Elsevier Editorial System(tm) for Research in Developmental Disabilities
Manuscript Draft

Manuscript Number: RIDD-D-14-00746R2

Title: Exploring different explanations for performance on a theory of mind task in Williams syndrome and Autism using eye movements

Article Type: Research Paper

Keywords: Theory of mind, eye movements, Williams syndrome, Autism Spectrum Disorders

Corresponding Author: Dr Dagmara Dimitriou, PhD

Corresponding Author's Institution: Institute of Education-UCL

First Author: Jo Van Herwegen, PhD

Order of Authors: Jo Van Herwegen, PhD; Tim Smith, PhD; Dagmara Dimitriou, PhD

Abstract: The current study explored the looking behaviours of young children with Autism Spectrum Disorders (ASD), Williams syndrome (WS), and typically developing (TD) children while they were administered a low-verbal Theory of Mind (ToM) task. Although ToM performance in both clinical groups was impaired, only participants with WS showed small differences in looking behaviour at the start of the video. Furthermore, while TD children who passed the ToM task looked longer at the original hiding place there was no such contrast in the clinical groups. This shows that looking behaviour in ASD and WS is not necessarily atypical when saliency aspects such as language, background, and colour are removed and that differences in looking behaviour cannot explain ToM performance.

***Cover Letter**

Cover Letter,

Dear Editor,

We would like to thank you for your detailed comments. We are rather ashamed about some of the mistakes in terms of grammar and spelling errors. Please note that we have carefully made all the corrections pointed out in your email.

With Kind Regards,

***Detailed Response to Reviewers**

Reply to reviewers' comments

We would like to thank the reviewers for their insightful comments and suggestions. We have carefully considered the comments and by inclusion of the revisions below we hope that we have answered the reviewers' comments and that these changes improve the manuscript satisfactorily.

With Kind Regards,

Highlights

- This is the first study to explore looking behaviour of children with ASD and WS during a Theory of Mind task containing dynamic stimuli
- The WS group have difficulties disengaging from faces rather than just a preference for faces per se
- The ASD group did not show avoidance of looking at the faces in favour of objects or looking at the background.
- The WS and ASD groups, groups showed similar looking behaviours when viewing social

Introduction

It is now well established that Theory of Mind (ToM) ability forms a platform for the development of socio-cognitive abilities, such as understanding of the intentions and behaviour of others (Premack and Woodruff 1978). The nature and theoretical interpretation of ToM has become a focus in developmental and neurocognitive research and has sparked a wide range of paradigms. Studies have shown a clear progression on the acquisition of ToM abilities with typically developing (TD) children aged 4 years and younger passing simple ToM tasks (Wimmer and Perner 1983; Liszkowski et al. 2008), while more complex ToM abilities continue to develop into adulthood (Apperly et al., 2011). In contrast, individuals with developmental disorders, such as Autism Spectrum Disorders (ASD) and Williams syndrome (WS), have been found to be impaired on ToM tests, even when their mental age exceeds that of a 4 year old (Baron-Cohen et al. 1997; Tager-Flusberg and Sullivan 2000).

WS is a rare neurodevelopmental disorder (about 1 in 20 000 live births) caused by a deletion of some 28 genes on the long arm of one copy of chromosome 7 at q11.23 (Donnai and Karmiloff-Smith 2000). Despite an overall lower IQ of 50-70, individuals with WS show an uneven cognitive profile, with good performance on receptive vocabulary and face recognition in contrast to non-verbal abilities such as number, planning, visuo-spatial abilities, and route learning (Van Herwegen et al. 2011). Although individuals with WS are inclined to be overly sociable (Mervis et al. 2000), their performance on tasks that assess cognitive aspects of social development is impaired. This has led researchers to conclude that the social profile in WS might be uneven as well with the social perceptual components being intact but the socio-perceptual component being impaired (Tager-Flusberg and Sullivan 2000). However, more recent studies have found that participants with WS also show difficulties on tasks that tap into socio-perceptual abilities (for example Plesa-Skwerer et al 2006).

In contrast, Autism Spectrum Disorder (ASD) is a common neurodevelopmental syndrome (1 in 100) characterised by two core impairments in communication or social behaviour and repetitive behaviours from early childhood onwards (American Psychological

Association 2013). While there are marked individual differences in the extent and quality of the symptoms amongst individuals with ASD, one of the most common features is a striking difficulty with social skills, including the ability to attend to faces and difficulties in emotion recognition (Klin et al. 2002; Riby et al. 2011a). However, not all studies have found that individuals with ASD have an aversion towards faces and atypical looking behaviour towards faces and impaired emotion recognition has been found for static stimuli but not always for dynamic stimuli (Back et al. 2007; Speer et al. 2007).

Although there are many contrasting features between the individuals with ASD and WS, there are also several commonalities in their behavioural and cognitive profiles (Lincoln et al. 2007). For instance, both groups show socio-communication problems such as delayed use of pointing, unusual eye contact and problems with joint attention (Charman et al.1997; Klein-Tasman et al. 2007; Laing et al. 2002). In addition, impairment on ToM tasks has been reported in both clinical groups. For example, Baron-Cohen and colleagues (1985) reported that 80% of individuals with autism failed the change-of-location task Sally-Ann in which participants are asked to follow a scenario where Sally leaves a marble in a basket and Anne moves the marble to a box while Sally is away. Individuals with ASD responded incorrectly to the false-belief question of the Sally-Ann task by answering that Sally should look in the box where the marble had moved to rather than in the box Sally believed the marble to be in. Similarly, participants with WS fail this task as well (Tager-Flusberg and Sulliven 2000; Van Herwegen et al. 2013). Because of the differences and similarities in their social and cognitive profiles, contrasting performance of individuals with ASD to WS allows us to explore the cognitive mechanisms that underlie task performance, including theory of mind tasks.

Several domain specific theories have been proposed as to why individuals with ASD and WS might fail ToM tasks: for example performance on ToM tasks might be caused by the language impairments observed in WS and ASD (Tager-Flusberg 2000; Happé 1995). Yet, others suggest executive functioning and the use of context, are better predictors (Pellicano 2010; Van Herwegen et al. 2013). Although several studies have shown evidence for a deficit to use context in WS and ASD (for a review see Bernardino et al. 2012 as well as

Happé and Frith 2006), studies in ASD have shown that this cannot explain performance on ToM tasks (Burnette et al. 2005; Happé 1997).

One reason why the current theories cannot describe the difficulties observed in ASD and WS is that they only focus on domain-specific areas of cognition to explain ToM deficits. Recent evidence has suggested that impairments in domain-general abilities, such as attention or where in a visual scene a person was looking for detailed information, can explain impairments in domain-specific areas later on in life (Karmiloff-Smith et al. 2012). Thus, it is possible that subtle differences in looking behaviour or where a person was looking for detailed information can provide valuable information about what strategies individuals use to complete a task and whether task approach in WS and ASD is typical or atypical which in turn might provide an explanation for the task performance in ASD and WS on ToM tasks.

Previous eye tracking studies in participants with WS and ASD have shown that they show atypical attention patterns during social tasks. Riby and Hancock (2009) found that in human videos individuals with WS aged 8 to 28 years old spend more time looking at faces and less at the actors' bodies compared to controls while participants with ASD of a similar age looked less at the faces compared to the controls. It has been suggested that these atypical looking behaviours are related to the social abilities of individuals with WS and ASD and that atypical attentional bias toward others' faces could contribute to atypical social orienting (Kikuchi et al. 2009). For example, Klin and colleagues (2002) showed that adolescents with ASD focussed more on the mouth and less on the eye region while watching black and white video clips from the film "who is afraid of Virginia Woolf", in contrast to TD controls. They also found that looking behaviour predicted social competence in ASD: those who spend longer fixating on the mouth were more sociable, while there was a negative correlation for sociability and the time fixating objects. This suggests that atypical looking behaviours might also cause problems for task performance on theory of mind tasks (see also, Senju et al. 2010). However, none of these studies have directly examined whether individuals with WS and ASD show atypical looking behaviour during a ToM task and whether these looking behaviours can actually explain ToM deficits.

The current study is the first to explore the looking behaviours of participants with ASD and WS whilst performing a ToM task. Based upon previous studies, it was hypothesised that those with WS would fail the test questions as they would have difficulty to disengage from social stimuli which would prevent them from focusing on where the object had been hidden as well as where the object had been moved to. In contrast, those with ASD would fail the task as they would favour non-social stimuli in the background and thus they would have insufficient information to infer the deception included in the ToM task, as information about deception is especially visible in facial expressions. Thus, for both developmental disorders it was predicted that atypical looking behaviour would impair their focus on important information in the scene resulting in incorrect cognitive interpretations about the outcome of the ToM situation.

Methods

Participants

Fourteen children (4 male) with WS, 13 children with ASD (12 male) and a control group of 14 (5 male) typically developing (TD) children took part in the study. TD children were recruited through local primary schools and parents of children with ASD were contacted through mailing lists of local support groups, special needs schools and groups. The Williams Syndrome Foundation, UK assisted with recruitment of children with WS. Children with WS had been diagnosed clinically as well as by means of the *fluorescence in situ hybridisation* (FISH) test for microdeletion of genes at the elastic locus (7q11.22-11.23). Children in the ASD group met established criteria for autism, such as those specified in DSM-5 (American Psychiatric Association 2013). They had an average score within the mild-to-moderate range for ASD (mean: 32.62, *SD*= 4.782, range: 27 to 40.5) on the Childhood Autism Rating Scale (CARS: Schopler et al. 1988). Ethical approval was granted by the Middlesex University Research Ethics Committee and supported by the Williams Syndrome Foundation, UK. Both parental informed consent and the child's assent were obtained prior to participation. All participants were white and came from a mainly working class background, with a similar Social Economic Status. There was no significant difference in chronological ages between the groups ($F(2,40)= 0.611$, $p= 0.542$). Table 1 provides an overview of the chronological ages of the three groups.

TABLE 1 ABOUT HERE

Clinical groups were matched to the TD children on chronological age since the dependent variable in the current study focuses on looking behaviours (and pointing), i.e., non-verbal and implicit abilities for which no direct mental age equivalent scores are available. However, as this is likely to disadvantage developmental groups who rarely perform at CA levels, performance on standardised tasks was obtained to evaluate their verbal and non-verbal abilities as well.

Materials

Background measures. Participants were administered the *British Picture Vocabulary Scale* (BPVS: Dunn et al. 1997) to obtain vocabulary comprehension scores and the *Raven's Coloured Progressive Matrices* (RCPM: Raven et al. 1990) to obtain non-verbal performance scores. Due to the fact that scores for some participants in the clinical group were low on these tasks, and no age equivalent scores were available for these scores, raw scores were used instead. Although this approach does not allow comparison of the verbal versus non-verbal abilities within each clinical group, it still allows investigation of any differences between the different groups.

Low-verbal theory of mind task. This task consisted of a false belief task administered in Van Herwegen et al 2013. Participants were asked to watch one change-location ToM task similar to the classic ToM story of Sally-Ann. In the video, two protagonists (either two girls or two boys) are in a room, where there is a table with a basket and a box on. Protagonist A or "the seeker" has an object (either an apple or a game console). The seeker puts the object in the basket. The seeker yawns, stretches, and leaves the room. While the seeker is away, protagonist B or "the mover" goes to the basket and moves the object into the box. Then the mover leaves the room through a different door. When the seeker returns, the participant is asked a prediction of action question: "Where will this girl/boy look for the apple/game?" (Prediction question). Next, the participant was asked a reality question ("Where is the apple/ game now?") and a memory question ("Where did the girl/boy put the apple/game?"). Participants answered these questions by pointing to a picture out of three options (a picture of the seeker looking into the basket, the seeker looking into the box and the seeker looking under the table). These pictures were presented from left to right on the screen in randomised order and a researcher who stood behind the participant recorded the participant's answer.

For the eye-movement analyses the following dynamic area of interests (AOIs) were identified: 1) the face of the mover, 2) the face of the seeker, 3) hiding place one, and 4) hiding place two (see Figure 1) using an in-house software tool, *Gazeatron* (Võ, et al., 2012).

FIGURE 1 HERE

In order to investigate any differences in looking behaviour during particular moments in the task, six 3-second scenes in the video were identified: 1) the seeker hiding the object, 2) the seeker leaves or moves backwards, 3) the mover retrieves the object, 4) the mover moves the object to new location, 5) the mover leaves the room or moves backwards, and 6) the seeker returns.

Apparatus

A Tobii eye-tracker was used to record eye position data at 120 Hz and the stimuli were presented on a 17–inch monitor. Eye-movement recordings were controlled with Tobii's Studio software (version 2.01) while the experiments were controlled by E-prime software version 2.0 professional software.

Procedure

Participants were seated facing the eye tracker monitor at a distance of 60 cm. A 5-point calibration was conducted before the participant watched the videos. Participants were asked to watch the video carefully as they would be asked some questions afterwards. The video took about 40 seconds. The percentage of lost data during the videos due to blinks or poor tracking did not differ significantly across the groups: TD group (mean = 18.19%, SD = 8.62), WS group (mean= 17.62%, SD= 12.25), ASD group (mean = 23.89%, SD = 19.18), all *t*s < 2, and p-values > 0.5. In order to control for differences in length of the AOIs and to take into account the differences in data loss looking behaviour to each of the AOIs was calculated as a proportion of the total looking time on screen.

Results

Background measures

A one-way ANOVA comparing BPVS results between the three groups, showed that there was a significant difference between the groups on the BPVS raw scores; $F(2,23.627) =$

6.278, $p = .007$[1], $\eta^2 = .17$. Games-Howell post-hoc comparisons showed that the WS group performed significantly lower compared to the TD group ($p = .004$). There were no other group differences (all p's > .05). On the RCPM, several children decided not to complete the test (two participants with WS aged 8;0 and 9;10 years and three with ASD aged 4;03, 5;09 and 9;11 years old). Again a one-way ANOVA showed a significant difference in RCPM scores between the three groups; $F(2,35)= 35.659$, $p = .008$, $\eta^2 = .26$. However, as there were unequal group sizes, Gabriel post-hoc analyses were run which revealed a significant difference in scores between the TD and WS group ($p = .006$) but not the other groups ($p >$ .05). These differences show that the WS group performed at a lower level than the control children (see Table 1). However, the fact that no differences were found between the ASD and WS group suggests that these two groups can be directly compared.

Behavioural performance on low-verbal Theory of Mind task

Although more participants with WS and ASD (number of WS = 64%; ASD = 38%) failed the prediction question in the false belief task compared to the TD group (21%), Fisher-Freeman-Halton tests[2] showed that this difference did not reach significance (exact $p =$ .068). However, there was a significant difference between the three groups for the reality question (exact $p = .015$) and the memory question (exact $p = .011$). As shown in Table 2, more children with WS failed the memory and reality questions compared to the two other groups. When performance on the prediction question was examined for those participants who passed the reality and memory questions only, a significant difference between the three groups (Fisher-Freeman-Halton test exact $p = .044$) was found with TD > ASD > WS.

TABLE 2 HERE

Overall differences in looking behaviour

Group differences for proportion of total dwell time (i.e. cumulative fixation durations across multiple visits to the AOI relative to the total gaze time on screen) were investigated for each of the AOIs across the entire video (Table 3). One-way ANOVA tests were carried

[1] Welsh ANOVA was used since the homogeneity of variances assumption was violated.

[2] These are extensions of the Fisher's exact test for a 2 by 3 design (see http://vassarstats.net/fisher2x3.html)

out to investigate any group differences and Welsh ANOVAs were calculated when the assumption of equal variances was violated.

TABLE 3 HERE

Comparisons between the three groups for overall time looking during the video did not show any differences for overall looking times; $F(2,40)= .614$, p= .547, $\eta^2= .033$ which showed that all three groups engaged with the task in a similar way. More detailed investigation of the eyemovements towards the different aspects of the video showed that there was a significant difference for the face of the seeker; $F(2, 22.096)= 4.096$, $p = .031$, $\eta^2 = .075$. Games-Howell post-hoc tests showed that the WS group looked longer at the face of the seeker compared to the TD group ($p = .026$). All other post-hoc comparisons and other ANOVA comparisons were non-significant (all p's > .05).

Next, it was investigated whether there were any differences between the three groups in looking behaviour towards each of the AOIs during certain scenes within the videos. Again there were no significant differences (all p's > .05).

FIGURE 2 HERE

Can looking behaviour explain performance on ToM task?

Next, it was investigated whether there were any differences in looking behaviour between those who passed and those who failed the prediction question in each condition. Comparisons within each group showed that those participants in the TD group who passed the prediction question looked longer ($t(12)= -2.209$, $p = .047$) at the original location (.062, SD = .045) compared to those who failed (.002, SD = .001). There were no significant differences in the clinical groups (all p's > .05) (see Figure 3).

FIGURE 3 HERE

Finally, any differences between those who passed the prediction question and those who failed the prediction question were investigated for each of the six scenes within each group separately. There was a significant difference in the TD group to the face of the mover hiding the game in scene 4 ($t(12)= 3.365$, $p = .006$) which was still significant when correcting the p-value for multiple comparisons. Those who failed the prediction question

looked longer at the face (.140, *SD* = .099) compared to those who passed (.022, *SD* = .038). Again, there were no differences in the ASD or WS group for any of the scenes.

Discussion

The current study was the first to explore looking behaviour while participants with ASD and WS were administered a ToM task containing dynamic stimuli. The videos in the current task did not include any spoken language and were filmed in black-and-white colours in order to prevent the eye movements being guided by saliency of colours or certain aspects of the narration (i.e., words, intonation, pauses, etc.). In contrast to previous studies (Klin et al. 2002; Riby and Hancock 2009), the current results provide little evidence to support different looking behaviours in the clinical groups. For example, those with WS did not show a consistent atypical preference for faces: although participants with WS looked longer at the face of the human actor at the front of the stage in the false belief condition. This confirms previous studies that individuals with WS have difficulties disengaging from faces rather than just a preference for faces per se (Riby et al. 2011b). Strikingly, participants with ASD did not show avoidance of looking at the faces in favour of objects or looking at the background. This might be explained by differences in the stimuli used in that in contrast to previous studies the current stimuli included a plain background and thus there was not much for the participants with ASD to look at. In addition, the current study included much young children compared those reported in previous studies (Klin et al. 2002; Riby and Hancock 2009). Finally, there are some important differences in the testing paradigm. For example, in the current study participants were asked to view the stimuli freely but they also knew they were going to answer some question about the story at the end and thus, the specific instructions given to the participant might have impacted on their looking behaviour, thus explaining differences in the gaze strategies in the current study compared to previous studies that have measured spontaneous behaviour. In addition, there is evidence that those with ASD and WS have problems with executive functioning and difficulties with integration of information (Bernardino et al. 2012; Happé and Frith 2006; Pellicano 2007; Rhodes et al. 2011). Therefore, the use of silent videos in the current study would have reduced memory and attention demands allowing the looking behaviours in the clinical groups to be undisturbed and similar to control participants. This observation is in line with a previous study by Kelly and colleagues (2013) who found that, although there were no basic oculomotor deficits in participants with ASD, eyemovements in tasks that included voluntary control aligned with

their language abilities: those with language difficulties had difficulties maintaining fixation and shifting gaze (see also Norbury et al. 2014). This has a number of consequences for educational programmes and future research as it suggests that the type of stimuli used has an important impact on performance and looking behaviour for participants with developmental disorders (see also van Buijsen et al. 2011; Van Herwegen et al. 2013). Yet, future studies will need to directly compare the looking behaviours on a verbal and low-verbal task in order to confirm this possibility. Overall, all participants followed the actions depicted in the videos by shifting their gaze to the appropriate objects and locations at the appropriate times, e.g. shifting their gaze to the new location of the object in scene 4 once the hider moves it from its original location.

Secondly, it was investigated whether there were any differences in looking behaviours between those participants who passed and those who failed the prediction question. In the TD group there was a significant difference in the amount of time spent looking to the original location of the object, in that knowledge about where the object was in the first instance provided necessary information about where the seeker would look for the object. There were no significant differences for the clinical groups between those who passed and those who failed the prediction question. This seems to suggest that although the looking behaviour in clinical groups is not necessarily different from typically developing controls, those with ASD and WS who pass the false belief task do not use the same strategies as typically developing children. Yet, there were a varied number of participants who passed and failed within each group and the number of TD participants who failed the false belief task was small. In addition, the current study could not exclude children who did not answer the memory and reality questions correctly (see Van Herwegen et al 2013 for such an approach) when comparing performance on the prediction question due to the uneven numbers per group when excluding these children. It can therefore be argued that the looking behaviours of children who fail the prediction question as well as the control questions could differ from those who passed the control questions. Therefore, further studies are necessary to confirm the current findings. In addition, there are a number of limitations in the current study. First of all, comparisons to previous studies are difficult, not only because of the small number of participants used in these studies in general, but also due to differences in severity of the disorder in the clinical groups across different studies. For example, the ASD group in the current study did not differ from the TD group for receptive vocabulary scores, which suggests that they included mainly high functioning participants with ASD. Yet, the scores

from the parental questionnaires CARS contradicts this. This implies that the ASD group might have been of very mixed ability. However, some participants in the clinical group did not complete the background tasks and for others no accurate age equivalent scores could be calculated, meaning that the contribution of cognitive abilities to ToM performance and looking behaviours is still unclear. Future studies would therefore need to investigate the contribution of cognitive as well as developmental levels to looking behaviours during ToM task performance. Secondly, current analyses of eye tracking data vary hugely and need further improvements. For example, in our study, the video was divided up into time frames thus only proportions of overall looking time to areas of interest were investigated which implicitly assumes that "more looking time is better". However, the ability to shift between different areas of interest might be more informative in a social situation than the length of looking time to particular areas.

The current study is the first to evaluate looking behaviour during a ToM task that involved dynamic stimuli in which the amount of sensory information was reduced. The results show similarities between WS and ASD, two clinical populations that are generally assumed to have opposite looking behaviours when viewing social scenes. In addition, there were no differences in looking behaviours between those who passed and failed the ToM tasks in these clinical groups. Thus, task performance cannot be explained by failure to observe information at the appropriate time and factors such as integration of information will need to be considered in future studies. The current findings add further data on syndrome-specific and syndrome-general differences between WS and ASD which is crucial to plan syndrome-specific interventions and educational programmes that would benefit the child.

References

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

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J. & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief-desire reasoning in children and adults. *Child Development, 82*(5), 1691-703.

Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animatedfaces in autism. *Child Development, 78*, 397–411.

Baron-Cohen, S., Jolliffe, T., Mortimore, C. & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry, 38*(7), 813-822.

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

Bernardino, I., Mouga, S., Almeida, J., van Asselen, M., Oliveira, G., & Castelo-Branco, M. (2012). A Direct Comparison of Local-Global Integration in Autism and other Developmental Disorders: Implications for the Central Coherence Hypothesis. *PLoS ONE, 7*(6), e39351.

Burnette, C. P., Mundy, P. C., Meyer, J. A., Sutton, S. K., Vaughan, A. E., & Charah, D. (2005). Weak central coherence and its relation to theory of mind and anxiety in autism. *Journal of Autism and Developmental Disorders, 35*(1), 63-73.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology, 33*, 781-789

Donnai, D., & Karmiloff-Smith, A. (2000). Williams syndrome: From genotype through to the cognitive phenotype. *American Journal of Medical Genetics, 97*(2), 164-171.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale II*. Windsor: NFER-Nelson Publishing Company Limited.

Happé, F. G. E. (1995). *Autism: An introduction to psychological theory*. Harvard: University Press.

Happé, F. G. E. (1997). Central coherence and theory of mind in autism: reading homographs in context. *British Journal of Developmental Psychology, 15*, 1-12.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of autism and developmental disorders, 36*(1), 5-25.

Karmiloff-Smith, A., D'Souza, D., Dekker, T., Van Herwegen, J., Xu, F., Rodic, M., & Ansari, D. (2012). Genetic and environmental vulnerabilities: the importance of cross-syndrome comparisons. *PNAS, 190*(2), 17261-17265.

Kelly, J. D., Walker, R., & Norbury, C.F. (2013). Deficits in volitional oculomotor control align with language status in autism spectrum disorders. *Developmental Science, 16*(1), 56-66.

Kikuchi, Y., Senju, A., Tojo, Y., Osanai, H., & Hasegawa, T. (2009). Faces do not capture special attention in children with autism spectrum disorder: a change blindness study. *Child Development, 80*, 1421-1433.

Klein-Tasman, B. P., Mervis, C. B., Lord, C. & Phillips, K. D. (2007). Socio- communicative deficits in young children with Williams syndrome: Performance on the autism diagnostic observation schedule. *Child Neuropsychology, 13*(5), 444-467.

Klin, A., Jones, W., Schlutz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809-816.

Laing, E., Butterworth, G., Ansari, D., Gso□dl, M, Longhi, E., Panagiotaki, G., Paterson, S., & Karmiloff-Smith, A. (2002). Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science, 5*(2), 233- 246.

Lincoln, A. J., Searcy, Y.M., Jones, W., & Lord, C. (2007). Social interaction behaviors discriminate young children with Autism and Williams Syndrome. *American Academy of Child and Adolescent Psychiatry, 46*, 323-331.

Liszkowski, U., Albrecht, K., Carpenter, & Tomasello, M. (2008). Infants' visual and auditory communication when a partner is or is not visually attending. *Infant Behavior and Development, 31*(2), 157-167.

Mervis, C. B., Robinson, B. F., Bertrand, J., Morris, C. A., Klein-Tasman, B. P., & Armstrong, S. C. (2000). The Williams syndrome cognitive profile. *Brain and cognition, 44*(3), 604-628.

Norbury, C., Gemmell, T., & Paul, R. (2014). Pragmatic abilities in narrative production: a cross-disorder comparison. *Journal of Child Language*, *41*(3), 485-510.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*(4), 974-990.

Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in Theory of Mind in autism. *Developmental Psychology 46*(2), 530-544.

Porter, M. A., & Coltheart, M. (2005). Cognitive heterogeneity in Williams syndrome. *Developmental Neuropsychology, 27*(2), 275-306.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Raven, J. C., Court, J. H., & Raven, J. C. (1990). *Manual for Raven's progressive matrices and vocabulary scales—section 2: Coloured progressive matrices*. Oxford: Oxford Psychologists Press.

Rhodes, S. M., Riby, D. M., Fraser, E. M., & Campbell, L. E. (2011). The extent of working memory deficits associated with Williams syndrome: Exploration of verbal and spatial domains and executively controlled processes. *Brain and Cognition, 77*(2), 208-214.

Riby, D., & Hancock, P. J. (2009). Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research, 53*(2), 169-181

Riby, D. M., Doherty-Sneddon, G., & Bruce, V. (2011a). Exploring face perception in disorders of development: Evidence from Williams syndrome and autism. *Journal of Neuropsychology, 2*(1), 47-64.

Riby, D. M., Jones, N., Brown, P. H., Robinson. L. J., Langton, S. R., Bruce, V., & Riby, L. M. (2011b). Attention to faces in Williams syndrome. *Journal of Autism and Developmental Disorders, 41*(9), 1228-1239.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *Child Autism Rating Scale.* Western Psychological Services Corporation.

Senju, A., Southgate, V., Miura, Y., Matsui, T., Hasegawa, T., Tojo, Y. & Csibra, G. (2010). Absence of spontaneous action anticipation by false belief attribution in children with autism spectrum disorder. *Development and Psychopathology, 22*(02), 353-360.

Plesa-Skwerer, D., Faja, S., Schofield, C., Verbalis, A., Tager-Flusberg, H. & Dykens, E.M. (2006). Perceiving Facial and Vocal Expressions of Emotion in Individuals With Williams Syndrome. *American Journal on Mental Retardation, 111*(1), 15-26.

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism*, 11, 265- 277.

Tager-Flusberg, H. (2000). Language and understanding minds: Connections in autism. *Understanding other minds: Perspectives from developmental cognitive neuroscience, 2*, 124-149.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition, 76*(1), 59-90.

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language and Hearing Research, 52*(2), 336-358.

Van Buijsen, M., Hendriks, A., Ketelaars, M., & Verhoeven, L. (2011). Assessment of theory of mind in children with communication disorders: Role of presentation mode. *Research in Developmental Disabilities, 32*(3), 1038-1045.

Van Herwegen. J., Dimitriou, D., & Rundblad, G. (2013). Performance on verbal and low-verbal false belief tasks in children with Williams syndrome. *Journal of Communication Disorders, 45,* 440-448.

Van Herwegen, J., Rundblad, G., Davelaar, E. J., & Annaz, D. (2011). Variability and standardized test profiles in typically developing children and children with Williams Syndrome. *British Journal of Developmental Psychology, 29*(4), 883-894.

Võ. M. L., Smith, T .J., Mital, P. K., & Henderson, J. M. (2012). Do the Eyes Really Have it? Dynamic Allocation of Attention when Viewing Moving Faces. *Journal of Vision, 12*(13), 1-14.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128.

Figure

Figure 1 Example of Areas of Interest for the false belief condition

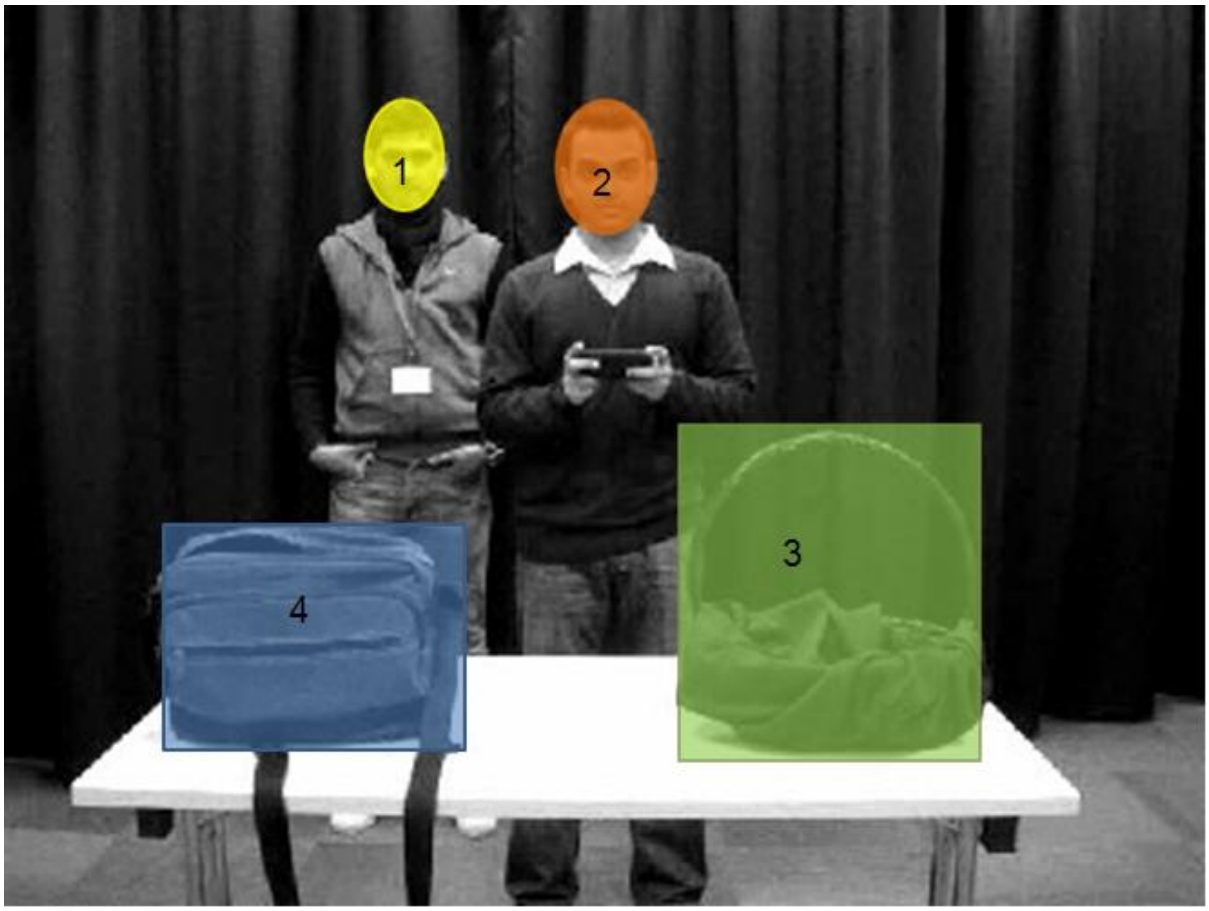


Figure 2 Dwell time to each AOI per group for each of the different scenes

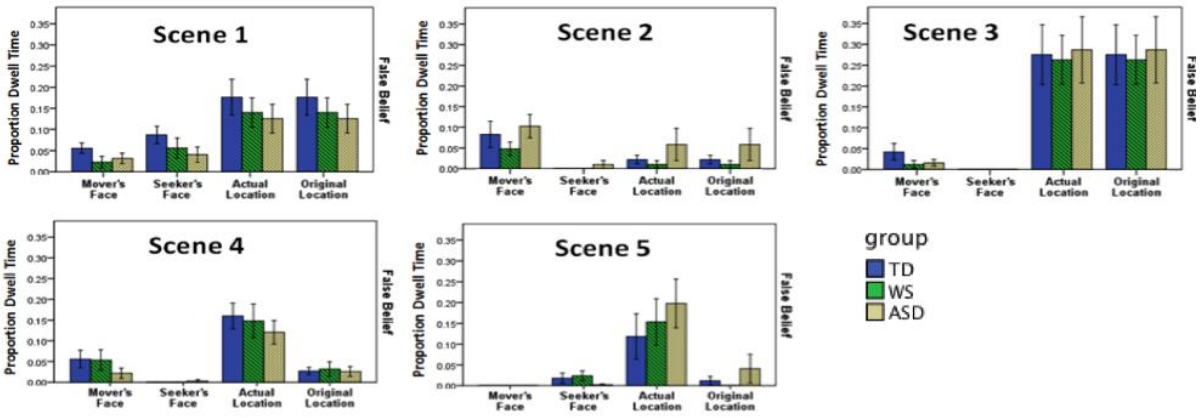
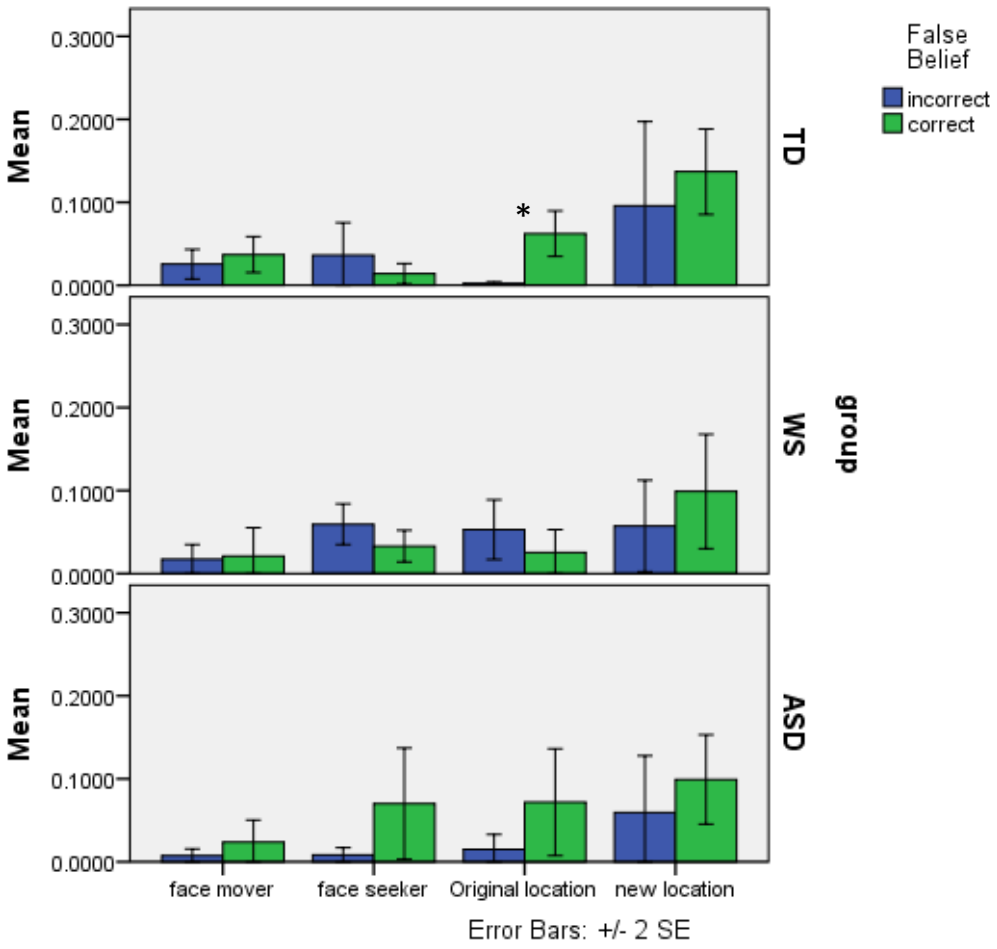


Figure 3 Proportion of total dwell time on each AOI for those who passed and those who failed the false belief condition per group (* $p < .05$).



Table

Table 1 Overview of chronological age (in years; months), raw BPVS and RCPM raw scores per group.

	Mean (SD; range)		
Group	CA	BPVS Raw score	RCPM Raw score
TD	7;0 (1;03: 4;08-9;04)	72.50 (15.698; 43-101)	27.07 (16.069; 12-35)
WS	7;06 (1;07: 5;0-10;07)	49.29 (18.424; 19-83)	13.25 (2.417; 8-17)
ASD	7;08 (1;08: 4;03-10;04)	61.15(29.664; 17-95)	22.20 (5.007; 13-30)

Table 2 Number of participants who passed and failed the reality, memory, and prediction of action questions as well as the prediction of action question (Prediction final) for those who passed the memory and reality questions.

Group	Reality		Memory		Prediction		Prediction final	
	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass
TD	0	14	1	13	3	11	2	11
WS	6	8	7	7	9	5	5	2
ASD	3	10	1	12	5	8	3	6

Table 3 Proportion of total dwell time per group per condition for each of the AOIs

Area of Interest	False Belief		
	Mean (SD)		
	TD	WS	ASD
1) the face of the mover	0.034 (0.032)	0.019 (0.030)	0.017 (0.030)
2) the face of the seeker	0.019 (0.024)	0.050 (0.034)	0.046 (0.079)
3) original location	0.049 (0.047)	0.043 (0.048)	0.050 (0.076)
4) new location	0.128 (0.084)	0.072 (0.080)	0.084 (0.076)