

**Flexibility in young people with Autism Spectrum Disorders on a card sort
task**

Phil Reed¹, Helen Watts¹, and Roberto Truzoli²

¹Swansea University, U.K.; ² Università degli Studi di Milano, Italy

Correspondence address: Phil Reed,

Department of Psychology,

Swansea University,

Singleton Park,

Swansea, SA2 8PP, U.K.

Short title: Behavioural Flexibility

Re-submitted for publication to: *Autism*,

March (2011)

Abstract

Adults with Autism Spectrum Disorders (ASD) have shown deficits in switching between rules governing their behaviour, as have high-functioning children with ASD. However, there are few studies of flexibility in lower-functioning children with ASD. The current study investigated this phenomenon with a group of low functioning children with ASD compared to a mental-aged matched comparison group. The ASD group learned an initial discrimination task as quickly as the matched comparison group, but when the rule governing the discrimination was shifted, the comparison group learned the task with fewer errors, and made the discrimination more quickly than the groups with ASD. These findings suggest that low-functioning children with ASD do display the predicted deficits in extra-dimensional shift.

Key words: Behavioural flexibility, discrimination learning, stereotyped behaviours, low functioning autism

A key part of the diagnosis of an Autistic Spectrum Disorder (ASD) is the presence of restricted, repetitive, and stereotyped patterns of behaviour, which often are manifested by preoccupations with restricted sets of interests and activities, and inflexible routines or stereotyped motor mannerisms, and preoccupations with parts of objects (DMS IV).

Although these behaviours are not restricted to ASD, they are found with high prevalence in this population (Lewis & Bodfish, 1998). Moreover, the finding of greater inflexibility in behaviour among individuals with ASD has formed the basis of a number of discussions of the underlying deficits in individuals with ASD (e.g., Lopez, Lincoln, Ozonoff, & Lai, 2005; Zelazo, Jacques, Burack, & Frye, 2002; see Hill, 2005; Geurtz, Corbett, & Solomon, 2009, for reviews). For example, behavioural inflexibility has taken to reflect problems with shifting attention (Gioia, Isquith, Kenworthy, & Barton, 2002), and with some, but not all, aspects of executive functioning (cf., Lopez et al., 2005; Pennington & Ozonoff, 1996), which is a key theory regarding the central deficits in ASD.

Given the potential importance of repetitive and inflexible behaviours for ASD, attention has focused on the development of experimental tasks in order to explore the factors associated with such inflexible behaviours (e.g., see Truzoli, 2002). A number of procedures have been utilised to investigate flexibility in performance, such as extra-dimensional shift tasks (e.g., Hughes et al., 1994; Truzoli, 2002; Yerys et al., 2009), and card sort tasks (e.g., Lopez et al., 2005; Ozonoff et al., 1991; Zelano et al., 2002), and these procedures have been reviewed by Hill (2004), and by Geurts et al. (2009), in relation to executive function problems in ASD.

The extra-dimensional shift procedure has been used to explore behavioural inflexibility in ASD, and, potentially, to relate it to clinical aspects of ASD (Lopez et al., 2005; South et al., 2007; Yerys et al., 2009), and neural functioning (Gottesman & Gould, 2003). However, despite the potential importance of the EDS paradigm, Yerys et al. (2009;

see also Geurts et al., 2009) have noted that the evidence regarding deficits in this task is not as clear as might be expected. Yerys et al. (2009) note that there are several demonstrations of impaired EDS in high-functioning adolescents and adults with ASD (Ozonoff et al., 2004), but inconsistent evidence for the effect in primary school age children and those with low functioning ASD (e.g., Happe et al., 2006; Lionello-DeNolf et al., 2008). Yerys et al. (2009) documented impaired EDS, and EDS-reversal, in older children with low functioning ASD, and noted that their performance on this task correlated with psychometrically-measured repetitive-behaviour problems. Unfortunately, it should be noted that Yerys et al. (2009) presented no data on the correlations between EDS performance and other psycho-metrically measured aspects of ASD.

Another task that is taken to reflect flexibility in reasoning and executive functioning is the Wisconsin Card Sort Task (WCST), and performance on this task has also been shown to be impaired in those with ASD (see Geurts et al., 2009; Ozonoff et al., 1991). In this task, individuals are required to learn to sort cards, differing from one another along three stimulus dimensions, into a category based on one of the dimensions. After successfully learning the task, the rule is changed, and the number of errors, and perseverative errors, are noted. Similar card sort tasks have also been shown to produce impaired performance in individuals with ASD when the rule governing the sort is changed (see Lopez et al., 2005; Zelazo et al., 2002). Although there is a less discrepant literature associated with these tasks than with the EDS task, the data from those with ASD are largely limited to adults and older children who are higher functioning (e.g., Lopez et al., 2005; Ozonoff et al., 1991). However, there are a few demonstrations of impaired ability to shift sort rules in more severely impaired adolescents (e.g., Zelazo et al., 2002). Lopez et al. (2005) noted that ability to shift rules during a card sort task correlates with restricted behavioural patterns in those with ASD.

If it is the case that problems with flexibility are core in ASD, they should be displayed in a wide variety of participants with ASD. However, both tasks discussed above provide only limited evidence about these abilities with lower functioning young children with ASD (but see Zelazo et al., 2002). Hence, the current study focused on behavioural flexibility in a group of children with low functioning ASD. The current experiment also focused on the use of a simple card sort task, as such tasks appear to present clearer evidence of an impairment (see Geurts et al., 2009, for discussion of the problems with EDS tasks). One procedural reason for this may be that card sort tasks are often presented manually, and it is suggested that computer-based tasks produce worse performance than table-top tasks in people with ASD (although Yerys et al., 2009, did use computerised presentation), and this might be especially so in low functioning young children. As Yerys et al. (2009) suggested that poor performance on EDS tasks may obscure group differences between tasks due to floor effects, a table-top procedure was thought prudent to adopt for the current card sort task.

Thus, the current study aimed to extend the current literature by investigating behavioural flexibility in a sample of low functioning, young children with ASD by using a card sort task, and aimed to relate this performance to psychometrically measured aspects of ASD symptomatology.

Method

Participants

Thirty children (15 with ASD and 15 mental-aged matched typically-developing children) participated (13 boys and 2 girls in each group). The group with ASD were all diagnosed as having either childhood autism or PDD:NOS by a paediatrician who was independent from the study. The diagnosis was made using a combination of DSM-IV criteria for these disorders and clinical judgement. An attempt was made to secure a

homogenous sample with ASD, and were selected according to the criteria of having low functioning ASD (GARS Autism Quotient > 100), and an IQ lower than 80.

To gauge the approximate severity of the ASD in the sample, the Gilliam Autism Rating Scale (GARS) was employed (parent-rated). The GARS measure showed that the mean (standard deviation) of the overall GARS score for this sample was 116.20 (\pm 8.44), indicating that this sample was of a higher than average autistic severity (the GARS has a standardized of 100, implying 'averagely autistic'; higher scores implying greater severity). The scores for the four sub-scales of the GARS, each representing a different aspect of the disorder, showed a similar pattern, in that all of these scores were higher than the average. The sub-scales have a standardized mean of 10, representing 'average severity' (higher scores implying greater severity). The mean sub-scale scores were: Stereotyped Behaviours = 12.87 (\pm 2.42); Communication Problems = 11.93 (\pm 1.58); Social Interaction Problems = 12.47 (\pm 2.20); and Developmental Disturbances = 12.47 (\pm 2.22).

The mental age of group with ASD was matched to the chronological age of the comparison group by use of the Leiter International Performance Scale-Revised (Leiter-R; Roid & Miller, 1997) for non-verbal IQ, and the British Picture Vocabulary Scale (BPVS, Dunn, Dunn, Whetton, & Pintilie, 2002). The mean chronological age for ASD group was 8.33 (\pm 0.77; range = 7.10 to 9.50) years, but their mean nonverbal IQ (Leiter overall score) was 71.27 (\pm 8.51, range = 54 to 80), giving a non-verbal mental age equivalent of 5.93 (\pm 0.86; range = 4.20 to 7.30) years. The verbal mental age of the group with ASD, measured by the BPVS, was similar to the non-verbal measure, and was 5.91 (\pm 0.85; range 4.20 to 7.20) years. The chronological age of the comparison group was 5.92 (\pm 0.83, range = 4.30 to 7.10) years, representing a good match on both variables to the group with ASD.

Materials

The Gilliam Autism Rating Scale (GARS; Gilliam, 1995) comprises four sub-scales, each describing behaviours that are symptomatic of ASD (*Stereotyped Behaviours, Communication Difficulties, Social Interaction, and Developmental Disturbances*). The raw scores from these sub-scales can be converted into standard scores (mean = 10, standard deviation = 3). These sub-scales combine to give an overall *Autism Quotient*; higher scores meaning greater autistic severity (mean = 100 [average autistic severity], standard deviation = 15). In terms of assessing the probability that an individual has ASD, an *Autism Quotient* score of between 90 to 110 means an ‘average’ probability of ASD, a score below 89 means that there is a ‘below average’ probability of ASD, and a score below 79 means that there is a ‘low’ probability that the individual has ASD (Gilliam, 1995). The scale is appropriate for persons aged 3 to 22 years old, and is completed by parents or professionals in about 10 minutes. Its internal reliability is 0.96, and it has high criterion validity with the Autism Behavior Checklist (0.94). While not a strong diagnostic tool, this measure can an estimate of the relative severity of the disorder (see Reed, Corness, & Osborne, 2010).

Leiter International Performance Scale-Revised (Leiter-R; Roid, & Miller, 1997) is a measure of intellectual abilities of individuals (aged 2 to 20 years) with significant communication disorders, cognitive delays, and various types of learning disabilities. It has been used extensively for the assessment of children with ASD (Tsatsanis, Dartnall, Cicchetti, Sparrow, Klin, & Volkmar, 2003). It consists of subtests that are organised into four domains: Visualisation, Reasoning, Memory, and Attention, from which only the visualisation and reasoning battery was used. Specifically, the subtests that were used were: Figure Ground, Design Analogies, Form Completion, Matching, Sequential Order, Repeated Patterns, Classification, and Paper Folding. The Leiter-R provides a total nonverbal IQ score (mean = 100 & SD = 15). The reliability for the full IQ for different age groups varies from

0.91 to 0.93, and the test-retest reliability varies from 0.61 to 0.90. It has a 0.85 correlation with the WISC-III full scale IQ measure.

British Picture Vocabulary Scale (BPVS; Dunn et al., 1982) is derived from the Peabody Picture Vocabulary Scale, and measures receptive language ability. The BPVS is standardized for use on children in the U.K. between 3 and 17 years old, and gives an age equivalent score for this ability. It has an internal reliability of 0.93, and has a 0.59 correlation with the Reynell Comprehension Scale.

Conditioning Materials. Two sets of sixteen cards (100mm x 75mm) were employed. In each set, each card showed one of four pictures (flower, tree, apple, car), in one of four colours (red, yellow, green, blue).

Procedure

Consent for participation was obtained from the parents of the children, who also completed and returned the GARS. The participants were tested while sitting at a table located in a quiet room that was free from distractions. Initially, the participants with ASD were given the Leiter and the BPVT to establish their mental age.

Phase 1 of the experiment used 16 cards described above, each showing one of four pictures (i.e. flower, tree, apple, and car), in one of four colours (i.e. red, yellow, green, and blue). The participants were given the following instructions: "*We are going to play a game. You have to put these cards into four piles, so that the cards that are like each other are put in the same pile. You do it like this.*" Then the experimenter shuffled the cards for 5s-10s, and placed them face down in a single pile on the table. The experimenter then picked up one card at a time, and sorted them into four piles, according to the rule that the children were going to be required to follow (i.e. by shape by half of the participants, or by colour for the other half of the participants).

After the instructions and demonstration, the experimenter shuffled the cards again, placed them face down in a single pile in front of the participant, who were asked to put them into four piles, in the same way as they had just seen demonstrated by the experimenter. The rule determining which response was correct was counterbalanced across the participants. For half the participants, sorting the cards into piles of different colours was considered a correct; for the remaining participants, sorting the cards by shape was considered correct. If each card was placed in the pile according to the correct rule, the participant received verbal praise. If the participant sorted the cards by another variable, then this was considered to be incorrect, and the participant was told that they had done it wrongly. Each trial lasted until all 16 cards had been placed into the four piles. Between each trial, the cards were collated into one large pile and shuffled. When the participant had successfully sorted the cards according to the appropriate rule for three consecutive trials, the phase ended.

Phase 2 commenced immediately after the termination of Phase 1, but this was not signalled to the participants (and no modelling of responses was performed). As in Phase 1, the participants were given sixteen shuffled cards (4 pictures in 4 colours) to sort into four piles. In this phase, a correct trial was when the participant had sorted the cards by the alternative dimension to that trained in Phase 1 (i.e. shape if they had previously sorted by colour; and colour if they had previously sorted by shape). Each trial lasted until all the 16 cards were placed in piles. As in Phase 1, the cards were shuffled between each trial. When 3 consecutive trials were correctly sorted, the experiment ended. A limit of 10 trials was placed on this phase to avoid unnecessary distress for those participants who could not master the task.

Results

 Figure 1 about here

Figure 1 displays the number of trials taken to reach criterion in each phase in the two groups. Inspection of these data shows little difference between the groups in acquiring the task in Phase 1. However, in Phase 2 the comparison group mastered the task in fewer trials than the group with ASD: the former group taking about the same number of trials to criterion as in Phase 1; whereas, the group with ASD required substantially more trials to reach criterion in Phase 2 compared to Phase 1.

These data were analysed by a two-factor mixed-model analysis of variance (ANOVA), with group (ASD versus comparison group) as a between-subject factor, and phase as a within-subject factor. This analysis revealed statistically significant main effects of group, $F(1,28) = 42.62, p < 0.001$, partial $\eta^2 = 0.60$, and phase, $F(1,28) = 97.47, p < 0.001$, partial $\eta^2 = 0.78$, and a statistically significant interaction between the two factors, $F(1,28) = 61.85, p < 0.001$, partial $\eta^2 = 0.69$. To further analyse the interaction, simple effects were conducted between the groups for each phase. In Phase 1, there was no statistically significant difference between the groups, $p > 0.09$, partial $\eta^2 = 0.09$. In Phase 2, the comparison group acquired the task statistically significantly faster than the group with ASD, $F(1,28) = 98.65, p < 0.001$, partial $\eta^2 = 0.79$. Further simple effect analyses were conducted separately for the two groups comparing their performance on each phase. These analyses revealed no statistically significant difference for the comparison group between the phases, $p > 0.20$, partial $\eta^2 = 0.13$, but the group with ASD made statistically significantly more errors in the second phase compared to the first, $F(1,14) = 147.84, p < 0.001$, partial $\eta^2 = 0.91$.

Figure 2 about here

Figure 2 shows two measures of perseverative responding both based on the first trial in Phase 2. The left side of Figure 2 shows the total number of perseverative errors made in the first trial of card sorting in Phase (defined as sorting cards by the original rule acquired in Phase 1). The right side of Figure 2 shows the number of times in a row that participants sorted the cards by the rule learned in Phase 1 on the first trial of Phase 2. Both measures show a greater number of perseverative errors for the group with ASD than for the comparison group. Separate ANOVAs conducted on these data shows statistically significant differences for the total number of perseverative errors, $F(1,28) = 27.18, p < 0.001, \eta^2 = 0.49$, and for initial perseverative errors in a row, $F(1,28) = 16.07, p < 0.001, \eta^2 = 0.37$.

Table 1 about here

Table 1 shows the correlational data for the participants with ASD between the four subscales of the GARS, as well as for the Leiter standard scores, and the verbal age as measured by the BPVS, and the four measures reported above (trials to criterion in Phase 1 and 2; the total perseverative errors in trial 1 of Phase 2, and the number of initial perseverative errors made in trial 1 of Phase 2). These data show that the stereotyped behaviours subscale was the only scale reliably associated with the measures. Although it may be worth brief note that the correlations for the Developmental Disturbance sub-scale of the GARS were all consistently just under the $p < 0.05$ level for statistical significance.

Discussion

The current study attempted to examine whether behavioural flexibility was impaired in a sample of lower functioning children with ASD, compared to a mental-aged matched comparison group, using a card sort task. The results showed that both groups learned an initial discrimination task as quickly as one other. This finding replicates previous findings about the ability of children with ASD to learn such discriminations (see Lionello-DeNolf et al., 2008; Yerys et al., 2009; Zelazo et al., 2002), and rules out potential explanations involving motor or perceptual deficits. However, in the second phase of the experiment, which involved shifting the rule by which the cards were sorted, the group with ASD showed more errors, and greater levels of perseverative responding.

The levels of flexibility were associated with psychometrically-measured aspects of the childrens' ASD related to stereotyped behaviours, which extends the findings documented by Lopez et al. (2005), and Yerys et al. (2009), to a lower functioning sample. Thus, these findings extend the range of individuals with ASD for whom behavioural inflexibility has been demonstrated, suggesting that this is a common problem for all with ASD, and also provide evidence that the current experimental task is strongly related to clinical aspects of ASD associated with stereotyped behaviours and restricted interests (see also Lopez et al., 2005).

The current task differs slightly from those previously employed in the study of behavioural flexibility, such as EDS (e.g., Hughes et al., 1994; Yerys et al., 2009), and the WCST (Ozonoff et al., 1991). In part, a new task was employed due to the nature of the sample being employed, as it was thought to be more suited to the abilities of this group. However, the new task also raises several possibilities for testing potential mechanisms underlying such abilities, and the reasons why the WCST appears to produce more reliable demonstrations of impaired flexibility than the EDS task. There are a number of differences

between the two tasks. One, noted above, concerns the mode of presentation; either through table-top (more typical of WCST) or computer (more typical of EDS) designs. A second difference concerns the fact that the EDS task presents one irrelevant dimension during training before transfer, whereas the WCST presents two irrelevant dimensions in this phase. It may be that presentation of multiple irrelevant cues produces differences in performance between the two tasks. For example, this may make the initial discrimination harder, producing greater numbers of exposures to the stimuli, and, subsequently, making it harder to learn about those stimuli (e.g., through latent inhibition, which refers to the process whereby exposure to irrelevant stimuli makes subsequent learning about those stimuli harder). Similarly, multiple exemplars of negative classes might make learning about the target dimension stronger, and subsequently harder to reverse.

The current task employed a card sort task like the WCST, but with only one irrelevant dimension (similar to that used by Zelazo et al., 2002; and also suggested as appropriate by Geurts et al., 2009), and found a difference in ability to perform the task between those with ASD and mental-aged matched comparison participants. This suggests that impairments in flexibility are demonstrable with one irrelevant dimension using a table top card sort procedure. However, these aspects of the task could be varied in further studies to narrow down other controlling variables in the task. This might be particularly interesting exploring the concept or category learning abilities in those with ASD, which appear to vary depending on the task, and this factor may be worth exploring in this context (e.g., see Bott et al., 2006).

It is also important to note that the GARS measure employed in this study has been shown to underestimate the severity of ASD in a number of reports, which is problematic for diagnostic purposes (see South et al., 2002). However, as diagnosis was not the purpose of this study, and since the GARS has been used in previous studies of interventions for ASD

(Reed et al., 2007; see also Jesner, Aref-Adib, & Coren, 2007, for a review of pharmacological studies that employ the GARS), it was considered suitable for use as a research tool in terms of demonstrating the relative relationship between stereotyped behaviours and performance on a behavioural flexibility task. Moreover, the use of different scales from previous studies (e.g., Lopez et al., 2002; Yerys et al., 2009) extends the generality of these findings, and demonstrates that they are not measure specific (see Osborne & Reed, 2009). Importantly, the current study demonstrated that such associations between psychometrically-measured aspects of ASD and task performance are limited to stereotyped behaviours, and do not appear to be associated with other aspects of the disorder.

In summary, the current study demonstrated behavioural inflexibility in a homogeneous group of low functioning young children with ASD. This extends the demonstration of this deficit from the existing literature. Moreover, the current findings also demonstrated an association between performance on the experimental task and a measure of stereotyped behaviour problems in ASD.

References

- Bott, L., Brock, J., Brockdorff, J., Boucher, J. & Lamberts, K. (2006). Perceptual similarity in autism. *Quarterly Journal of Experimental Psychology*, **59**, 1-18.
- Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale* (Revised ed.). Slough, Bucks: NFER-Nelson
- Dunn, L.M., Dunn, L.M., Whetton, C., & Pintilie, D. (1982). *The British Picture Vocabulary Scale*, NFER- Nelson.
- Gilliam, J.E. (1995). *Gilliam Autism Rating Scale*. Austin, TX.: Pro-Ed.
- Gioia, G.A., Isquith, P.K., Kenworthy, L., & Barton, R.M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, **8**, 121-37.
- Gottesman, I.I., & Gould, T.D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, **160**, 636-645.
- Geurts, H.M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in Autism. *Trends in Cognitive Sciences*, **13**, 74-82
- Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in Autism Spectrum Disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, **61**, 25-39.
- Hill, E.L. (2004). Evaluating the theory of executive dysfunction in Autism. *Developmental Review*, **24**, 189-233.
- Hughes, C., Russell, J. & Robbins, T.W. (1994). Evidence for executive dysfunction in Autism, *Neuropsychologia*, **32**, 477-92.
- Jesner, O., Aref-Adib, M., & Coren, E. (2007). Risperidone for autism spectrum disorder. *Cochrane Database Syst Review*, **24**, CD005040

- Lewis, M.H., & Bodfish, J.W. (1998). Repetitive behavior disorders in autism. *Mental Retardation and Developmental Disabilities Research Reviews*, **4**, 80–89
- Lionello-DeNolf, K.M., McIlvane, W.J., Canovas, D.S., de Souza, G.D., & Barros, R.S. (2008). Reversal learning set and functional equivalence in children with and without Autism. *Psychological Record*, **58**, 15-36.
- Lopez, B.R., Lincoln, A.J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of autistic disorder. *Journal of Autism and Developmental Disorders*, **35**, 445-460.
- Osborne, L.A., & Reed, P. (2009). The relationship between parental stress and behavior problems of children with autistic spectrum disorders. *Exceptional Children*, **76**, 54-73.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, **32**, 1081–1105.
- Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R.M., et al. (2004). Performance on Cambridge Neuropsychological Test Automated Battery subtests sensitive to frontal lobe function in people with Autistic Disorder: Evidence from the collaborative programs of excellence in Autism. *Journal of Autism and Developmental Disorders*, **34**, 139-150.
- Pennington, B.F. & Ozonoff, S. (1996). Executive functions and developmental psychopathologies. *Journal of Child Psychology & Psychiatry*, **37**, 51-87.
- Reed, P., Osborne, L.A., & Corness, M. (2010). Effectiveness of special nursery provision for children with autistic spectrum disorder. *Autism*, **14**, 67-82.
- Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale—Revised*. Wood Dale, IL:

- South, M., Ozonoff, S., & McMahon, W.M. (2007). The relationship between executive functioning, central coherence, and repetitive behaviors in the high-functioning Autism Spectrum. *Autism*, **11**, 437-451.
- Truzoli, R. (2002). *Sviluppare il pensiero flessibile. Un programma d'intervento*. Roma: Armando Editore.
- Tsatsanis K.D., Dartnall N., Cicchetti D., Sparrow S.S., Klin A., Volkmar F.R. (2003). Concurrent validity and classification accuracy of the Leiter and Leiter-R in low functioning children with Autism. *Journal of Autism and Developmental Disorders*, **33**, 23-30.
- Yerys, B.E., Wallace, G.L., Harrison, B., Celano, M.J., Giedd, J.N., & Kenworthy, L.E. (2009). Set-shifting in children with autism spectrum disorders: Reversal shifting deficits on the Intradimensional/Extradimensional Shift Test correlate with repetitive behavior. *Autism*, **13**, 523-539.
- Zelazo, P.D., Jacques, S., Burack, J.A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with Autism-Spectrum Disorders. *Infant and Child Development*, **11**, 171–195.

Figure Captions

Figure 1. Number of trials to criterion for both groups in both phases of the study (error bars = standard deviation).

Figure 2. Number of perseverative errors for trial 1 of Phase 2; left panel = total perseverative errors; right panel = initial number of perseverative errors in a row (error bars = standard deviation).

.

Figure 1

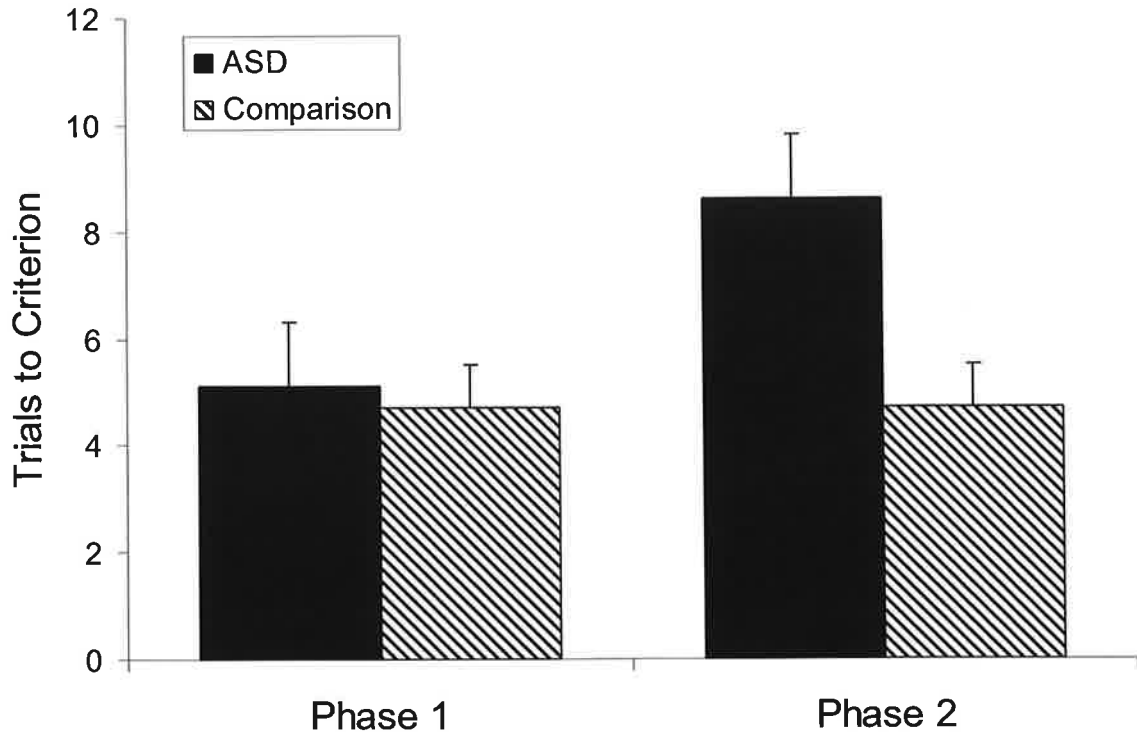


Figure 2

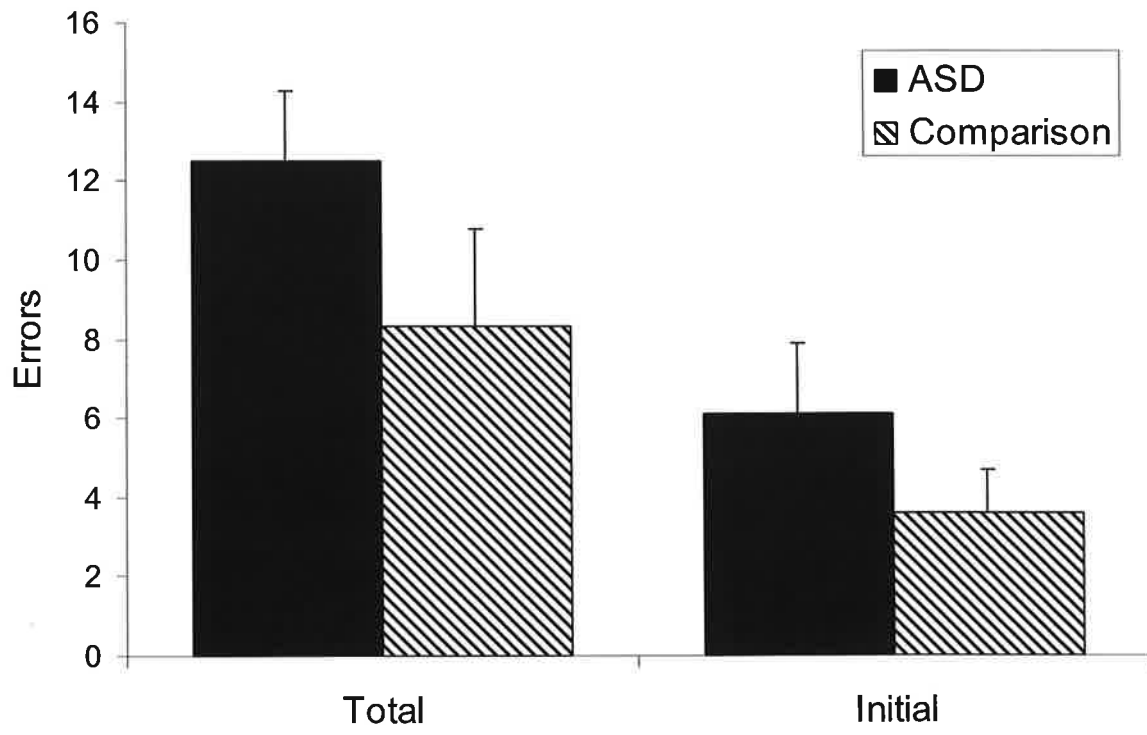


Table 1: Pearson's correlations between GARS subscales, the Leiter and BPVS scores, and the performance indicators (Trials to Criterion in Phases 1 and 2; Total Perseverative Errors in Trial 1 of Phase 2; and Initial Perseverative Errors in Trial 1 of Phase 2).

	Trials to Criterion Phase 1	Trials to Criterion Phase 2	Total Perseverative Errors (Trial 1 Phase 2)	Initial Perseverative Errors (Trial 1 Phase 2)
Stereotyped Behaviours	0.47	0.60**	0.69**	0.53*
Communication Difficulties	-0.20	0.02	- 0.19	-0.50
Social Interactions	0.18	0.07	0.26	0.10
Development Disturbances	0.50	0.47	0.32	0.46
Leiter Standard	-0.15	-0.29	-0.38	0.02
BPVS (age)	0.29	0.34	0.32	0.11

* $p < 0.05$, ** $p < 0.01$