Short-Term Memory in Autism Spectrum Disorder

Marie Poirier

Jonathan S. Martin, Sebastian B. Gaigg, and Dermot M. Bowler

City University London

Running head: VERBAL SHORT-TERM MEMORY IN ASD

Date: 04 Oct. 2010
Abstract

Three experiments examined verbal short-term memory in comparison and Autism Spectrum Disorder (ASD) participants. Experiment 1 involved forward and backward digit recall. Experiment 2 used a standard immediate serial recall task where, contrary to the digit-span task, items (words) were not repeated from list to list. Hence, this task called more heavily on item memory. Experiment 3 tested short-term order memory with an order recognition test: each word list was repeated with or without the position of two adjacent items swapped. The ASD group showed poorer performance in all three experiments. Experiments 1 and 2 showed that group differences were due to memory for the order of the items, not to memory for the items themselves. Confirming these findings, the results of Experiment 3 showed that the ASD group had more difficulty detecting a change in the temporal sequence of the items.

KEYWORDS: AUTISM SPECTRUM DISORDER, SHORT-TERM MEMORY
Short-Term Memory in Autism Spectrum Disorder

With respect to memory functioning, research on Autism Spectrum Disorder (ASD) has produced a complex pattern (Mayes & Boucher, 2008; Russell, Jarrold, & Henry, 1996). For example, the ASD literature suggests episodic memory difficulties in the face of preserved memory for facts (Bowler & Gaigg, 2008), impairments in the use of organisation strategies to support recall (Minshew & Goldstein, 2001), and preserved rote memory (Hermelin & O’Connor, 1970). It is also traditionally thought that short-term/immediate memory is preserved in ASD – or proportional to general cognitive ability (Boucher, 2001; Bennetto, Pennington, & Rogers, 1996). However, a number of considerations suggest that this conclusion may be inappropriate.

First, a close examination of key studies reveals methodological shortcomings that invite a re-assessment. For example, some of the most influential and oft-cited papers in the field concluded that verbal memory span is intact in children with ASD (O’Connor & Hermelin, 1965; Hermelin & O’Connor, 1967, 1970). These studies, however, compared samples that were equated on digit span before immediate memory for words was tested. Currently, all models of STM assume that digit span and immediate verbal recall rely on the same mechanisms (Baddeley, 2000; Brown, Neath, & Chater, 2007). If ASD and comparison samples are equated on digit span and memory span for words is tested, a null effect appears as the only likely possibility. Below, we examined STM in ASD with groups equated on general cognitive ability – but they were not matched on digit span.

Consider also the hypothesis that suggests that ASD entails selective memory deficits, and in particular, problems with memory for when events occurred [or with temporal-contextual information] (Boucher, 2001). For example, when asked what activities they had carried out during a given period, children with autism performed poorly, and this was not
because they had forgotten the activities themselves (Boucher, 2001). To the extent that STM tasks involve remembering stimuli in relation to their time of appearance, this view predicts an impairment in ASD. Typically, assessments of immediate memory require the recall of items in their order of presentation. To recall items in order, their temporal or ordinal relationships must be retrieved. In effect, current computational models of STM often rely on time-based context information to predict serial recall performance (Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999).

A final perspective suggesting STM difficulties in ASD is offered by Gaigg, Gardiner & Bowler (2008). They propose that the pattern of memory functioning in ASD results from problems with processing the relations amongst elements of experience—whilst memory for the elements themselves is preserved. As with Boucher (2001), this account predicts that individuals with ASD will struggle to encode the temporal/positional relationships in STM tasks despite memory for the items themselves.

When these views are considered along with the methodological concerns mentioned above, the case for re-visiting STM in autism seems compelling. Verbal STM is thought to be important in daily activity and associated with language development; it is generally thought to contribute to higher-order cognitive functions (Baddeley, 2000; Jarrold, Thorn, & Stephen, 2009). Uncovering a STM deficit in ASD would help us to understand some of the characteristic behaviours of ASD, potentially shedding some light on how best to communicate information and suggest some remedial action (we return to this in the discussion).

In this paper, we investigate STM in individuals with ASD with normal levels of intelligence and typical language development. Hence, we can rule out the possibility that any STM disadvantage uncovered is due to general cognitive difficulties. Testing such high-
functioning individuals constitutes a stringent test of the predicted STM difficulty because, usually, normal IQ and language development is not associated with verbal STM problems.

Method: General

All experiments involved adults with ASD, diagnosed by local health authorities and/or experienced clinicians in accordance with DSM-IV-TR criteria (American Psychiatric Association, 2000). For 31 of the total 43 ASD participants (2 participated in all three and 9 in two of the experiments), diagnosis was confirmed through administration of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989). The detailed clinical records of a further 7 individuals left no doubt as to the accuracy of their diagnosis and although no detailed records were available for the remaining 5 individuals, they were recruited through specialist group homes or support groups that cater specifically for those with ASD. In the expert opinion of the fourth author (DB), these individuals exhibited clear behavioural manifestations consistent with a diagnosis of ASD. Since the results reported below did not change if these individuals were excluded, they were retained in the final sample. The comparison group was recruited via local newspaper advertisements, and brief interviews ensured that no participant had a history of neurological or psychiatric illness. Individuals gave their informed consent and were paid standard University fees (£7/hr) for their participation.

The ASD and comparison participants were group matched on age and prorated verbal IQ – we used a verbal IQ score that excluded the scores on digit span (WAIS-R<sup>uk</sup> or the WAIS-III<sup>uk</sup>). Therefore, we avoided equating samples on the memory process we were studying. The groups never differed on full scale or performance IQ. Participants were
individually tested. Finally, in all cases where this was relevant, we used the prorated verbal IQ as a covariate. The ANCOVAs in all three experiments called upon this covariate.

Experiment 1

In this study, participants recalled digit sequences; two versions of the task were used: forward recall and backward recall. We chose to test both as this is typically done in standard cognitive assessments. A span test usually involves interrupting testing after two or three lists have been recalled incorrectly; also, scoring is global: each list is given a score of 0 or 1, irrespective of the number of errors. Here, to ensure sensitivity or our measures, we used 16 trials per direction and did not use global scoring.

Method

Participants

Sixteen ASD participants (12 men, 4 women) and 16 comparison individuals (11 men, 5 women) were recruited for this experiment. Table 1 (first two columns) provides descriptive statistics for both groups.

Insert Table 1 about here

Materials

This study involved visually presented digits (between 1 and 9). Presentation was computer controlled. Random sequences of digits were presented in Arial 72 font, at a rate of one per second, in the middle of the screen.

For forward trials, individuals were instructed to recall the numbers orally, in the same order as they appeared. For backward trials, they were told to verbally recall the numbers ‘back to front’, starting with the last digit and ending with the first. Before the start
of each task, the experimenter made sure the participant understood the instructions. For forward recall, the list-length was seven while for backward recall six digits were presented; task order was counterbalanced across participants. Pre-testing established that these list-lengths equated the difficulty of the tasks without producing floor/ceiling effects. Each participant completed 12 trials per recall direction.

Results & Discussion

We first examined overall correct recall in position, where to be considered correct, digits had to be reported in their original position. Both item errors (omissions, intrusions) and order errors (third digit recalled fourth) contribute to this score. STM research has shown the importance of considering these two dimensions separately; several variables affect item recall without interfering with order memory while others have the reverse effect (Saint-Aubin & Poirier, 1999; Majerus, 2008). Moreover, the theoretical perspectives reviewed above suggest that order information recall may present a particular challenge for the ASD group relative to memory for the items themselves. Therefore, two further scores were examined: 1) correct item recall: this is the proportion of digits correctly recalled irrespective of order; 2) proportion of order errors: this is the number of misplaced digits (order errors) divided by the item recall score. This score provides the proportion of correct items that were not recalled in the correct position. It controls for differences between groups and individuals in item recall and provides a more adequate measure of order memory than a simple error count (Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999).

For overall correct in position, a 2 (forward vs. backward) x 2 (group) mixed ANCOVA showed that only the group factor \( F_{(1,29)}=8.10, \ p=0.008, \ d=0.97^1 \) was significant. Comparison participants had better mean performance (0.81, SD=0.15) than the ASD group.

---

1 Cohen’s \( d \) is a measure of effect size
STM in ASD

(0.64, SD=0.20). There was no effect of recall direction \(F_{(1,29)}<1, p=0.94\), an unsurprising result given backward and forward versions were equated for difficulty by adjusting list-lengths. The interaction was also non-significant \(F_{(1,29)}<1, p=0.61\).

For item recall (correct recall, irrespective of order), there were no significant effects [task \(F_{(1,29)}<1, p=0.67\); group \(F_{(1,29)}=1.14, p=0.29\); interaction \(F_{(1,29)}<1, p=0.79\)]. For the proportion of order errors, there was a clear effect of group \(F_{(1,29)}=11.27, p=0.002, d=1.09\) but again no effect of task \(F_{(1,29)}<1, p=0.73\) and no interaction \(F_{(1,29)}<1, p=0.75\). The ASD group produced more order errors per word recalled (mean=0.28, SD=0.16) than the comparison group (mean=0.13, SD=0.11). This error analysis implies that the group difference in overall performance is attributable to order memory, not memory for the items per se.

Considering that these findings disagree with the received view in the field, a replication of the order memory findings is required. One could argue that Experiment 1 placed little demand on item memory as items were taken from a small and familiar set of digits. The implication is that if performance relied more on item memory, group differences in both item and order recall might be revealed. Experiment 2 addressed this issue.

**Experiment 2**

Experiment 2 required the serial recall of six-word lists and instead of oral recall, written recall was used. Different words appeared on every trial, so that both the items presented and the order in which they appeared had to be remembered, providing a better test of item memory than the task in Experiment 1.

**Method**

**Participants**
Twenty-two ASD participants (16 male, and 6 female) and 22 comparisons (17 male, 5 female) participated in this experiment. Table 1 (middle columns) presents the mean ages and IQ scores for both groups.

Materials

A pool of 288, two/three syllable words were selected from the MRC psycholinguistics database (Coltheart, 1981). They had a Brown (1984) verbal frequency average of 12.7 and a Kuçera and Francis (1967) written frequency average of 94.1 per million. They were randomly sorted into 48, six-word lists; the lists were checked to avoid semantically or otherwise related words. This process was repeated until there were four sets of 48 lists. One set was randomly selected (without replacement) per participant. A program controlled the display of words (Arial 24 font; centred). For each participant, the order of the words within lists and the order of the lists were randomised. A further 12 words were selected to construct two practice trials. Participants provided responses on specially prepared answer sheets.

Procedure

After instructions, two practice trials were provided. The experimenter ensured that the instructions were being followed and 48 experimental trials followed. Participants read words aloud as they appeared, each presented for 1000 msec, followed by a 500 msec pause. After six words, a question mark indicated recall could commence. Words were written down in their order of appearance: participants recalled the first word first, and then moved through the list sequentially. If they were unable to recall a word, the corresponding space was left blank.

Results & Discussion
The same scores as in Experiment 1 were analysed (overall correct in position, item recall irrespective of order, and proportion of order errors). As list-length in this experiment was constant, serial position could be included in the correct-in-position analysis.

For correct-in-position, a mixed-model ANCOVA was run including the between-subjects factor group and the within-subjects factor serial position (1 to 6). There was a main effect of group ($F_{(1,41)}=5.10$, $p=0.03$, $d=0.67$). Average performance was 0.51 (SD=0.19) for the ASD group and 0.63 (SD=0.17) for the comparison group. There was the typical serial position effect ($F_{(5,205)}=5.34$, $p<0.001$) and no interaction ($F_{(5,205)}<1$, $p=0.89$). Figure 1 presents these results.

Concerning the number of items correctly recalled irrespective of order, the groups were not significantly different [group ANCOVA ($F_{(1,41)}=1.81$, $p=0.19$)]. The comparison group remembered 74% (SD=0.13) of the presented words while the ASD group recalled 69% (SD=0.12).

The proportion of order errors was analysed with a mixed-model ANCOVA; there was a significant group effect ($F_{(1,41)} = 7.98$, $p=0.007$, $d=0.81$). The mean proportion of items recalled in an incorrect position was higher for the ASD group (0.27, SD=0.17) than for the comparison participants (0.16, SD=0.09).

These results replicate and extend the findings of Experiment 1 – individuals with a diagnosis of ASD show a reliable and sizeable decrement when it comes to short-term order memory for verbal information.

**Experiment 3**

In Experiment 3, the above findings were further explored through an order recognition test. Six words were presented one at a time and then presented sequentially
again; 50% of the time, the positions of two adjacent items were swapped. The participants’ task is to detect any change in the temporal sequence of the items. If the ASD group has more difficulty remembering temporal order, we would expect change detection performance to suffer.

**Method**

**Participants**

Eighteen individuals with ASD (12 male, and 6 female) and 18 typical individuals (13 male, 5 female) participated. Table 1 (last columns) presents the mean ages and IQ scores for both groups.

**Materials**

A pool of 300, two/three syllable words was selected from the MRC Psycholinguistics Database (Coltheart, 1981). These words had a Brown (1984) verbal frequency average of 10.4, and a Kuçera and Francis (1967) written frequency average of 84 per million. The words were randomly sorted into 50 six-word lists which were reviewed to avoid semantic or other relationships. For each participant, word order within lists and list order were randomised. Two further similar word lists were created for practice trials. A program controlled word display (Arial 24 font, centre screen) for study and test stages.

**Procedure**

Following instructions and any questions about the procedure, participants read out the words as they appeared for 1000 msec with a 500 msec pause. After the sixth word, three black squares (2000 msec) indicated the end of the list. One second later, the six words were presented again in the same manner; 500 msec after the last word, two buttons appeared in the centre of the screen, labelled ‘same’ and ‘different’. Participants clicked on ‘different’ if they thought two words had been swapped; otherwise, they clicked on ‘same’. In 50% of the
trials (n=25) both lists were identical; if not, the positions of two adjacent words were inter-changed. Which lists were different and which positions were swapped was randomly determined per participant; however, each possible swap (2-1, 3-2, 4-3, 5-4, 6-5) was tested 5 times.

Results & Discussion

The proportion of correct responses was calculated for same trials as well as for each change position. The serial/change position data (only available for the change trials) was first examined. A 5 (change positions) x 2 (group) mixed design ANCOVA found no reliable effect of position ($F(4,132) < 1, p=0.66$) or group x position interaction ($F(4,132) < 1, p=0.80$); there was however, a significant group difference ($F(1,33)=6.2, p=0.02, d=0.72$). As there was no effect of position, averages across positions are reported: correct detection scores were 0.68 (SD 0.21) and 0.80 (SD 0.11) for the ASD and comparison groups respectively.

These results might be attributable to a propensity to answer ‘same’ in the ASD group, but the ASD group mean was numerically lower for ‘same’ responses (0.70, SD=0.17) relative to the comparison group (0.76, SD=0.15). An ANCOVA comparing the proportion of ‘same’ answers between groups did not produce a significant effect ($F(1,33)=1.0, p=0.26$). To confirm these findings, following Pastore, Crawley, Berens, and Skelly (2003), the $d'$ (d-prime) and decision bias (propensity to say ‘same’) of each participant was calculated and groups were compared. The mean $d'$ scores were 1.15 (SD=0.91) and 1.70 (SD=0.88) for the ASD and comparison groups, respectively, and their decision bias scores were 0.5 (SD=0.46) and -0.6 (SD=0.24). The between group ANCOVA on the $d'$ scores showed there was a reliable group difference ($F(1,33)=4.33, p=0.04, d=0.61$), confirming that the comparison group showed superior change detection. There was no reliable difference for the decision bias score ($F(1,33)=1.02, p=0.32$).

General Discussion
Until now, when verbal STM in ASD has been discussed, the received view was that people with ASD did not show a deficit or if they did, the difficulty was commensurate with their intellectual abilities. We have argued that this characterisation needs to be reviewed. By testing individuals with normal IQ and language development, we put the hypothesis of a STM difficulty in ASD to a stringent test.

In three experiments, the comparison group performed better than the ASD group. Experiment 1 used forward and backward digit recall and Experiment 2 tested immediate word recall, with different words on every trial. Both studies showed the groups recalled a similar number of items; the difference between groups depended on how well the order of the said items was remembered. In Experiment 3, words were provided at the point of recall—but the position of two of the items could be inter-changed. Detecting these changes proved to be more challenging for ASD participants although groups did not differ on decision bias.

The present results clearly suggest that ASD is characterised by atypical verbal STM—which is produced by order memory difficulties. This is at odds with tests of STM for ASD and comparison groups where no difference is found. We suggest that a number of factors may contribute to these null effects. In most cases, the objective was not to investigate STM so samples were matched on verbal IQ or full scale IQ, including digit span. Also, a Wechsler style span task is typically used where testing is interrupted after a few errors and scoring is list-wise: the same score is given if 1 error or 4 errors are made. This implies considerable loss of data and reduces sensitivity to group effects. The combined effect of a more global measure, a smaller number of trials, and matching procedures that partial out some of the difference of interest is probably sufficient to produce null effects.

On the other hand, the order memory pattern revealed here is consistent with findings reported in relation to long-term memory. Gaigg et al. (2008) used a free recall task
involving items from various categories. They found that relational processing—processing centred on the relations between items, or on the relations between items and the context in which they appear, rather than on item-specific features—was not called upon to the same degree by ASD participants. Bowler and Gaigg (2008) reviewed the evidence on relational processing problems in ASD and concluded that there is considerable support for this idea. They propose that, in ASD, there are difficulties with processing relationships amongst elements of experience whilst memory for the elements themselves is preserved.

The results reported in the current paper suggest that relational processing difficulties may extend to the short-term domain. Morin, Poirier, Fortin and Hulme (2006) discussed item and relational processing in STM, suggesting that the view is useful in explaining many effects observed with immediate serial recall; in this context, item order is considered as a relational characteristic.

The results reported here are also of interest because recent theoretical views of verbal STM insist on its role in maintaining the order of recently presented information (Majerus, 2008; Thorn & Page, 2008; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008). Majerus (2008) proposed that to-be-remembered items are represented in long-term memory networks while STM’s main role is to briefly maintain the temporal relationships between events until further processing or retrieval can take place. In this light, our findings suggest that in ASD, the representation of items in long-term memory is typical – but the function more specific to STM is not.

Although speculative, a further interpretation is that ASD is associated with temporal processing difficulties. Converging lines of evidence have identified brain regions thought to be important in timing, including the frontal cortex, hippocampus, basal ganglia, and cerebellum (Meck, 2005). Notably, disruptions in all four of these structures have been linked to autism. Majerus (2008) and Burgess and Hitch (1999) among others, have suggested that
verbal STM relies on the joint encoding of separate representations – one for the items and one for their temporal context. At the point of retrieval, order recall relies on the encoded temporal information. Assuming the representation of the temporal context is less precise in ASD, one would predict the type of order errors reported here. Moreover, these difficulties would not be unique to STM, an expectation that is coherent with many research findings in the area. Precise time processing is related to many behaviours such as fine motor control, speech, etc (Meck, 2005). Other work has also led to this suggestion; for example, Gepner and Feron (2009) offered a hypothesis accounting for various aspects of ASD behaviour, suggesting that many difficulties could be attributable to temporo-spatial processing disorders. Finally, there is also empirical evidence pointing to time processing difficulties in ASD (e.g. Martin et al, 2010). Further research will be necessary to determine to what degree verbal STM problems and other characteristics of ASD are associated with impaired temporal processing.

What implications do our findings have at a more clinical level? We would offer that more research is required before firm recommendations can be made; however, we can draw on research in other areas. For example, Gathercole and Alloway (2006) review short-term and working memory impairments in neurodevelopmental disorders, including Down’s syndrome, William’s syndrome, specific language impairment, and attention-deficits. They include research-based suggestions for remedial support and conclude by suggesting that: “…to minimise the adverse consequences for learning and educational progress that result from these impairments, early diagnosis followed by remedial support that targets relevant domains of learning is strongly recommended.” (p. 12). In particular, as verbal STM is associated with language development, rehearsal training and vocabulary development support is recommended. The present results suggest that these recommendations should be
extended to children and adults with ASD although again, further research would be highly desirable.

**Conclusion**

We have presented clear evidence that contrary to the received view in the field, individuals with a diagnosis of ASD show a STM difficulty specifically related to memory for the order of the to-be-remembered items. These findings, along with a number of others in the field, suggest relational processing may be atypical in ASD; although speculative, a time processing deficit hypothesis also fits well with our results.
References


Author’s note

This research was supported by grant RES-000-22-1063 from the Economic and Social Research Council to Marie Poirier and Dermot Bowler. Thanks to Ms Pritha Dhir for her invaluable help with scoring response protocols.

Correspondence should be addressed to Marie Poirier, Psychology Department, City University, Northampton Square, London, United Kingdom, EC1V 0HB. E-mail:

M.Poirier@city.ac.uk.
Table 1. Descriptive statistics, ASD and comparison groups: Experiments 1, 2, and 3

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ASD</td>
<td>Comparison</td>
<td>ASD</td>
</tr>
<tr>
<td></td>
<td>(n = 16)</td>
<td>(n = 16)</td>
<td>(n = 22)</td>
</tr>
<tr>
<td>Chronological age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>31.6</td>
<td>34.8</td>
<td>37.6</td>
</tr>
<tr>
<td>SD</td>
<td>11.4</td>
<td>10.8</td>
<td>13.3</td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>100.3</td>
<td>102.4</td>
<td>106.9</td>
</tr>
<tr>
<td>SD</td>
<td>16.2</td>
<td>11.9</td>
<td>18.8</td>
</tr>
<tr>
<td>Range</td>
<td>80-122</td>
<td>88-120</td>
<td>80-138</td>
</tr>
<tr>
<td>Verbal IQ*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>100.6</td>
<td>101.8</td>
<td>109.8</td>
</tr>
<tr>
<td>SD</td>
<td>13.9</td>
<td>12.3</td>
<td>18.3</td>
</tr>
<tr>
<td>Range</td>
<td>81-121</td>
<td>84-122</td>
<td>83-142</td>
</tr>
<tr>
<td>Performance IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>100.6</td>
<td>103.1</td>
<td>108</td>
</tr>
<tr>
<td>SD</td>
<td>17.9</td>
<td>11.3</td>
<td>19.4</td>
</tr>
<tr>
<td>Range</td>
<td>74-129</td>
<td>89-122</td>
<td>86-139</td>
</tr>
</tbody>
</table>

* Prorated Verbal IQ: does not include digit span
Figure captions

*Figure 1.* Experiment 2: Proportion correct recall in position by serial position and group; error bars represent the standard error.