Taking an alternative perspective on language in autism
Overweg, Jennigje

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 4. Spatial language reveals perspective-taking difficulties in children with and without autism
4.1 Abstract
The interpretation of *in front of* and *behind* depends on the relative spatial positions of the speaker and hearer. Interpreting these spatial prepositions from another person’s perspective requires visual perspective taking, which is cognitively demanding and involves several higher-order cognitive processes. Given suggested difficulties with visual perspective taking and cognitive development in children with autism spectrum disorder (ASD), we hypothesized that these children less often than their typically developing (TD) peers interpret spatial prepositions from another person’s contrasting visual perspective. We investigated the interpretation of spatial prepositions from the child’s own visual perspective as a hearer and from the speaker’s contrasting visual perspective in 48 Dutch-speaking children with ASD and 43 TD children (age 6-12). Additional tasks were administered to measure Theory of Mind (ToM), cognitive inhibition, cognitive flexibility and working memory. We found that children with ASD and TD children alike mainly interpreted spatial prepositions from their own visual perspective instead of from the speaker’s contrasting visual perspective. No differences between ASD and TD primary school-aged children emerged, suggesting that an egocentric bias towards one’s own visual perspective is not limited to young children or children with ASD and can be observed in children in general until at least adolescence. Our findings clarify why visual perspective taking is cognitively demanding. We showed an influence of cognitive inhibition, cognitive flexibility, and, to a lesser extent, ToM on the interpretation of spatial prepositions. We propose that hearers need ToM to understand that perspectival differences are relevant for the interpretation of spatial prepositions, and need cognitive inhibition and cognitive flexibility to shift between visual perspectives.

4.2 Introduction
The process of taking another person’s perspective takes time to develop (Birch & Bloom, 2004). In the seminal visual perspective taking study of Piaget and Inhelder (1956), children judged how a doll perceived a landscape by choosing a picture that showed the doll’s viewpoint among a set of pictures with several possible viewpoints. They found that typically developing (TD) children up to 7 chose the alternative viewpoint that showed the landscape
from their ‘egocentric’ visual perspective. Following this study, researchers focused on visual perspective taking in children younger than 7 years old (Flavell et al., 1981; Masangkay et al., 1974; Moll et al., 2013; Newcombe, 1989). For example, Moll et al. (2013) showed that in simple visual perspective taking tasks with only two answer options, 4-year-old children can take another person’s visual perspective. In this study, 4-year-olds understood that while they themselves saw an object in blue, the adult who looked through a yellow colour filter saw the same object in green. Yet, other studies observe that children as old as 10, and even adults, show an egocentric bias (in errors and reaction times) when judging another person’s perspective in simple visual perspective taking tasks (Surtees & Apperly, 2012; Surtees, Samson, & Apperly, 2016). Thus, these findings suggest that an egocentric bias towards one’s own visual perspective can be observed throughout childhood and into adulthood.

Taking another person’s visual perspective is cognitively demanding (Qureshi, Apperly, & Samson, 2010; Surtees, Apperly, & Samson, 2013). Despite the large body of research on visual perspective taking, little is known about which cognitive abilities are involved. In this study, we investigate multiple higher-order cognitive processes that could be needed to take another person’s visual perspective. Firstly, we suggest that Theory of Mind (ToM), the cognitive ability to mentally take the perspective of other people to understand their beliefs, desires and intentions (Wimmer & Perner, 1983), is involved. Some studies indicate a relation between ToM and visual perspective taking (Aichhorn, Perner, Kronbichler, Staffen, & Ladurner, 2006; Farrant, Fletcher, & Maybery, 2006). We hypothesize that hearers need ToM to understand that another person’s visual perspective on a situation can differ from one’s own.

Besides ToM, we suggest that Executive Functioning (EF) is needed to take another person’s visual perspective. EF refers to cognitive processes like working memory (WM), inhibition and flexibility, that allow for the flexible alteration of thought and behaviour in response to changing contexts (Welsh & Pennington, 1988). We expect that WM, the capacity for processing and temporarily holding information in mind (Baddeley, 1986), is involved. Possibly, hearers with low WM capacities are unable to simultaneously keep their own perspective and another person’s perspective in mind and therefore judge the situation from their own visual perspective. Furthermore, in line with previous studies (Qureshi et al.,
2010; Surtees & Apperly, 2012), we propose that cognitive inhibition, the ability to suppress irrelevant information (Dagenbach & Carr, 1994), is needed. That is, taking another person’s perspective involves inhibiting one’s own response prompted by one’s own perspective (Friedman & Leslie, 2005). Yet, inhibiting may not be sufficient. We suggest that cognitive flexibility, the ability to shift between different thoughts or actions (Scott, 1962), is needed to shift to another person’s visual perspective.

The process of taking another person’s visual perspective could be especially demanding for children with autism spectrum disorder (ASD). In a meta-study, Pearson and colleagues (2013) compared 11 studies of visual perspective taking in 8- to 13-year-old children with ASD. Their results indicated that children with ASD understand that other persons can see an object differently from how they themselves see that object. However, they have problems understanding how another person sees that object. In a later study, Pearson and colleagues (2016) find that, compared to a younger control group of TD children, primary school-aged with ASD do not have problems with visual perspective taking. These findings suggest that visual perspective-taking abilities in primary school-aged children with ASD are not impaired, but possibly delayed compared to their TD peers. This, in turn, may be due to difficulties with EF and ToM understanding which have often been reported at different ages in children with ASD (Baron-Cohen et al., 1985; Christ et al., 2011; Happé, 1994; Hill, 2004b; Luna, Doll, Hegedus, Minshew, & Sweeney, 2007).

Together, these findings suggest that it is fruitful to extend our research on children with ASD to everyday life instances of visual perspective taking. In their daily communication children encounter many situations that require them to take another person’s visual perspective. Suppose Peter and Jane are looking for a ball and Peter tells Jane ‘The ball is behind the tree’. The interpretation of behind depends on the relative spatial positions of Peter and Jane. The preposition behind classifies the nature of the spatial relation between the ball and the tree as behind from one visual perspective, while simultaneously the ball is located in front of the tree from the opposite visual perspective. Thus, from Jane’s visual perspective the ball may be in front of the tree. This means children must be able to take another person’s visual perspective in order to fully acquire these prepositions. TD children start to comprehend and produce the spatial prepositions in front of and behind in relation
to their own body by the age of 4 (Washington & Naremore, 1978). Around the age of 5, they learn to use these prepositions in relation to other objects (Bowerman, 1996). The interpretation of in front of and behind has not been studied in children with ASD so far. Nor has it been studied if potential visual perspective-taking impairments are related to EF and ToM abilities.

The present study investigates: 1) whether Dutch-speaking children with ASD interpret voor ‘in front of’ and achter ‘behind’ from another person’s visual perspective less often than their TD peers; and 2) which cognitive processes are needed to shift perspective and interpret these prepositions from another person’s visual perspective. To this end, primary school-aged children with and without ASD are asked to interpret spatial prepositions in situations where the child’s own perspective is the only salient visual perspective (part 1) or in situations where the speaker’s contrasting visual perspective is also present (part 2). Also, children’s interpretations are investigated when these two situations alternate, which best resembles everyday life (part 3). What makes each situation complex is that the child is not told whether the speaker describes the situation from the speaker’s own visual perspective or from the child’s contrasting visual perspective as a hearer. To interpret in front of and behind from the speaker’s visual perspective, the child could engage in perspective taking to interpret the preposition from the speaker’s perspective. In a pilot study with 24 TD adults (mean age 22;5), we found that in the situations in part 1 of the task, as expected, adults mainly interpret in front of and behind from their own visual perspective (Self perspective 97%; Other perspective 3%). In contrast, in the situations in part 2 of the task, adults overwhelmingly interpret in front of and behind from the speaker’s contrasting visual perspective (Self perspective 9%; Other perspective 91%). The adults were not tested on part 3 of the task. The order of the two parts of the tasks was balanced across participants, but did not have an effect on the responses. Thus, to achieve a mature understanding of the prepositions in front of and behind, in part 2 of the task the child should shift from its own visual perspective to the speaker’s contrasting visual perspective.

In this study, we hypothesize that in situations where perspective taking is optional, such as with perspective-dependent spatial prepositions, children prefer to interpret spatial prepositions from their own perspective, and children with ASD even more so than their TD
peers. These expectations are based on the general difficulty of taking another person’s visual perspective and the possibly delayed development of visual perspective-taking skills in children with ASD. We expect that ToM, inhibition, flexibility and WM are needed to shift perspective to interpret spatial prepositions from another person’s visual perspective. These cognitive processes may explain why children with ASD interpret spatial prepositions from another person’s visual perspective less often than their TD peers. Also, based on cognitive flexibility deficits that children with ASD show in daily life (Hill, 2004b), we expect that during constant switching between perspectives (part 3) group differences will be larger.

4.3 Method

4.3.1 Participants

Forty-eight children with ASD and 43 TD children were tested. Only monolingual Dutch-speaking children with no diagnosis of any language disorder were included. All children in the ASD group were diagnosed with ASD by clinicians on the basis of the DSM-IV-TR criteria (APA, 2000) and had an IQ of >75 derived from a clinically administered full IQ test. Additionally, in all children, the Autism Diagnostic Interview Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999), two subtests (Vocabulary and Block Design) of the WISC-III-NL to estimate IQ (Kort et al., 2002), and the Peabody Picture Vocabulary Test to measure Verbal Ability (VA) (PPVT-III-NL; Schlichting, 2005) were administered by certified professionals. Two children from the ASD group were excluded because they met neither the ADOS nor the ADI-R criteria for ASD (cf. Risi et al.’s ASD2 criteria, 2006). One child from the TD group met the ADOS criteria for ASD and was therefore excluded, leaving 46 children with ASD (mean age=9.4; SD=2.2) and 42 TD children (mean age=9.2; SD=2.0) for further analysis. Table 4.1 provides the group descriptives.

Children with ASD and their parents were recruited via outpatient clinics for child and adolescent psychiatry in the north of the Netherlands and a national website for parents with children with ASD. TD children were recruited via information in newsletters and brochures at schools in the north of the Netherlands. Children were tested individually on a single day in a quiet room with two experimenters present. All children participated in a
### Table 4.1 Description of Participants With Autism Spectrum Disorder (ASD) and Typically Developing (TD) Participants

<table>
<thead>
<tr>
<th>Background variables</th>
<th>ASD (N=46)</th>
<th>TD (N=42)</th>
<th>Group differences (General Linear Model ANOVA analyses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (boys:girls)</td>
<td>39:7</td>
<td>34:8</td>
<td>Ns</td>
</tr>
<tr>
<td>Chronological Age (Year;Month)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>9;4 (2;2)</td>
<td>9;2 (2;0)</td>
<td>ns</td>
</tr>
<tr>
<td>Range</td>
<td>6;0-12;5</td>
<td>6;2-12;7</td>
<td></td>
</tr>
<tr>
<td>Clinical diagnosis of ASD subtype according to DSM-IV criteria (N):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autistic Disorder</td>
<td>4</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Asperger’s Disorder</td>
<td>2</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>PDD-NOS</td>
<td>42</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Number of participants meeting ASD2 criteria on:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS and ADI</td>
<td>33</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>ADOS only</td>
<td>10</td>
<td>1 (excluded)</td>
<td>-</td>
</tr>
<tr>
<td>ADI only</td>
<td>3</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>No ASD on ADOS and ADI</td>
<td>2 (excluded)</td>
<td>42</td>
<td>-</td>
</tr>
<tr>
<td>Estimated IQ (WISC)(^b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>99.87 (16.92)</td>
<td>113.21 (13.86)</td>
<td>TD &gt; ASD(^{***})</td>
</tr>
<tr>
<td>Range</td>
<td>66.65-145.48</td>
<td>72.71-145.48</td>
<td></td>
</tr>
<tr>
<td>Verbal ability score (PPVT)(^c)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>104.48 (13.9)</td>
<td>113.62 (11.53)</td>
<td>TD &gt; ASD(^{**})</td>
</tr>
<tr>
<td>Range</td>
<td>77-139</td>
<td>87-138</td>
<td></td>
</tr>
</tbody>
</table>

Note. ANOVA = analysis of variance; ns= nonsignificant; PDD-NOS = pervasive developmental disorder— not otherwise specified; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders (4\(^{th}\) edition); ADOS = Autism Diagnostic Observation Schedule; ADI = Autism Diagnostic Interview.

\(^a\) The ASD2 criteria of Risi et al. (2006) are as follows: “a child meets criteria on Social and Communication domains or meets criteria on Social and within 2 points of Communication criteria or meets criteria on Communication and within 2 points of Social criteria or within 1 point on both Social and Communication domains” (Risi et al., 2006; p.1100). \(^b\) Estimated IQ on the basis of two subtests of the Dutch version of the Wechsler Intelligence Scale for Children (WISC-III-NL; Kort et al., 2002). \(^c\) Normed verbal ability score from the Dutch version of the Peabody Picture Vocabulary Test (PPVT-III-NL; Schlichting, 2005); **p<.01; ***p<.001
larger study on language and communication in ASD. The medical ethical committee of the University Medical Hospital Groningen evaluated this study as not falling under the Medical Research Involving Human Subjects Act (WMO). Nevertheless, we followed the required procedures and obtained informed consent from parents.

4.3.2 Outcome measure

Spatial prepositions task. Participants watched short films in which two objects were visible on a table and a woman described the position of these objects. Participants had to press one of two buttons (agree versus disagree) to indicate whether the spoken description matched the film. In the Self Perspective condition (block 1), the woman uttering the description is not visible in the film and children were asked whether the spoken description matched the film according to themselves. In the Other Perspective condition (block 2), the woman is visible in the film and children were asked whether the spoken description matched the film according to the woman. We designed the task such that the woman described the situation from her own perspective. In block 3 perspectives were alternated in a semi-random fashion. Figure 4.1 shows stills from the films in both perspectives.

Figure 4.1 Stills from the films in the Self perspective (left) and the Other perspective (right) condition accompanied by the spoken description ‘The ball is behind the hat’
First, three practice items were administered to check whether the participants understood the task. The test session consisted of 48 items, distributed across 4 conditions (Perspective-independent prepositions in Self or Other perspective, Perspective-dependent prepositions in Self or Other perspective). The Perspective-independent prepositions *in* and *on* served as control items (their interpretation is independent of the speaker’s and hearer’s spatial position). The Perspective-dependent prepositions *in front of* and *behind* served as test items. For each condition, in half of the items the spoken description matched the film; in the other half it did not. Stimuli were presented and data recorded using the computer software E-Prime 2.0 (Schneider et al., 2002).

### 4.3.3 Cognitive tasks

**ToM.** The Bake Sales task, a second-order false belief (FB) Task adopted from Hollebrandse, van Hout and Hendriks (2014) was used to test ToM. Their stories were modelled after Perner and Wimmer’s (1985) “ice cream truck story”. The task consisted of eight stories, each of which contained a first-order FB question (involving the belief of one other person) and a second-order FB question (involving the belief of another person about a third person). The measures of ToM1 and ToM2 were calculated using the ACC on the eight first-order FB questions and the eight second-order FB questions, respectively.

**Cognitive inhibition.** In the Flanker test (Amsterdam Neuropsychological Test battery (ANT) version 2.1; De Sonneville, 1999) participants had to identify the colour of a target stimulus that was surrounded by eight distractors (flankers). The target colour was associated with the left (red) or right (green) button. The flankers were in the same colour as the target (compatible) or in the colour that was associated with the opposite response of the target (incompatible). The participant received 40 compatible and 40 incompatible items. The mean ACC and mean reaction time (RT) of cognitive inhibition was measured by subtracting the mean ACC or RT on compatible trials from the mean ACC or RT, respectively, on incompatible trials (resulting in the congruency effect; see Mullane, Corkum, Klein, & McLaughlin, 2009).

**Cognitive flexibility task.** To test cognitive flexibility, a classical switch task was used (Rogers & Monsell, 1995). This task is an adaptation of the gender-emotion switch task (De
Vries & Geurts, 2012). Pictures of round or square figures, in black or white, were displayed on the computer screen. Participants had to press the left or right button to report the shape (round or square) or the colour (black or white) of the figure. The cue at the top indicated whether the shape or colour had to be reported. The test part consisted of 72 switch trials (switching from colour to shape or vice versa). The mean ACC and mean RT of switch costs was measured by subtracting the mean ACC or RT on repeat trials from the mean ACC or RT, respectively, on switch trials (cf. De Vries & Geurts, 2012).

Working Memory. In the N-Back task (Owen et al., 2005) participants had to remember pictures presented on a screen and indicate per picture if that picture matched the picture of the current trial (0-back) or the picture one (1-back) and two (2-back) trials before. Each participant received a practice session of 15 trials and a test session of 60 trials per condition. The mean accuracy (ACC) on the 2-back condition was calculated as a measure of WM.

4.3.4 Data analysis

The data of the spatial preposition task were analyzed using Generalized Linear Mixed Models (GLMMs), using a logit link to accommodate the repeatedly measured (48 trials) binary outcome variable Accuracy (0 for incorrect; 1 for correct) (Heck et al., 2012; Jaeger, 2008). Compound symmetry was used as the covariance matrix type. We set out with a full factorial model with Block (Standard vs. Switch), Perspective (Self vs. Other) and Preposition (Perspective-independent vs. Perspective-dependent) as within-group factors and Group (TD vs. ASD) as between-group factor. Age was mean-centered and additionally included. Interactions with no effect on Accuracy (p > .05) were removed from the model one by one, choosing the largest p-value for removal, after which we refitted the model. This resulted in model 1, which shows the extent to which Accuracy was predicted by Block, Perspective, Preposition, Group, and Age, as well as the relevant (p < .05) interactions.

Next, the seven parameters derived from the False belief task (ToM1 and ToM2), Flanker task (Cognitive inhibition ACC and Cognitive inhibition RT), cognitive flexibility task (Switch costs ACC and Switch costs RT) and N-Back task (WM) were mean-centered and, one by one, examined as main effects and in interaction with the significant predictors from
model 1 in seven separate analyses. The data of 3 participants (2 ASD and 1 TD) were missing in the Cognitive inhibition ACC and RT analyses, leaving the data of 44 ASD and 41 TD participants. In each separate analysis, interactions that had no effect on Accuracy (p > .05) were removed from the model. Based on the outcomes of these analyses per predictor, we combined the cognitive processes with (main or interaction) effects on Accuracy (p < .05) and added these with the significant predictors of model 1 in a model with multiple predictors to evaluate their effects adjusted for one another (cf. Kuijper, Hartman, & Hendriks, 2015; Overweg, Hartman, & Hendriks, 2018). This resulted in model 2, which shows the relevant cognitive processes that had an effect on the interpretation of spatial prepositions.

Finally, the parameters from the WISC (estimated IQ) and PPVT (VA) were mean-centered and included in two separate analyses in model 1. If they had an effect on Accuracy (p < .05), they were added to model 2 to check whether these general background variables changed the effects found in model 2.

4.4 Results

Figure 4.2 presents the mean proportions of correct responses in all conditions, separately for the ASD and TD groups. As expected, children performed (almost) at ceiling in the Perspective-independent control conditions (with the prepositions in and on), indicating that they understood the task.

Model 1 showed main effects of Perspective and Preposition and interactions of Perspective*Block and Perspective*Preposition. As expected, children mostly interpret Perspective-dependent prepositions in the Other perspective from their own perspective. Hypothesized lower Accuracy in the ASD group compared to the TD group in the Switch block was not found. In contrast, both groups improved significantly with Perspective-dependent prepositions in the Other perspective condition in the Switch block compared to the Standard block. We also found a main effect of Age, indicating that the older the child, the better their overall performance on the task. No main effect of Group nor interactions with Group were found (all p-values >.05). Table 4.2 lists all effects of model 1.

Next, we examined one by one, which cognitive processes were associated with Accuracy. These separate analyses indicated main effects of ToM2 (B=0.991; SE=0.451;
Figure 4.2 Accuracy of the interpretation of spatial prepositions in all four conditions per group (TD vs. ASD) in each Block (Standard vs. Switch)

Model 2 showed a main effect of ToM2, indicating that children with low second-order ToM understanding show a lower Accuracy overall on the task. Model 2 also showed interaction effects of Cognitive inhibition RT*Preposition and Switch costs RT*Perspective*Preposition. Children who are able to inhibit quickly show a higher Accuracy in the interpretation of Perspective-dependent prepositions, as is shown in Figure 4.3. Children who are slower switchers have more problems with the interpretation of Perspective-dependent prepositions in the Other perspective, as is shown in Figure 4.4.
Spatial language reveals perspective-taking difficulties in children with and without autism

**Figure 4.3** Change in estimated mean of Accuracy on Preposition (Independent vs. Dependent) plotted for Cognitive inhibition as low RT (1 SD below mean) and high RT (1 SD above mean) (mean=0; SD=38.95)

**Figure 4.4** Change in estimated mean of Accuracy on all four conditions plotted for Switch Cost at low RT (1 SD below mean) and high RT (1 SD above mean) (mean=0; SD=70.38)
### Table 4.2 Estimated Effects of Models 1 and 2 on the Interpretation of Spatial Prepositions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimate</th>
<th>SE</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.705**</td>
<td>0.298</td>
<td>3.794**</td>
<td>0.307</td>
</tr>
<tr>
<td>Group</td>
<td>-0.440</td>
<td>0.312</td>
<td>-0.354</td>
<td>0.311</td>
</tr>
<tr>
<td>Age</td>
<td>0.016*</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>Perspective</td>
<td>0.789**</td>
<td>0.262</td>
<td>0.750*</td>
<td>0.266</td>
</tr>
<tr>
<td>Block</td>
<td>-0.063</td>
<td>0.167</td>
<td>-0.092</td>
<td>0.166</td>
</tr>
<tr>
<td>Preposition</td>
<td>-1.477**</td>
<td>0.299</td>
<td>-1.651**</td>
<td>0.291</td>
</tr>
<tr>
<td>Perspective*Block</td>
<td>0.410*</td>
<td>0.203</td>
<td>0.444*</td>
<td>0.203</td>
</tr>
<tr>
<td>Perspective*Preposition</td>
<td>-3.023**</td>
<td>0.456</td>
<td>-2.947**</td>
<td>0.458</td>
</tr>
<tr>
<td>ToM2</td>
<td></td>
<td></td>
<td>1.072*</td>
<td>0.473</td>
</tr>
<tr>
<td>Cognitive inhibition RT</td>
<td></td>
<td></td>
<td>0.009*</td>
<td>0.004</td>
</tr>
<tr>
<td>Switch costs RT</td>
<td></td>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>Cognitive inhibition RT*Preposition</td>
<td></td>
<td></td>
<td>-0.011*</td>
<td>0.005</td>
</tr>
<tr>
<td>Switch costs RT*Preposition</td>
<td></td>
<td></td>
<td>0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>Switch costs RT*Perspective</td>
<td></td>
<td></td>
<td>0.005</td>
<td>0.003</td>
</tr>
<tr>
<td>Switch costs RT<em>Perspective</em>Preposition</td>
<td></td>
<td></td>
<td>-0.013*</td>
<td>0.006</td>
</tr>
</tbody>
</table>

* p=<.05; **p=<.01

Finally, the potential effects of the background variables IQ and VA on Accuracy were checked. These analyses per predictor indicated a main effect of VA (B=0.025; SE=0.012; p=.046) and no effect of IQ (all p-values >.05). In model 3, we combined the main effect of VA together with the effects of model 2. However, VA did not add to the explanation of Accuracy (p=.164), indicating that model 2 described our data best.

Note that the main effects of Perspective and Preposition and the interactions Perspective*Block and Perspective*Preposition remained significant in model 2. This indicates that ToM, cognitive inhibition and flexibility could not fully explain the relative difficulty to understand perspective-dependent prepositions. However, with the addition of these cognitive processes, the main effect of Age disappeared (p >.05). Together, the results
show that second-order ToM understanding, cognitive inhibition and cognitive flexibility contribute to the interpretation of spatial prepositions. Individual differences therein explain why older children perform better than younger children.

4.5 Discussion
We investigated the interpretation of spatial prepositions by primary school-aged children with and without ASD. Children with ASD and also their TD peers interpreted perspective-dependent spatial prepositions from their own visual perspective instead of from the speaker’s contrasting visual perspective. We found that older children performed better than younger children. Contrary to our expectations, no differences between children with ASD and their TD peers emerged. For both groups, cognitive inhibition and cognitive flexibility were needed to shift perspective and interpret these prepositions from another person’s visual perspective and second-order ToM was associated with the interpretation of spatial prepositions in general. No role of WM was found.

Children with ASD and TD children alike had difficulties ignoring their own perspective when the speaker’s contrasting visual perspective was present, which would allow them to shift perspective to achieve a mature interpretation. Children were getting better with age, but an egocentric bias persisted. This suggests that an egocentric bias towards one’s own visual perspective can be observed until at least adolescence, or perhaps even adulthood (Surtees & Apperly, 2012; Surtees et al., 2016). Like Pearson et al. (2016), we found that children with ASD resemble their TD peers in their difficulties in taking another person’s visual perspective. Possibly, visual perspective-taking differences between individuals with and without ASD are best observed in adulthood. Schwarzkopf and colleagues (2014) found that, compared to a control group, adults with ASD have more problems ignoring their own perspective when judging another person’s visual perspective.

Contrary to our a priori expectations, children with ASD, and especially TD children, improved their interpretation of prepositions from another person’s perspective when they constantly had to switch between perspectives. The improved performance in the switch block could be (in part) a learning effect, since this block was always presented last. However, it is plausible that through switching the child became aware of the difference
between their own perspective and the speaker’s perspective. The awareness of different visual perspectives could help children in their development of visual perspective taking. This may explain why some studies report that 4-year-old TD children take another person’s visual perspective (Masangkay et al., 1974; Moll et al., 2013). In these studies, children are first explained how each situation can be observed and then are explicitly asked how they themselves as well as the other person perceive each situation. This creates an explicit awareness that in each situation the child’s own perspective and the other person’s perspective can differ, and that perspectival differences are relevant. In our task, it was not explicitly pointed out how each situation could be observed. Also, we designed our task such that children did not know from whose perspective the speaker described the situation, as is usually not the case in daily life. Yet, the switch block may have aided them in perspective-switching.

Previous studies on visual perspective taking have suggested that taking another person’s visual perspective involves cognitive processes like inhibition (Qureshi et al., 2010; Schwarzkopf et al., 2014). Our results provide further insight into the specific cognitive processes involved in shifting to another person’s contrasting visual perspective. We confirmed the role of inhibition. Our findings suggest that children with and without ASD need cognitive inhibition to suppress their own visual perspective when a different perspective is available, even when it is not necessary to actually shift to this other perspective. Also, as expected, cognitive flexibility is involved. Our findings suggest that cognitive flexibility is involved in order to shift from one’s own (egocentric) perspective to the other person’s perspective to interpret a perspective-dependent preposition from that person’s perspective. Although a relation between cognitive flexibility and perspective-taking abilities has been suggested (Geurts et al., 2010), we are the first to show this relation empirically. Yet, these cognitive processes only partly explained what seems to be needed to shift perspective and to interpret spatial prepositions from another person’s visual perspective. Additional cognitive processes, such as mental rotation (Pearson et al., 2016), may play a role.

Whereas our findings suggested that inhibition and flexibility are needed to shift perspective to interpret perspective-dependent spatial prepositions, ToM turned out to be
Spatial language reveals perspective-taking difficulties in children with and without autism

relevant for the interpretation of spatial prepositions in general. An explanation for the more
general effect of ToM, compared to specific effects of inhibition and flexibility, is that ToM
possibly is only needed for the awareness that different visual perspectives are available and
for the calculation of each perspective. According to Aichhorn et al. (2006), visual
perspective taking tasks require ToM, but only to realize that other persons can have
different perspectives on the world because they represent the world differently. In line with
this, we propose that ToM is needed in the interpretation of spatial prepositions to calculate
how another person perceives the situation and whether their perspective differs from one’s
own. This explanation implies that the process of visual perspective taking is divided into
different steps (or sub-processes), a view endorsed by Leslie and colleagues (2005). They
claimed a distinction between the calculation process of what someone sees, knows or
thinks and the selection process of making a judgement on the basis of this information.

After calculating another person’s perspective, one needs to select that perspective when
this perspective differs from one’s own. Our findings suggest that in this selection process,
inhibition is needed to inhibit an egocentric perspective and flexibility is needed to shift to
the other person’s perspective and select this perspective. Qureshi et al. (2010) provide
support for this interpretation, showing that EF processes, like inhibition, are only involved
in the selection between perspectives, and not in the calculation of another person’s
perspective.

To conclude, we have shown that children with and without ASD until at least the age
of 12 predominantly interpret spatial language from their own visual perspective. Children
with ASD and TD children alike showed an egocentric bias towards their own visual
perspective. These results suggest that perspective-taking difficulties are not limited to
young children or children with ASD and can be observed in children in general until at least
adolescence. Our findings shed light on why visual perspective taking is cognitively
demanding. Hearers seem to need ToM to become aware of different perspectives, and
require inhibition and flexibility skills to interpret spatial prepositions from the other
person’s perspective.