Using inclusive sampling to highlight specific executive functioning impairments in autism spectrum disorder

Maggie McGonigle-Chalmers, Fiona McCrohan

Psychology Department, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK

Objectives: The aim of the study is to help identify the nature of impaired executive functioning (EF) in children with autism spectrum disorder (ASD). It is also argued that participant sampling by age alone should inform experimental research on EF, as selection through IQ matching may weaken any experimental effects.
Methods: Sixteen children with ASD across a wide range of Nonverbal IQ (NVIQ) and 16 neurotypical control children matched on age alone were given two different types of computerized sequencing game. Both required top-down organization, but in one case the sequence had to be self-generated while in the second it had to be learned and strictly followed. Measures of learning success in relation to NVIQ, and information processing demands were made.
Results: Children with ASD were significantly impaired on the first task only, especially when the processing demands were increased. The effects were particularly pronounced for children with below average NVIQ.
Conclusions: The study indicates a selective problem with self-organized sequencing in ASD with implications for certain real world contexts, but also points to a need for more inclusive sampling of children in order to fully expose specific executive impairments in autism spectrum disorder.

Keywords: autism spectrum disorder, executive functioning, sequencing, nonverbal IQ, working memory

Children with autism spectrum disorder (ASD) show sufficient atypicalities in their cognitive functioning to feature in specific 'cognitive' accounts of the autistic phenotype even when their full-scale IQ falls within the normal range, i.e. those diagnosed as 'high functioning' (see Rajendran and Mitchell 2007 for a review). Difficulties with executive (goal-directed) tasks that require a combination of forward planning, response flexibility and working memory (WM) are particularly implicated (O'Hearn *et al.* 2008; Ozonoff *et al.* 2004; Pennington and Ozonoff 1996). But although such executive problems continue to be identified in children diagnosed with autism or Asperger Syndrome, their precise root causes and how these might impact on lower functioning children remains to be clarified.

One problem arises from sample selection. To effect a closer comparison with typically developing controls (likely to have a higher mean IQ), child and adolescent participants are often matched on the basis of IQ as well as by age — a practice that has been described as likely to 'wash out' the very impairments that are under investigation (Bardikoff and McGonigle-Chalmers 2014; McGonigle-Chalmers and McSweeney 2013). This in itself could go some way to explaining why executive difficulties are not invariably reported (Barendse *et al.* 2013; O'Hearn *et al.* 2008) and when they are, are sometimes reported as 'subtle' in nature (Goldberg *et al.* 2005). To enable such matching furthermore, the sampling may be restricted to children with a full scale IQ above 80 (e.g. Geurts *et al.* 2004; Semrud-Clikeman *et al.* 2010; Sinzig *et al.* 2008). This automatically demands a reasonably high verbal as well as nonverbal IQ, excluding many language delayed children with autism, even when the study is specifically aimed at nonverbal intelligence such as visual search (Horlin *et al.* 2016; Keehn and Joseph 2016). Where the IQ range is allowed to extend into the borderline, below average bound, moreover, effects of lowered IQ on test scores are sometimes described as 'due' to IQ rather than to autism (Mari *et al.* 2003). It has even been suggested that training of cognitive skills in below average intelligent children with ASD may be 'a less fruitful endeavour' than with more intelligent children (Rommelse *et al.* 2015). One argument put forward by these authors for this suggestion is that there are 'cognitively different profiles' affecting individuals at the high and low end of the IQ spectrum.

A second issue is one that applies to all research on Executive Functioning (EF) in Autism Spectrum Disorder

Correspondence to: Maggie McGonigle-Chalmers, Psychology Department, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK. Email: M.McGonigle@ed.ac.uk
The research represented in this article was conducted by the second author for the degree of Master of Science, Edinburgh University.

DOI 10.1080/20473869.2017.1288887

(ASD), and that is the specification of what is being measured and what is thought to be impaired. There is strong agreement there are at least three separable components of executive functioning, comprising set-shifting, updating and response inhibition (Miyake *et al.* 2000). There has also been a long-standing consensus that these are all implicated in executive impairments in ASD (Frith 2003; Russell 1997) and can separately or interactively explain the broader concepts of inflexibility and poor goal-directed planning. For example, impairments in set-shifting (from one type of correct response to another) have been identified in tasks such as the WCST (Kaland *et al.* 2008; Ozonoff 1995; Pennington and Ozonoff 1996; Rumsey and Hamburger 1988) and the ID/ED task (Hughes *et al.* 1994; Ozonoff *et al.* 2004; Pascualvaca *et al.* 1998). Difficulties in response inhibition (also a possible cause of inflexibility) have been captured in e.g. the Go-NoGo, and Stroop tests (Robinson *et al.* 2009; Verté *et al.* 2006) and the 'box' task devised by Russell *et al.* (1991). Failure of response inhibition has even been thought to lie at the root of difficulties in the famous false belief tasks used to assess Theory of Mind in autism (see McGonigle-Chalmers 2015 for a brief summary). Whilst set-shifting and inhibitory impairment could both lead to an apparent failure of information 'updating', this has also been identified as a core failure in its own right. Restrictions on working memory in individuals with ASD using tasks ranging from standard digit span to visual search (Goldberg *et al.* 2005; Pellicano *et al.* 2010) are regularly reported (see Kercood *et al.* 2014 for a recent review), and are thought to reflect neurological abnormalities in development extending into adolescence (O'Hearn *et al.* 2008).

It is now commonplace, therefore, for studies on EF in autism to feature standardized test batteries based on traditional tests tapping into these 'core' abilities, but findings are mixed and even 'often contradictory' (Craig *et al.* 2016). In recent a recent review of EF in autism and related disorders, these authors considered a total of 26 studies using measures ranging from parent teacher questionnaires to Stroop, Towers, WCST and CANTAB (Craig *et al.* 2016). Collectively these were aimed at testing (among others) working memory, response flexibility, and forward planning, but in each case, some studies showed a deficit for ASD children as compared with TD controls whilst others did not. Apart from the difficulties in comparing across disparate studies (where methodology and sampling practices can vary), some of the traditional tests they employ (such as the WCST) are thought to tap into multiple abilities (Ozonoff 1995; Snyder *et al.* 2015) — a problem that has been described as 'task impurity' (Roelofs *et al.* 2015). As Snyder *et al.* (2015) put it, traditional tasks can be 'too broad to answer fine-grained questions about specific aspects of EF'. A further problem in comparing across different tasks is that impairment is found mainly (and sometimes only) where there is a heavy load on the information to be monitored (Goldberg *et al.* 2005; McGonigle-Chalmers *et al.* 2008; Robbins 1997). Informational load has in fact been identified by Barendse and colleagues as the main determiner of whether or not spatial working memory will be reported as impaired in ASD (Barendse *et al.* 2013).

Group measurement based on a simple scoring algorithm, furthermore, can obscure the important detail regarding how different individuals actually engage with a task. This is particularly relevant where participants with a low IQ are concerned as it could impede an answer to whether deficits are qualitatively different for children in different ability groups. Autism-related task difficulties could be mild or subtle in others but similar in nature to those expressed more severely in other individuals. Alternatively, impairment in children with below average NVIQ may indeed represent something qualitatively different. Either conclusion requires the combination of a task that focuses on a particular skill and an analysis of task performance that goes beyond a simple accuracy score. It may even require scrutiny at the level of individual participants (Geurts *et al.* 2004) in order to be sure that task instructions were fully understood, which would be particularly relevant where participants with a low IQ are concerned.

Sequencing tasks have been found to be a transparent means of identifying specific aspects of EF impairment in ASD. Sequencing atypicalities have been noted since Frith reported differences between autistic and neurotypical participants in tasks ranging from serial pattern prediction to verbal recall (Hermelin and O'Connor 1970). In each case, the difficulty was related to the extent of overall organization imposed by the participant whilst recalling or enacting a sequence. For example, in a task in which colored counters were disclosed one at a time, predictions by individuals with autism as to what would 'come next' were perfectly valid but restricted to immediately preceding elements, whilst control participants predicted on a gradually emerging global pattern (Frith 1970).

These atypicalities represent what later became known as 'weak central coherence' (Happé and Frith 2006) and occur where it is important to process information that connects items according to an overall rather than a local organizing principle. This is sometimes described as top-down control (Greenaway and Plaisted 2005). Whilst related to response flexibility, this concept bears particularly on the ability of the participant to consider connections that go beyond the immediate or local. Laboratory tasks have identified selective difficulties in picture sequencing by participants with ASD when connections have to be made according to an overall narrative as in the 'strange stories' test (Happé 1994). It has also been found in terms of a failure by participants with ASD to use a suitably broad hierarchical categorizing principle in order to be maximally efficient at selecting the order of questions during the Twenty Questions Task (TQT) (Alderson-Day 2011). Similarly, anecdotal difficulties have been observed in managing to make broad efficient searches for items in information rich environments such as supermarkets (http://www.autism.org.uk/living-with-autism/understanding-behaviour/organising-sequencing-and-prioritising.aspx). Implied in these tasks is the participant's ability to *self*-impose an organizing strategy on the items, i.e. evaluating the broader picture oneself before settling on the most appropriate organizing principle.

Where self-organization is made the clear objective of the experimental task, studies have yielded significant effects of a clinical diagnosis of autism. One case is the self-ordered sequencing task (SOPT) devised by Petrides and Milner (1982) where the objective is to point each time to a different object from a spatially randomised array of pictures across successive presentations, i.e. to impose a strategic non-reiterative search. This has been found to be impaired in school age children with ASD using nameable stimuli (Joseph *et al.* 2005) as well as abstract shapes (Verté *et al.* 2006).

Sequencing tasks can also be easily manipulated in terms of informational load (found to impact particularly on children with ASD in the SOPT task). The number of items to be sequenced was found to explain a very significant impairment in size sequence learning by children with ASD (McGonigle-Chalmers *et al.* 2008). These stimuli, however, varied only on one dimension. An obvious reason why information load ought, in principle, to impact particularly on sequencing with more complex stimuli is that they may need top-down control. Hierarchical nesting (as in tinned goods, fruit, peaches, etc. in the supermarket example) is a method of maintaining coherence across multiple complex items, but it requires combining broader categorical connections with local ones and remembering this structure whilst the search is ongoing.

Although commonplace in everyday situations, inviting sequencing through categorical nesting is unusual in autism research. One exception was a computer-game based study where the stimuli were chosen to elicit such nesting when attempting to put them in a sequence (McGonigle-Chalmers and Alderson-Day 2010). It was expected that this would be enabled by self-organization into e.g. groups of shapes then colors. But, by contrast with the study on size sequencing, there was no main effect of clinical diagnosis of ASD in children aged from 9 to 16 years. However, as conceded in the article, although intended to induce complex, hierarchical processing, the task may not have succeeded in that objective, as the stimulus set allowed children to opt for a simple one-dimensional solution. That is, shape or color could be employed as an organizing principle that would reduce working memory demands to manageable limits, without the need to nest items within these groupings.

Using a similar task, the current study sought to make the information processing demands more explicitly multidimensional and top-down in nature. This was effected by subdividing the set into only two colors so that the remaining subsets (6 or 8) would require further specification by both shape and size to minimize the working memory demands. The current study also sought to clarify the difference between self-organized vs. imposed sequential control. Accordingly, a fixed search task using similar multidimensional stimuli was included, but employing a hierarchically organized sequential rule that the participants were obliged to follow. Whilst this task required the participants to flexibly follow changes in the sequential order with increasing task difficulty, they did not have to impose the plan themselves. The property in common to both tasks is that they can only be completed if three dimensions can be systematically maintained in working memory (color, shape, and size).

Finally, in order to consider the broader relationship between task performance and other indices of intellect, all children were assessed in terms of nonverbal IQ (NVIQ), allowing for a full range of scores from the borderline normal of 70 (below average) to 130 (above average). Participants were selected exclusively on the basis of age and diagnosis, and no children with ASD were excluded for purposes of 'matching' with controls.

The study below focuses specifically on comparing self-organized vs. imposed sequencing under conditions of high information load on stimulus classification and the need for top-down control. Whilst hypothesizing that the pressure on multiple classification would selectively affect children with ASD, particularly in a self-organizing task, it was an open question as to whether inclusive sampling would illuminate, or obscure, any such impairment in the clinical group.

Method

Participants

A mixed clinical group of 16 children with ASD (15 male, 1 female) and control group of 16 typically developing (TD) children (11 male, 5 female) were recruited locally. The ASD participants were pupils at a special education school with an intake tending towards higher functioning children on the autistic spectrum,[1] while the TD participants were recruited from two mainstream high schools. The ASD participants were all diagnosed by a multi-disciplinary team using Autism Diagnostic Interview-Revised (ADI-R) (Lord *et al.* 1994). According the criteria in use at the time (which allowed a differential diagnosis of Asperger's Syndrome — a category no longer acknowledged in the DSM-5), 13 participants were diagnosed with autism whilst, three were diagnosed with Asperger's Syndrome. Both the ASD and the TD group were aged between 12 and 15 years; the ASD mean age was 13.08 (SD = .862), and the TD mean age was 12.58 (SD = .90). There was no significant age difference between the groups: t (30) = 1.69, $p > .05$).

Nonverbal IQ

Whilst the Special school held records of some previous IQ testing, these were variable in terms of age of testing and instrument used. To ascertain the nonverbal intelligence levels of both groups at the time of testing, the Kaufman Assessment Battery for Children (K-ABC II) was used (Kaufman and Kaufman 2004). The nonverbal index alone was used because of its general relevance to our nonverbal task, and also because of the high variability in language onset in the ASD and AS participants. The five subtests of the K-ABC II (story completion, triangles, block counting, pattern reasoning and hand movements) were administered on a different day to the computer game to ensure that cognitive fatigue was not an issue for either of the groups.

Design

The experiment was based on a mixed design using one inter-subject factor (ASD v TD) and one intra-subject factor (free v fixed search). All subjects performed the free search task first, as the second would have acted as a clue to solving the first. Increasing levels of difficulty were presented for each task, and the main dependent measures were number of attempts made at each task level and percentage correct. Ancillary measures were the number of correct touches made before exiting a trial and improvement across trials (a measure of learning).

Stimuli

The stimulus set consisted of two groups of 16 shapes: square, triangle, circle, and star in the colors red and green and sizes large and small, or rectangle, spiral, diamond, and stellate in colors blue and yellow, also in sizes large and small. Half the participants were randomly assigned to one set of stimuli (red and green shapes) in the free search levels, and contrasting shapes and colors (blue and yellow) in the fixed search levels, while other participants were presented with the reverse shape and color allocation.

Tasks

Free search

This was a self-ordered task using shape icons on a touch screen. The first level — FreeSearch1 (hereafter FS1) used three shapes each with two colors and two sizes presented in a random layout. The task was to simply select (click on) each of the 12 items once only in any order, but the shapes moved around to a new random presentation after every selection. The learning criterion was to successfully complete this task twice consecutively, after which the next task (FS2) was presented. Although participants were free to attempt the task as often as they liked, a skip mechanism controlled by the researcher was available throughout testing. This was employed to reduce frustration of participants — in particular among the ASD group — and permitted them to proceed to the next task, irrespective of performance on the previous task. Thus FS1 was followed (irrespective of whether criterion was reached) by exactly the same task (FS2) but with one new shape (in two colors and two sizes) added, totaling 16 in all. An example is shown in Fig. 1.

Fixed search

This required participants to follow a fixed sequential order across three levels of difficulty (4, 8 and 16 items), using the new stimulus set. Because it would be almost impossible to determine the correct sequence from a large set, the sequence progressed in complexity, updating the basic shape order at each new level by adding size variation and then color variation, denoted hereafter as FxS1, FxS2, and FxS3. The new variants produced a consistent hierachical nesting as depicted in Fig. 2.

The learning criterion of two consecutive successful trials applied throughout these trials and children continued until they finished the game. The skip mechanism was not necessary for this task.

Game features

The computer program FLASH was used to develop the game, which was presented on a portable 13 in. *Eee* computer. Participants used the mouse to select the shapes. The game was designed to be as entertaining as possible, using interactive features, and both visual and auditory feedback on immediate touch-by-touch and trial-by-trial basis. Immediate feedback was illustrated by a high-pitched tone of praise; 'Well done, you scored a point!' On completion of correct sequence, participants were visually rewarded with the creation of half of a rainbow. An error would result in the loss of this rainbow. Thus, two consecutive

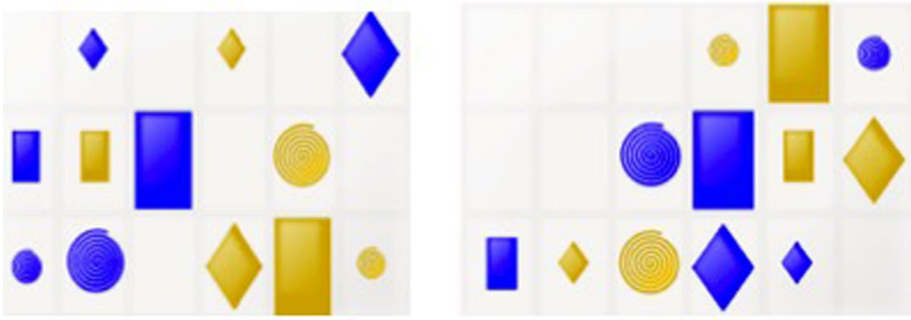

Figure 1 Example of the screen layout for the search tasks showing 12 items on two consecutive trials. In Free Search (FS1 and FS2), 12 then 16 items were presented and the participant had to choose their own route through the set. In Fixed Search (FxS1, FxS2, and FxS3), 4, 8 and 16 items were displayed in a similar manner but the correct sequence was predetermined.

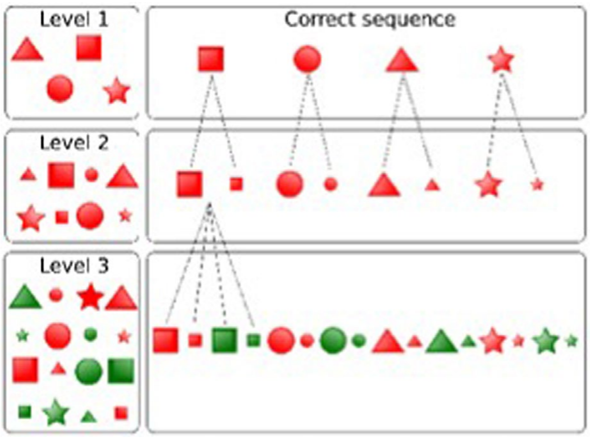

Figure 2 Diagrammatic representation of how the stimulus features were added in to the stimulus set at each level of the Fixed Search task (FxS).

completions were necessary to gain a full rainbow and allow progression onto the next level.

General procedure

Each participant was tested individually in a quiet and familiar setting of the school. The 'Rainbow Collector' game simply began by displaying the instructions: 'Touch every shape once and as fast as you can. Complete the puzzle twice in a row to collect a rainbow'. This was illustrated on the screen for as long as the participants needed to study them and the instructions were repeated by the experimenter. When the participant was comfortable with the rules, they initiated the game by pressing a start button.

When the game reached the Fixed Search task, the participants were informed that this time 'There are only four shapes, but they go in a special order. You must try to find this order to win (the rainbow)'. The instruction was repeated for the two subsequent levels (FxS2 and FxS3). The total game duration was contingent on the participant's performance; the average was between 15 and 25 min.

Ethical approval

The selection and recruitment of participants as well as the proposed task and procedures were formally approved by the standing Ethics Committee at the Department of Psychology. In addition, approval was granted by head teachers and relevant staff at the schools, who then obtained signed letters of informed consent from parents. The staff did not judge written consent from the children to be necessary, but all children were asked if they wished to play the game before testing commenced.

Results

NVIQ

The mean Standard Score for the ASD group was 90.25 (SD = 19.0), while the mean for the TD group was 102.88 (SD = 10.29). Both groups had normally distributed scores (above threshold for $p = .01$ in both cases) on a Shapiro-Wilk test: ASD $W(19.0) = .91$; Controls $W(10.29) = .93$ and were significantly different on a t test for independent samples $t(30) = 2.34$, $p < .05$. — a large effect (Cohen's $d = .82$) The ASD group could be split into 8 participants with a NVIQ at or below the 'normal' threshold of 85 ($M = 76$, SD = 6.87) and 8 participants above the threshold ($M = 104.5$, SD = 16.2). (All control children were at or above the cut-off.) The influence of nonverbal IQ on performance was considered along with the main analyses.

Free search

The number of children reaching criterion in FS1 was 12 out of 16 for the ASD group and 16 out of 16 for the TD group. Of the four of the ASD children who failed, two went on to attempt FS2 but failed to meet criterion and a further three of the previously successful children in this group also failed to pass on FS2. Thus a total of only nine of 14 ASD children passed FS2. Of the seven who failed, four were in the lower NVIQ group (the remaining three had Standard Scores ranging from 86 to 100). All control children passed FS2.

Table 1 shows the number of attempts by all children as well as by those who failed, together with their mean percentage of correct trials for both FS1 and FS2. Mean percentage correct scores are depicted in Fig. 3 for all children ranked by success in order to illustrate the group differences across all individuals for FS1 (but not FS2).

Normality assumptions were violated (at $p = .01$) for scores based on number of attempts in FS1 for both ASD and control groups: $W(23.94) = .83$; $W(9.28) = .77$ respectively. They were also violated (at $p = .05$) for percentage correct for control participants. For this reason and also because of the relatively small sample sizes, non-parametric tests were used to investigate group effects on both these measures. A Mann-Whitney U test indicated a significant effect at $p < .05$ of group on the number attempts by participants to complete FS1: ($U = 73.5$) — a medium effect ($d = .78$), but not FS2 ($U = 107.5$). This difference between groups was present also for percentage correct

Table 1 Performance scores as a function of group and success/failure on the Free Search tasks (FS1 and FS2).

	FS1				FS2			
	No. attempts		% correct trials		No. attempts		% correct trials	
	ASD	Cntrls	ASD	Cntrls	ASD	Cntrls	ASD	Cntrls
Successful	16.7	11.12	27.7	36.8	7.7	8.6	61	38.1
Failed	49.2		5.8		9.2		4.5	
Overall mean	24.81	11.12	20.9	36.8	8.3	8.6	39.3	38.1

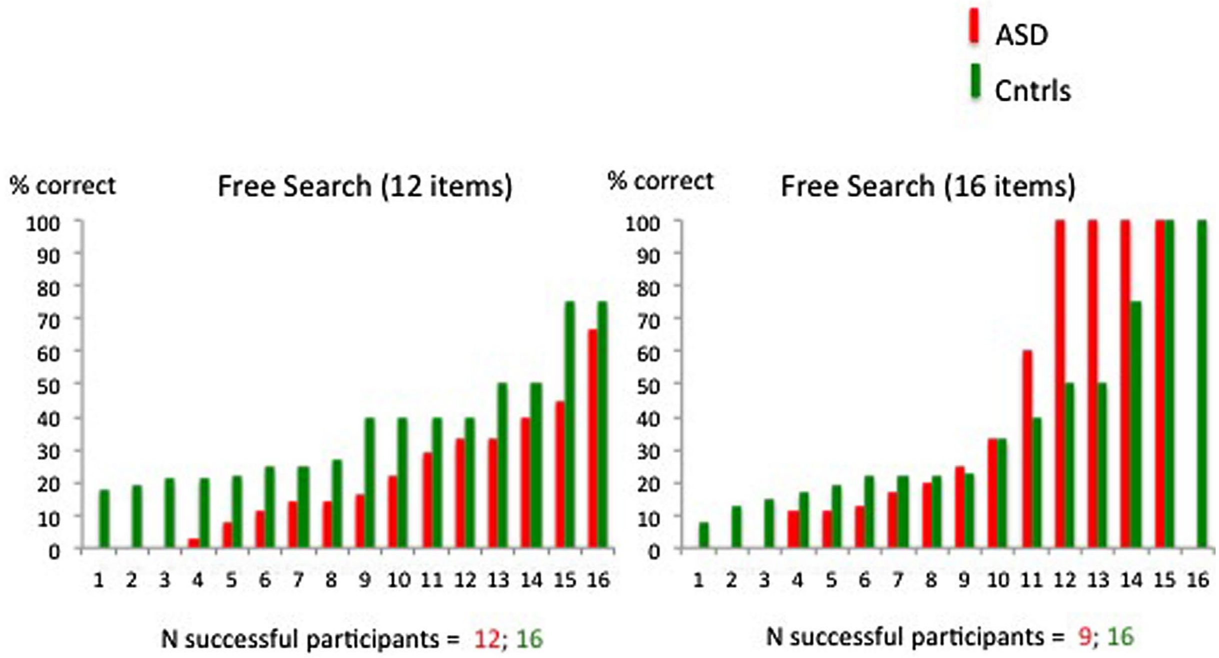

Figure 3 Percentage correct by each individual on both levels of the Free Search task in rank order of success.

in FS1 ($U = 64.5$) — a large effect ($d = .93$), but not FS2 ($U = 103.5$).

These group differences were found to be related to the NVIQ scores for the ASD group. A Spearman's test indicated a significant correlation between nonverbal IQ and percentage correct for the ASD group in FS1 ($r = .64, p < .01$), but not FS2 ($r = .25, p > .05$), but not for the control group for either level ($r = .08, p = .78$; $r = .11, p > .05$). There was also a significant inverse correlation between NVIQ and number of attempts for ASD in FS1 ($r = -.67, p < .01$), but not for FS2 ($r = -.14, p > .05$), but not for the control group for either level ($r = -.13, p = .63$; $r = -.05, p = .84$). Age was not correlated with percentage correct for either group or task (all p values > .1).

The effect of NVIQ on scores on FS1 can be illustrated in terms of the subset falling above and below the normal threshold of 85 as Fig. 4 shows. Mann-Whitney U tests indicated no significant difference between the two ASD sub-groups ($U = 18$; $p > .05$), but did show a significant difference between the lower NVIQ subgroup (only) and the control group ($U = 21.5, p > .05$), the latter showing a large effect size ($d = 2.01$).

The order in which both ASD and TD individuals selected the stimuli was examined for the classification strategy used by both groups. 76.9% of ASD and 83.3% of TD individuals used shape as the primary classification criterion; 15.4% ASD participants and 16.6% TD children categorized the set using color. 7.7% of the ASD population did not commit to any one technique. Whilst the selection of the secondary and tertiary basis for ordering the color (or shape) categories was usually consistent within a trial there was little evidence that success was predicated on maintaining a particular fixed sequence across trials.

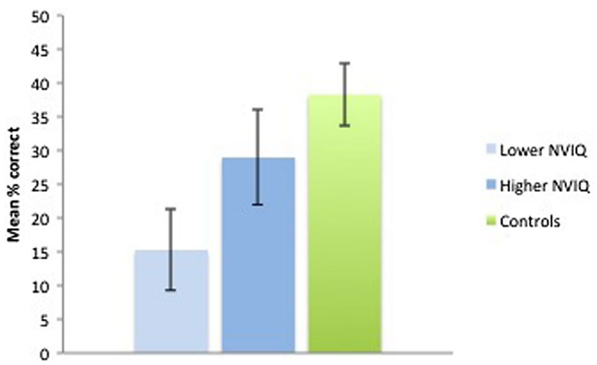

Figure 4 Percentage correct on FS1 (and standard error) for the two ASD sub-groups compared with controls.

Summary and further analyses

The results from the free sequencing tasks show impairment in the ASD group. The correlations between performance and NVIQ and the fact that only the lower NVIQ subgroup were significantly different from controls on percentage correct shows that this subgroup is contributing in the main (though not exclusively) to the deficit shown by ASD children on this task. It is therefore important to know if this is simply due to an inability to understand the task. To examine this in greater detail, learning assessments were carried out using as a measure the actual number of correct touches made on each trial as the game progressed. Once again, the data are divided into the lower and higher scoring NVIQ subgroups and were plotted in terms of number of successful touches on each trial on the two levels of the task. Figure 5 shows the learning

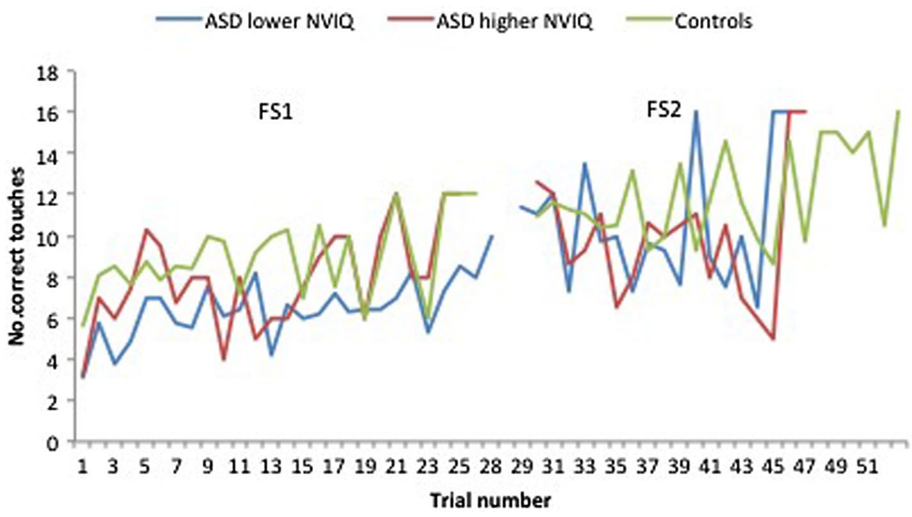

Figure 5 Learning profiles for participants as a function of sub-group during FS1 and FS2, showing the mean number of successful touches made on a given trial before an error was committed.

Table 2 Performance scores as a function of level and group on fixed search tasks (FxS1, Fxs2 and FxS3)

	FxS1		FxS2		FxS3	
	No. attmpts	% correct	No. attmpts	% correct	No. attmpts	% correct
Controls	15.4	36	7.9	41.7	7.8	46.6
ASD	12.6	31.2	7.6	46.9	5.9	50.5

curves from which there is no suggestion that lower NVIQ children were different from the other children in terms of game progression.

Showing evidence of learning in all three groups, statistical inferences are hard to draw from this data because of variability across participants in terms of total number of attempts and the fact that one high NVIQ participant and two low NVIQ participants did not attempt FS2. To avoid this problem, whilst utilizing the extra sensitivity offered by the correct touch data, the numbers of successful touches made on the first three trials only of the first (FS1) task were combined because all participants completed at least three trials on this task. This also allowed a specific focus on the initial response to this self-organizing task before there was any opportunity to adopt a routinized strategy. The average numbers of successful touches (and SDs) across these first trials were: 12.6 (6.2), 16.3 (8.7) and 22.2 (6.8) for lower NVIQ, higher NVIQ and controls respectively. These data were normally distributed (at $p < .05$) and one-tailed paired comparisons were made across sub-groups. The higher and lower NVIQ ASD sub-group were not significantly different from one another on this measure; $t(14) = .95$, $p > .05$, but both the lower and the higher NVIQ subgroups were significantly different from the control children on this measure: $t(22) = 3.34$, $p < .01$ and $t(22) = 1.83$, $p < .05$, respectively. The effect sizes for these results were large ($d = 1.5$) and medium ($d = .68$).

Free search summary

Children with ASD show an impairment relative to controls in self-organized sequencing, and show less task participation as the informational load increases. Sub-group analyses suggest that the lower NVIQ children contributed more strongly to this effect than the children in the normal range, but their learning profiles did not indicate an inability to understand the task.

Fixed search

All children in both groups achieved learning criterion on all three levels of the task. Table 2 shows mean number of attempts on each level as well as percentage correct.

The data of particular interest in this task are where the stimulus properties are nested (levels 2 and 3) and especially FxS 3 (16) items as this level represented the same three varying stimulus properties as the free search task; namely shape color and size. Percentage correct for FxS levels 2 and 3 (8 and 16 stimuli) are depicted in Fig. 6.

Once again normality assumptions were violated for number of attempts by ASD and also controls (at $p = .05$ and $p = .01$): $W(3.29) = .88$ and $W(6.1) = .79$ respectively, as well as for percentage correct for controls (at $p = .01$): $W(21.5) = .81$. Non-parametric tests were therefore used to investigate group differences as for free search.

A Mann-Whitney U test indicated no significant difference between the ASD participants and TD controls using number of attempted trials as a measure across FxS1 ($U = 104.5$, $p = .3$), FxS2 ($U = 114.5$, $p = .6$), FxS3 ($U = 117.5$, $p = .7$) nor on percentage of trials correct on FxS1 ($U = 114.5$, $p = .6$), FxS2 ($U = 101$, $p = .3$), or FxS3 ($U = 123$, $p = .8$). The maximum possible on all tasks was 100% and as only three children achieved this across all levels, this is cannot be sourced to a ceiling effect.

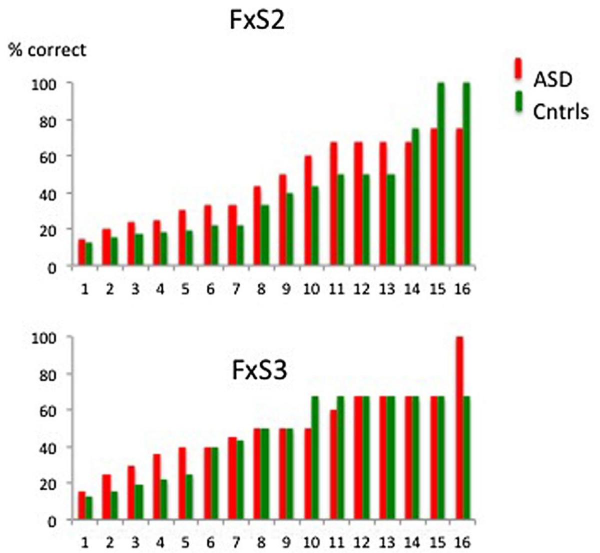

Figure 6 Percentage correct by each individual on levels 2 and 3 of the Fixed Search task in rank order of success.

A Spearman correlation indicated that NVIQ did not correlate significantly with the percentage correct in any of the fixed search levels for ASD: FxS1 ($r = .297$, $p = .09$); FxS2 ($r = 0.68$, $p = .7$); FxS3 ($r = .224$, $p = .22$), or for TD: FxS1 ($r = .06$, $p = .82$); FxS2 ($r = -0.05$, $p = .86$); FxS3 ($r = .04$, $p = .88$).

Fixed search summary

The results for the fixed search task clearly contrast with those for the free sequencing task and there was no basis or rationale for exploring differences in learning or carrying out sub-group analyses as the groups were highly similar on the main parameters of performance.

Discussion

Children with ASD form a highly heterogeneous population. Even children with ASD in mainstream or in special schools with a supposedly ‘higher functioning’ intake will have a nonverbal IQ extending into the moderate to below normal range. Such children are often excluded from research on cognitive functioning in autism because they do not meet the criteria for matching with control participants. Yet this selectivity of sampling could cause executive difficulties in autism to be under-estimated as IQ tests themselves make various demands on task comprehension and working memory, which, if factored out by matching, could dilute the effects of task variables. Indeed lower NVIQ could arise precisely because children with ASD can experience difficulties with certain executive tasks and in that regard are potentially valuable participants in studies of executive skills.

In this study, we included all the children from a special school that were matched to the neurotypical sample on age alone, resulting in half the sample having a NVIQ below the cut-off for normal. As a group, the children with ASD displayed high levels of difficulty during self-organized sequencing of complex objects. Expressed simply in pass/fail terms, the participants with ASD were significantly less likely than age matched controls to discover an organized route by means of which complex shapes could be touched in a non-reiterative order. The number of attempts at the first level and the percentage correct on each trial also reflected a group difference, and learning showed slower progress in terms of number of correct touches per attempt. NVIQ was correlated with basic performance measures, and the sub-group with the lower NVIQ appeared to be the main contributors to the effect. However, children in the higher NVIQ sub-group were intermediate between these children and the control children on all measures. By contrast, all ASD participants performed as well as control children on a second fixed sequencing task using similar stimuli, stepping through task levels involving 4, then 8, then 16 stimuli.

It is important to know whether our inclusive sampling has helped reveal a latent problem with self-organization in children with ASD or whether it simply introduced unprincipled performance impairment (arising from generally lowered intelligence) and was cognitively ‘different’ from that shown by their higher NVIQ peers. The results from our study suggest the former for several reasons. First, the subset of higher NVIQ children was intermediate between control children on all measures. Their significantly lower scores on the first three trials (compared to controls) suggests that they, too, were challenged by the information processing demands of FS1, but were quicker to find a successful strategy. Second, the lower NVIQ subgroup showed that they were on a similar course to both the higher NVIQ children and the control children as the game progressed, even though they were less likely to pass. Finally, but significantly, the impairment was selective to the free search condition only; all group differences disappeared using equally complex stimuli even though the (fixed) task structure was entirely new[2].

In short the study showed evidence of difficulty in free (as opposed to fixed) sequencing that was specific to the ASD group even though it varied in severity. There is no suggestion, furthermore, that any of the ASD group simply failed to engage with the task. All were willing participants, and as the learning data show, were prepared to make considerable efforts to solve it, despite suffering sometimes from what Barendse *et al.* (2013) cite as the ‘overwhelmed feeling’ arising from the strain of complex processing. Indeed those who attempted the harder FS2 level (beyond a few trials) went on to converge with controls as if they had managed to establish a successful ‘rule of thumb’ by this stage, thereby diminishing the group difference found initially. It is also consistent with other reports that spontaneous memorizing strategies can emerge during learning even with children with moderately impaired IQ (Bebko and Ricciuti 2000), and indicates that a game context can be an effective way to encourage these.

During the fixed search task, all participants engaged with the sequential rule at every level, and there were no group differences on any measure. Here it must be stressed that the information-processing load on the final level was exactly equivalent to that in FS2. In the transition from level 2 to level 3 in the fixed task, furthermore, sequential links had to be actively inhibited in order to incorporate the new dimension of color. In the two tasks we presented, therefore, we seem to have captured not only a decided and specific impairment in children with ASD with regard to self-organized search only, but also a 'switch' to an unimpaired ability in the group as a whole during the fixed sequence learning task. We now ask how this finding fits with the symptomatology and theories of EF associated with autism.

A crucial aspect of the free search task is that it needs to be based on at least some element of looking ahead to the total set in order to be sure to explicitly monitor shape, color and size in the selection. Too local or too cursory a consideration will mean that the feature order in working memory will be underspecified and thus subject to unprincipled variation within and across trials. An actual example from a participant in the ASD group on the first trial is as follows: large yellow spiral; large blue spiral; small blue diamond; large yellow rectangle; large blue rectangle; small yellow diamond; large yellow rectangle (then exit following an error after 7 touches); whilst from a control participant (who exited after 9) we have: large yellow rectangle; large yellow diamond; large yellow spiral, small yellow rectangle, large blue rectangle; small blue rectangle; small blue diamond; large blue spiral; small blue rectangle. Even in this random comparison, it can be seen that the more effective response involves keeping two features constant (consistently the same or consistently changing) whilst varying the third.

With regard to traditional concepts of EF impairment in ASD, therefore, we can eliminate the some of the broadly defined factors as causal to the results. We can tell from the fixed search tasks that it is neither the monitoring nor the encoding of all three features per se that is the problem. To solve this second task, in fact, the precise feature order has actually to be encoded in WM and into long-term memory (as well), as the very same rule applies across trials within a level. Thus WM in terms of storage difficulties would not be a candidate explanation. As the rule changes across levels, WM in terms of 'updating' was apparently not a problem either. Likewise, response flexibility in terms of set-shifting and inhibiting previous responses does not in itself to be an issue in the fixed search task, as inter-stimulus connections have to be disrupted at each level to incorporate new ones.

As for top-down control, it would seem that if this is provided by the trained structure of the task, then participants with ASD are even able to follow a strict hierarchical rule where shape is consistently maintained as the broadest (top) level of categorization, followed by size and then color. Where the deficiency arises in children with ASD is in generating a suitable organizing principle for themselves. In terms of a core deficit, this points back to WM but in a very precise sense. As suggested elsewhere (McGonigle-Chalmers *et al.* 2008), it is not the retrospective contents of WM that seem to be affected, but rather but rather on-line 'prospective' working memory. In our task, it is to think ahead for oneself as to how to incorporate all the object properties into a plan — even if it is only for a given trial. The more information that has to be considered, the greater the difficulty.[3] This conclusion also concurs with other evidence from studies of top-down self-organization in ASD such as the Towers tasks and the self-ordered pointing task where, as here, finding a deficit can depend on the amount of information that has to be kept in memory whether in terms of moves or number of items (Joseph *et al.* 2005; Robbins and Sahakian 1994).

There are two new implications that arise from our study on self-orgnisation. The first is with regard to how it might impact on the way individuals with ASD cope with everyday cognitive demands. Although presented as a game, our free search task highlights the real-world need to impose an organized structure on things and activities in everyday life. Here it is relevant to remember that individuals with ASD are characterized as likely to adopt fixed routines and those with Asperger's Syndrome, in particular, to enjoy memorizing by rote. By itself, this could be used to argue that the relative strength on our fixed task simply reflected the natural predilections or 'cognitive 'style' symptomatic of the ASD group (Happé and Frith 2006). However, this was not expressed as a group difference with regard to the TD controls in contrast with the significant difficulty found at a group level in the free search task. In the light of this, it is surely more likely that fixed and routinized solutions that become symptomatic of ASD are compensatory in nature. In other words, it may be exactly their weakness on open-ended processing that makes individuals with ASD seek solutions in everyday life that are less costly in cognitive effort. This is an intellectual disability insofar as there are very few real world situations that favor the fixed over the flexible, with the exceptions of shopping lists and other such material. Real-world environments are commonly populated by complex objects that sometimes need to be sought for in a principled way, such as in a supermarket. Here, the organizing search principle requires exactly the same sort of multi-dimensional nesting of features that we used in our study, and it is notable that training regimes for such situations emphasize the importance of rehearsing sequences (Mechling 2004). We would thus add to this the importance of encouraging individuals to construct sequences for themselves.

A second main implication from our study is that free sequencing difficulties we observed were likely to have been milder and more 'subtle' had we selected children exclusively from the normal to above average NVIQ range. The practical implications of this are important, as all too often autism is viewed as developmental disorder that is sometimes associated with a (separate) 'learning disability' (Hill 2004). Not only can this impede our understanding of how the clinical and intellectual characteristics are interwoven, it also tends to relegate the lower IQ children to beyond the parameters of some of the research. This study suggests a continuity rather than discontinuity in terms of cognitive profiling across the spectrum from which much could be learned. But the more practical implication arising from it is that children at the lower end of the NVIQ distribution may need particular help and training in real world situations requiring top-down organization.

Whilst pointing the way towards the need for more research on self-organization skills in individuals with ASD, enlarging the sample size in a future study would permit more rigorous comparisons across low and high performing sub-groups. As for verbal IQ, this is a measure is as likely to confound than clarify the findings, for reasons to do with heterogeneity of language onset within ASD populations (Alderson-Day 2011). There is also a lack of any evidence that verbal IQ impacts directly on executive performance (Joseph *et al.* 2005). But the role of language is nevertheless an important area of research pointing to a possible role of verbal mediation or inner speech in nonverbal executive tasks (Whitehouse *et al.* 2006). This possibility that has been suggested by Joseph *et al.* (2005) in relation to finding impaired performance by children with autism on a SOPT task involving nameable stimuli. This could be explicitly tested using SOPT or a task such as our own, using dedicated methods such as articulatory suppression (Wallace *et al.* 2009).

In conclusion, using an inclusive sampling strategy helped us to identify a particular working memory problem relating to top-down self-organized planning by children with ASD. The selective nature of this difficulty was made particularly manifest because of the marked contrast between free and fixed search scores in children with a NVIQ toward the lower bound of the normal range. It thus argues for more inclusivity of sampling in future research attempting to highlight and remediate executive impairments in autism. Given the evidence of learning even by the lowest scoring participants, it also indicates that game-based sequencing tasks could be an effective training environment in which to prepare children on the autistic spectrum for the uncertainties of the real world.

Acknowledgements

We should like to sincerely thank the pupils and staff at Kaimes Special School and Gracemount High School, Edinburgh and also Dunblane High School, Dunblane for their patience and co-operation. We should also like to acknowledge Karl Monsen for his prototyping and programming of the games used in this study.

Conflict of interest

No potential conflict of interest was reported by the authors.

Notes

1 Children with moderate to severe learning difficulties attended a different school in Central Edinburgh for children with ASD and Pervasive Developmental Disorder (PDD).

2 Positive transfer from the two tasks is no more likely than negative transfer as the tasks were similar in terms of stimulus properties but different in terms of how they should be sequenced across the three levels.

3 This could manifest itself both in terms of being able to carry out a thorough visual search as well as well as in constructing a mental plan — although these are themselves almost certainly inter-related.

ORCID

Fiona McCrohan http://orcid.org/0000-0002-3607-6168

References

Alderson-Day, B. 2011. Verbal problem-solving in autism spectrum disorders: A problem of plan construction? *Autism Research, 4*, 401–411.

Bardikoff, N. and McGonigle-Chalmers, M. 2014. Testing nonverbal IQ in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*, 1200–1207.

Barendse, E. M., Hendriks, M. P. H., Jansen, J. F. A., Backes, W. H., Hofman, P. A. M., Thoonen, G. … Aldenkamp, A. P. 2013. Working memory deficits in high-functioning adolescents with autism spectrum disorders: neuropsychological and neuroimaging correlates. *Journal of Neurodevelopmental Disorders, 5*, 14–20.

Bebko, J. M. and Ricciuti, C. 2000. Executive functioning and memory strategy use in children with autism. *Autism, 4*, 299–320.

Craig, F., Margari, F., Legrottaglie, A. R., Palumbi, R., de Giambattista, C. and Margari, L. 2016. A review of executive function deficits in autism spectrum disorder and attention-deficit/hyepractivity disorder. *Neuropsychiatric Disease and Treatment, 12*, 1191–1202.

Frith, U. 1970. Studies in pattern detection in normal and autistic children. *Journal of Experimental Child Psychology, 10*, 120–135.

Frith, U. 2003. *Autism: Explaining the enigma*. Malden, MA: Blackwell.

Geurts, H. M., Verte, S., Oosterlaan, J., Roeyers, H. and Sergeant, J. A. 2004. How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal of Child Psychology and Psychiatry, 45*, 836–854.

Goldberg, M. C., Mostofsky, S. H., Cutting, L. E., Mahone, E. M., Astor, B. C. and Denckla, M. B. 2005. Subtle Executive impairment in children with autism and children with ADHD. *Journal of Autism and Developmental Disorders, 35*, 279–293.

Greenaway, R. and Plaisted, K. 2005. Top-down attentional modulation in autistic spectrum disorders is stimulus-specific. *Psychological Science, 16*, 987–994.

Happé, F. 1994. An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*, 129–154.

Happé, F. and Frith, U. 2006. The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*, 5–25.

Hermelin, B. and O'Connor, N. 1970. *Psychological experiments with autistic children*. London: Pergamon Press.

Hill, E. L. 2004. Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*, 189–233.

Horlin, C., Black, M., Falkmer, M. and Falkmer, T. 2016. Proficiency of individuals with autism spectrum disorder at disembedding figures: A systematic review. *Developmental Neurorehabilitation, 19*, 54–63.

Hughes, C., Russell, J. and Robbins, T. W. 1994. Evidence for executive dysfunction in autism. *Neuropsychologia, 32*, 477–492.

Joseph, R. M., Steele, S. D., Meyer, E. and Tagerflusberg, H. 2005. Self-ordered pointing in children with autism: failure to use verbal mediation in the service of working memory? *Neuropsychologia, 43*, 1400–1411.

Kaland, N., Smith, L. and Mortensen, E. L. 2008. Brief report: Cognitive Flexibility and focused attention in children and adolescents with asperger syndrome or high-functioning autism as measured on the computerized version of the wisconsin card sorting test. *Journal of Autism and Developmental Disorders, 38*, 1161–1165.

Kaufman, A. S. and Kaufman, N. L. 2004. *Kaufman assessment battery for children: Second edition (KABC-II)*. Circle Pines, MN: AGS Publishing.

Keehn, B. and Joseph, R. M. 2016. Exploring what's missing: What Do Target Absent Trials Reveal About Autism Search Superiority? *Journal of Autism and Developmental Disorders, 46*, 1686–1698.

Kercood, S., Grskovic, J. A., Banda, D. and Begeske, J. 2014. Working memory and autism: A review of literature. *Research in Autism Spectrum Disorders, 8*, 1316–1332.

Lord, C., Rutter, M. and Le Couteur, A. 1994. Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Mari, M., Castiello, U., Marks, D., Marraffa, C. and Prior, M. 2003. The reach-to-grasp movement in children with autism spectrum disorder. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 393–403.

McGonigle-Chalmers, M. 2015. *Understanding cognitive development*. London: SAGE publications, pp.315–320

McGonigle-Chalmers, M. and Alderson-Day, B. 2010. Free classification as a window on executive functioning in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 40*, 844–857.

McGonigle-Chalmers, M., Bodner, K., Fox-Pitt, A. and Nicholson, L. 2008. Size Sequencing as a window on executive control in children with autism and asperger's syndrome. *Journal of Autism and Developmental Disorders, 38*, 1382–1390.

McGonigle-Chalmers, M. and McSweeney, M. 2013. The role of timing in testing nonverbal IQ in children with ASD. *Journal of Autism and Developmental Disorders, 43*, 80–90.

Mechling, L. C. 2004. Effects of multi-media, computer-based instruction on grocery shopping fluency. *Journal of Special Education Technology, 19*, 23–34.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. and Wager, T. D. 2000. The unity and diversity of executive functions and their contributions to complex 'Frontal Lobe' tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

O'Hearn, K., Asato, M., Ordaz, S. and Luna, B. 2008. Neurodevelopment and executive function in autism. *Development and Psychopathology, 20*, 1103–1132.

Ozonoff, S. 1995. Reliability and validity of the Wisconsin card sorting test in studies of autism. *Neuropsychology, 9*, 491–500.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R. M., Klin, A. and Wrathall, D. 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 network. *Journal of Autism and Developmental Disorders, 34*, 139–150.

Pascualvaca, D. M., Fantie, B. D., Papageorgiou, M. and Mirsky, A. F. 1998. Attentional capacities in children with autism: Is there a general deficit in shifting focus? *Journal of Autism and Developmental Disorders, 28*, 467–478.

Pellicano, E., Smith, A. D., Cristino, F., Hoode, B. M., Briscoe, J. and Gilchrist, I. D. 2010. Children with autism are neither systematic nor optimal foragers. *PNAS, 108*, 421–426.

Pennington, B. F. and Ozonoff, S. 1996. Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51–87.

Petrides, M. and Milner, B. 1982. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*, 249–262.

Rajendran, G. and Mitchell, P. 2007. Cognitive theories of autism. *Developmental Review, 27*, 224–260.

Robbins, T. W. 1997. Integrating the neurobiological and neuro-psychological dimensions of autism. In: J. Russell, ed. *Autism as an executive disorder*. Oxford: Oxford University Press.

Robbins, T. W. and Sahakian, B. J. 1994. Computer methods of assessment of cognitive function. In J. R. M. Copeland, M. T. Abou-Saleh and D. G. Blazers, eds. *Principles and practice of geriatric psychiatry*. Chichester: John Wiley & Sons Ltd, pp.205–209.

Robinson, S., Goddard, L., Dritschel, B., Wisley, M. and Howlin, P. 2009. Executive functions in children with autism spectrum disorders. *Brain and Cognition, 71*, 362–368.

Roelofs, R. L., Visser, E. M., Berger, H. J. C., Prins, J. B., Van Schrojenstein Lantman-De Valk, H. M. J., and Teunisse, J. P. 2015. Executive functioning in individuals with intellectual disabilities and autism spectrum disorders. *Journal of Intellectual Disability Research, 59*, 125–137.

Rommelse, N., Langerak, I., van der Meer, J., de Bruijn, Y., Staal, W., Oerlemans, A. and Buitelaar, J. 2015. Intelligence may moderate the cognitive profile of patients with ASD. *PLOS ONE, 10*, e0138698.

Rumsey, J. M. and Hamburger, S. D. 1988. Neuropsychological findings in high-functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology, 10*, 201–221.

Russell, J. 1997. *Autism as an executive disorder*. New York, NY: Oxford University Press.

Russell, J., Mauthner, N., Sharpe, S. and Tidswell, T. 1991. The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331–349.

Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A. and Butcher, B. 2010. Executive functioning in children with asperger syndrome, ADHD-combined type, ADHD-predominately inattentive type, and controls. *Journal of Autism and Developmental Disorders, 40*, 1017–1027.

Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. and Lehmkuhl, G. 2008. Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child and Adolescent Psychiatry and Mental Health, 2*, 4.

Snyder, H. R., Miyake, A. and Hankin, B. L. 2015. Advancing understanding of executive function impairments and psychopathology: Bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology, 6*, 1–24.

Verté, S., Geurts, H. M., Roeyers, H., Oosterlaan, J. and Sergeant, J. A. 2006. Executive functioning in children with an autism spectrum disorder: Can we differentiate within the spectrum? *Journal of Autism and Developmental Disorders, 36*, 351–372.

Wallace, G. L., Silvers, J. A., Martin, A. and Kenworthy, L. E. 2009. Brief report: Further evidence for inner speech deficits in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 1735–1739.

Whitehouse, A. J., Maybery, M. T. and Durkin, K. 2006. Inner speech impairments in autism. *Journal of Child Psychology and Psychiatry., 47*, 857–865.