential interpretation in the mismatch condition. As expected, we also found that all three groups performed almost perfectly on the reflexive conditions and on the production task. That children with ASD and TD children show a similar Pronoun Interpretation Problem corroborates recent findings by Perovic et al. (2013). Perovic et al. (2013) consider the Pronoun Interpretation Problem to be pragmatic in nature (cf. Chien & Wexler, 1990). At first glance, this leaves unexplained why there are no differences between children with ASD and TD children in their interpretation of pronouns. After all, why would children with ASD, who are known for their pragmatic deficits, not make more errors in pronoun interpretation than TD children? And by extension, why would children with ADHD, who also have pragmatic difficulties, not show differential performance in pronoun interpretation? Perovic et al. (2013) argue that there may be different kinds of pragmatics: a kind of pragmatics related to social rules and a kind of pragmatics more directly related to language (cf. Schaeffer, 2003). This latter so-called 'linguistic pragmatics' may not be affected in ASD, according to Perovic et al (2013).

As an alternative interpretation, we propose that the absence of differences in pronoun interpretation between the ASD group, the ADHD group, and the TD group reflects an absence of grammatical, rather than pragmatic problems in the ASD and ADHD groups. When interpreting a pronoun, a listener must take into account the perspective of a generic speaker in order to determine the optimal interpretation of

the pronoun. That is, the listener must apply the relevant grammatical constraints for comprehension, but at the same time take into account the constraints for production that are relevant for a speaker. This grammatical perspective taking may be different from taking the perspective of the actual speaker (who may be sitting in front of the listener). The latter form of perspective taking is much more challenging, since it differs per speaker and per situation. In contrast, grammatical perspective taking may be less demanding. It only involves taking into account the perspective of a generic speaker. This generic speaker should be understood as contrasting with the actual speaker and may be the listener imagining himself or herself in the role of the speaker. This form of perspective taking does not vary per situation and therefore could be automatized (cf. the computational simulations of Van Rij et al., 2010). Our results may thus provide support for the claim that grammatical perspective taking is unaffected in children with ASD and ADHD. This corroborates previous findings of comparable grammatical performance in ASD children, ADHD children, and TD children (Geurts & Embrechts, 2008; Geurts et al., 2004; Helland et al., 2012; Kim & Kaiser, 2000). In contrast, taking the perspective of a specific speaker may be involved in pragmatic processes such as turn-taking and coherence, which are found to be impaired in ASD and ADHD (e.g., Geurts et al., 2004; Green et al., 2013).

This study was the first to investigate mechanisms hypothesized to be involved in pronoun interpretation. Our results may also apply to other linguistic phenomena that are likely to require listeners to take into account the grammatical perspective of the speaker, such as word order (de Hoop & Kramer, 2006), scalar reasoning (Noveck, 2001), and contrastive stress (Hendriks, de Hoop, Krämer, de Swart, & Zwarts, 2010). Furthermore, the results of the present study are in line with a study on the production referring expressions in subject position (e.g., *he* versus *the pirate*) by the same group of children as in the present study (Kuijper, Hartman, & Hendriks, 2015). That production study showed that children with ASD, children with ADHD and TD children as speakers take into account the listener's grammatical perspective in their choice of referring expressions.

Although we did not find an association between working memory and performance in pronoun interpretation, it should be kept in mind that working memory is a broad concept and many different tasks for its measurement have been developed. In our study, an n-back task with non-verbal stimuli (pictures) was used. It is possible that working memory tasks with verbal stimuli are associated with pronoun comprehension. However, meta-analyses showed that both working memory tasks with verbal stimuli and with non-verbal stimuli relate to general language comprehension (Daneman & Merikle, 1996) and that both give rise to similar activation patterns in neuroimaging studies (Owen et al., 2005). This suggests that a working memory task with verbal stimuli probably would not be related to performance on pronoun interpretation either. Likewise, Theory of Mind and inhibition also include a range of

different aspects. We chose to use a False Belief understanding task in our study to capture Theory of Mind, because in our view this is the Theory of Mind ability that is most closely related to grammatical perspective taking. Previous studies found significant relations between several aspects of language and False Belief understanding (for an overview, see Milligan, Astington, & Dack, 2007). Yet it would be worthwhile to examine the relation between pronoun interpretation and other Theory of Mind skills, such as required in Happé's (1994) strange stories. To capture response inhibition, we used the Stop Task, which has been found to be a reliable and valid measure of response inhibition (Kindlon, Mezzacappa, & Earls, 1995; Tannock, Schachar, Carr, Chajczyk, & Logan, 1989). However, it may be that other inhibition skills (e.g., cognitive inhibition) are associated with pronoun interpretation.

It should be noted that the children in our study were older (6-12 years old) and, correspondingly, had higher accuracy scores in the pronoun mismatch condition (about 80% correct) than the children in many other studies of the Pronoun Interpretation Problem in TD children (generally aged up to 7 years and performing up to 75% correct; Chien & Wexler, 1990; Spenader et al., 2009; Van Rij et al., 2010). However, we still found differences between performance in pronoun interpretation and performance in reflexive interpretation, indicating that our children did not yet reach adult-like performance in pronoun interpretation. Nevertheless, investigation of the role of Theory of Mind and inhibition in pronoun interpretation in a younger group of children may further strengthen our conclusions.

Most of the ASD children in our study could be classified as 'language normal' (based on the PPVT and the vocabulary subtest of the WISC-III cf. Kjelgaard & Tager-Flusberg, 2001). Perovic et al. (2013) found that the linguistic performance of ASD children with general language impairment differed from the linguistic performance of ASD children without general language impairment. However, they only found differences in the interpretation of reflexives, while both groups of ASD children performed similarly on the interpretation of pronouns. A remarkable difference between the study of Perovic et al. (2013). and the present study is the type of task that is used. Perovic et al. (2013) used a picture selection task, which tests for preference of interpretation, rather than for acceptability of interpretation, as in the picture verification task we used. On the basis of their experiment, it can be concluded that ASD children with language impairment have a preference for a disjoint interpretation of pronouns and reflexives. However, it is not clear whether these children would also accept a coreferential interpretation. To further unravel differences between ASD children with and those without language impairment, in grammatical ability as well as the nature of the Pronoun Interpretation Problem, it would be recommended to test these children on their interpretation of pronouns and reflexives with a picture verification task.

In summary, the current study provides insights in the mechanisms underlying the Pronoun Interpretation Problem. We found that both Theory of Mind and inhibition skills were associated with performance on pronoun interpretation. This yields support for Hendriks and Spenader's (2006) grammatical explanation of pronoun interpretation that a listener must take into account the perspective of the speaker and thus block the incorrect interpretation for the pronoun. Our finding may stimulate further research into the extent to which perspective taking and inhibition play a role in other aspects of language such as word order, scalar reasoning, and contrastive stress. Furthermore, our study showed that the grammatical performance of children with ASD or ADHD was comparable to that of typically developing children: the three groups demonstrated a similar delay in their interpretation of pronouns. Neither of the groups showed problems in the production of pronouns and reflexives. This suggests that children with ASD and children with ADHD do not have more problems than TD children in taking into account the grammatical perspective of a generic speaker, despite their possible difficulties in perspective taking with specific conversational partners.