J Autism Dev Disord (2010) 40:878–889
DOI 10.1007/s10803-010-0937-7

ORIGINAL PAPER

# Assessing Recollection and Familiarity in Autistic Spectrum Disorders: Methods and Findings

**Sally Bigham · Jill Boucher · Andrew Mayes · Sophie Anns**

Published online: 21 January 2010


**Abstract** We hypothesise that of the two processes underlying declarative memory, recollection is impaired in high-functioning autism (HFA) whereas recollection and familiarity are impaired in low-functioning autism (LFA). Testing these hypotheses necessitates assessing recollection and familiarity separately. However, this is difficult, because both processes contribute to performance on standard memory tests. Moreover, tests must be suitable for use with young or intellectually disabled participants. This study aimed to develop tests of recollection and familiarity separately, and to make preliminary tests of our hypotheses. We developed a temporal source memory task to assess recollection in LFA, and a shape recognition task to assess familiarity and an action recall task assessing recollection in HFA. The methods and implications of the results are discussed.

**Keywords** Declarative memory · Recollection · Familiarity · Source memory · Recall · Recognition

## Introduction

The study of memory in people with autistic spectrum disorders (ASDs) is important because any atypicalities of memory will affect how and what people with ASDs learn, and this will be evident in the course and outcomes of behavioural and brain development, and in ways in which individuals with ASDs experience and respond to the external world.

Memory in able individuals with ASDs, whether diagnosed with Asperger syndrome (AS) or with high-functioning autism (HFA) defined in terms of the triad of impairments plus language and cognitive abilities within normal limits, has been quite extensively studied over the last two decades (for reviews see Boucher and Bowler 2008; Boucher and Mayes 2010). Behavioural findings on memory in HFA[1] have been variously interpreted. Most commonly, Tulving's (1985) taxonomy of memory systems is used as the basis of interpretation, in particular his division of explicit, self-aware memory—'declarative memory'—into memory for unique personally experienced events as opposed to memory for factual, decontextualized information including word meanings. According to Tulving's definitions, the processes subserving memory for personally experienced events constitute the episodic memory system, whereas the processes subserving memory for factual information constitute the semantic memory system. Utilising this distinction, evidence on declarative memory in HFA has frequently been interpreted as reflecting a combination of impaired episodic memory with spared semantic memory (e.g., Bowler et al. 2000; Ben-Shalom 2003; Toichi and Kamio (2003); Salmond et al. 2005; Bowler and Gaigg 2008). Bowler and Gaigg (2008; Gaigg et al. 2008; see also BenShalom 2003) propose that

S. Bigham (✉)
Department of Psychology, Bournemouth University,
Fern Barrow, Poole House, Poole, Dorset BH12 5BB, UK
e-mail: SBigham@bournemouth.ac.uk

J. Boucher · S. Anns
City University, London, UK

A. Mayes
University of Manchester, Manchester, UK

[1] 'High functioning autism' (HFA) is used here to refer to individuals with the triad of diagnostic impairments, but no intellectual disability or language impairment. It therefore includes individuals who may be diagnosed with Asperger syndrome.

impaired episodic memory derives from a hippocampal-related impairment of relational memory (such as is important for processing personally experienced events) with spared memory for single items or simple item-item associations (such as may subserve memory for decontextualized facts).

Interpretation of findings on memory in HFA as reflecting impaired relational memory with spared single-item memory is essentially close to the other major interpretation in the literature, namely that argued for by Minshew and her colleagues over many years (e.g., Minshew and Goldstein 2001; Williams et al. 2006). According to this interpretation, the uneven pattern of declarative memory abilities in high-functioning individuals with ASDs reflects a difficulty in processing complex information of all kinds, co-existing with intact ability to process simple, unstructured information.

In a wide-ranging discussion of the neuropsychology of declarative and working memory profiles in HFA, Joseph et al. (2005) note a further distinction commonly made in the literature on declarative memory. This is the distinction between 'recollection' and 'familiarity' (the latter sometimes referred to as 'recognition memory'). Recollection is defined as a kind of recall in which a recognised stimulus cues recall of contextual information experienced within the episode in which the stimulus was encountered. Familiarity is defined as a conscious feeling that one has experienced a stimulus before without necessarily recalling any other information. Thus, familiarity generally relates to single precepts or items (including complex items such as scenes). Joseph et al. further note that recollection is argued to be crucially dependent on the hippocampus, whereas the sense of familiarity that accompanies recognition is more dependent on extra-hippocampal rhinal and temporal cortex (Aggleton and Brown 1999, 2006; Montaldi et al. 2006; Mayes et al. 2007; Kirwan et al. 2008). Joseph et al. suggest that a combination of impaired recollection with intact familiarity fits well with the uneven memory profile in HFA. In particular, impaired recollection is consistent with impaired performance on tests of episodic memory (which involves the recall of contextual 'relational' information), whereas intact familiarity is consistent with good recognition memory and memory for decontextualized 'single item' facts.

The distinction between recollection and sense of familiarity is one that we have used in our own interpretation of memory profiles across the spectrum (Boucher et al. 2008a, b; Boucher et al. 2010). We have suggested that whereas recollection is mildly impaired in individuals with HFA leaving familiarity intact (as proposed by Joseph et al. 2005), both recollection and familiarity are impaired in individuals with low-functioning autism (LFA), with impaired familiarity being a major cause of the language and learning impairments associated with LFA. The suggestion of a dual impairment of both recollection and familiarity in LFA is novel and potentially important. Moreover the hypothesis of a selective impairment of recollection in HFA, leaving familiarity intact, although suggested in a speculative discussion by Joseph et al. (2005), has not been tested. Both these hypotheses are therefore in need of testing.

To do this it is necessary to assess recollection and familiarity separately from each other. However, this is difficult to do. This is because recollection and familiarity generally both contribute to performance on standard experimental tests of recall and recognition, whether the material to be remembered relates to a personally experienced event or to a decontextualized fact, word, single item or scene. The 'remember-know' paradigm (Gardiner and Java 1993) was specifically designed to assess recollection and familiarity separately. This paradigm is, however, dependent on complex verbal instructions and verbal understanding, and is not suitable for testing young children or intellectually disabled individuals. Moreover, the paradigm has been criticised on the grounds that it does not discriminate between the contributions of recollection and familiarity as satisfactorily as first appeared to be the case (Dunn 2004; Gardiner et al. 2006). The aim of the research presented here was, therefore, to develop methods of assessing recollection and familiarity separately in young or intellectually disabled participants, such as could be used to test our hypotheses relating to recollection and familiarity across the spectrum.

The work was carried out in two phases. These are reported separately below, because although producing one interesting finding, work carried out in Phase 1 was not entirely satisfactory in ways that are instructive and which led directly to formulating new and more successful paradigms in Phase 2.

## Phase 1

In Phase 1 we attempted to develop two parallel source memory tasks: a temporal source memory recall task to assess recollection, and a spatial source memory recognition task to assess familiarity. However, despite extensive piloting of a variety of methodological modifications, we were not able to devise a pair of comparable tasks that we could use with both high- and low-functioning ability groups simultaneously without producing ceiling effects in able young children or floor effects in teenagers with intellectual disability. We therefore narrowed our initial aim, and focussed on using a test of temporal source memory to assess recollection in LFA only. This is reported as Experiment 1, below.

## Experiment 1: A Test of Recollection in Teenagers with Low-Functioning Autism

In this experiment we used a temporal source memory (TSM) recall task to assess recollection of contextual information in teenagers with LFA with sufficient language to co-operate with formal testing. We compared the performance of this group with performance in an ability-matched group of young typically developing (TD) children and an age- and ability-matched group of teenagers with intellectual disability (ID) without autism. A study of temporal source memory using a slightly different method showed impairment in participants with HFA (Bennetto et al. 1996). Moreover, a study using the remember-know paradigm demonstrated impaired recollection in adults with Asperger syndrome (Bowler et al. 2000). On the basis of these findings on individuals with HFA, and in line with our own hypothesis, we predicted that: (1) the LFA group would be impaired relative to the TD group; (25) the LFA group would also be impaired relative to the ID group; and (3) there would be no significant difference between the performance of the TD and ID groups.

### Method

*Participants* Three groups took part in the study: a group of teenagers with LFA ($n = 29$); a group of TD children ($n = 23$) equated with the LFA group for verbal and nonverbal mental ages; and a group of teenagers with ID without autism ($n = 24$) equated with the LFA group for chronological age (CA), verbal and nonverbal ability. Baseline tests of verbal and nonverbal ability were the British Picture Vocabulary Scale (BPVS; Dunn et al. 1997) and the Wechsler Abbreviated Intelligence Scale (WASI; Wechsler 1999). Descriptive data for the three groups are shown in Table 1.

All the young people in the LFA group were attending special schools catering for students with intellectual disabilities, with or without autism. All had been diagnosed as autistic by experienced psychiatrists and clinical psychologists using DSM-IV (APA 1994) criteria, and all had scores of 30 or above on the Childhood Autism Rating Scale (CARS; Schopler et al. 1988) as completed by their class teacher at the outset of the study. All had verbal quotients of less than 75, with an age equivalent of more than 5;0 years on the BPVS.

Children in the TD control group were attending mainstream junior schools serving mixed catchment areas; had no record of autistic features of behaviour; and were described by their teachers as having no significant social, emotional, or cognitive problems. These children were selected to be of average ability, with verbal quotients on the BPVS of between 85 and 115.

The young people in the ID control group attended the same special schools as the LFA group and had no record of autistic features of behaviour. All had verbal age-equivalents of more than 5;0 years on the BPVS.

Informed consent to include children in this research was obtained from parents or guardians via each child's school prior to any child being seen for baseline testing. Following baseline testing, each child was asked if they would like to come and work with the Tester again, and told that it was alright to say 'no' if they did not want to. In this way, 'informed consent' was obtained from each child. No child declined to participate.

*General Points of Procedure* Children were tested individually in a quiet room, usually in their own school, but exceptionally in the child's own home. Tester and child sat opposite each other at a table. A stopwatch was used to maintain presentation rates and retention intervals. Verbal input from the tester during testing was restricted to phrases such as 'keep looking' or 'you're doing well' if any child's attention appeared to wander. A pre-prepared score sheet was used. Children were praised at the end of the task.

*Materials and Procedure* A pre-test was given prior to the main test. Two questions testing understanding of 'before' and 'after' were used: "When you go swimming do you take your socks off before or after you get into the water?" "Do you dry yourself before or after you've had a swim?" Correction was provided if necessary. To assess understanding of the left–right = earlier-later convention, a strip of card with three rectangles drawn on it was prepared with the words 'BEFORE' and 'AFTER' written in the left and right hand rectangles, leaving the centre one empty. The strip of card was placed in front of the child. The Tester placed the number 5 in the middle rectangle, and said: "If you count from 1 to 10, does 4 come before 5 or does it come after 5?" The Tester asked the child to place the number 4 and then the number 7 in the correct position on the strip. The same procedure was used to test placement of three pictures depicting a meal. No child failed the pre-test.

Practice in the procedure to be used in the main test was given to ensure that all children fully understood what they were required to do. Six everyday objects such as a comb and a pen were used, plus a Mars chocolate bar, and the strip of card used in the pre-test. The procedure used during practice was identical to that described below for the test proper. Two practice runs were given, with additional instruction and support provided on the first run, if needed to ensure correct responding. Any child who failed on more than 2 items in each of the practice runs was deemed not to understand the procedure. Two participants from the LFA

**Table 1** Phase 1: Participant details, including between-group similarities and differences on baseline measures

| Group (n) | LFA (29)<br>$\mu$ (sd)<br>(Range) | TD (23)<br>$\mu$ (sd)<br>(Range) | ID (24)<br>$\mu$ (sd)<br>(Range) | Group comparisons |
|---|---|---|---|---|
| Male/female | 2/27 | 5/18 | 7/17 | |
| CA in months | 173.5 (21.0) | 96.5 (18.4) | 171.2 (14.9) | LFA and ID equated |
| Range | 127–206 | 66–134 | 146–207 | |
| BPVS score | 76.2 (13.4) | 79.8 (15.9) | 76.9 (11.9) | Groups equated |
| Age equivalent | $\mu$ 7yrs 6ms | $\mu$ 8yrs 0ms | $\mu$ 7yrs 7ms | |
| Range | 53–101 | 58–108 | 54–100 | |
| n Tested | 29 | 23 | 24 | |
| WASI Vocabulary | | | | |
| $\mu$ score (sd) | 17.9 (5.9) | 23.1 (8.1) | 18.3 (4.5) | Groups equated |
| Range | 7–35 | 11–42 | 13–35 | |
| n Tested | 29 | 23 | 24 | |
| Similarities | | | | |
| $\mu$ score (sd) | 15.3 ( 5.4) | 20.4 (5.5) | 16.7 (4.3) | LFA and ID equated |
| Range | 1–24 | 12–31 | 9–28 | TD and ID equated |
| n Tested | 27 | 20 | 24 | TD > LFA ($p < .01$) |
| Blocks | | | | |
| $\mu$ score (sd) | 15.8 (16.5) | 15.0 (9.1) | 13.4 (12.1) | Groups equated |
| Range | 0–60 | 4–33 | 4–51 | |
| n Tested | 27 | 23 | 24 | |
| Matrices | | | | |
| $\mu$ score (sd) | 14.7 (7.8) | 16.3 (5.7) | 11.7 (5.6) | Groups equated |
| Range | 3–29 | 8–24 | 5–28 | |
| n Tested | 29 | 19 | 23 | |

group, 1 from the TD, and 3 from the ID group were excluded. Recalculation of the baseline scores shown in Table 1 excluding these participants showed no significant changes in group matching.

For the main test, a set of 16 everyday objects was assembled, different to those used for practice, plus a tube of Smarties and the strip of card used in the pre-test. At study, the Tester said: "I've got a whole lot of new things, different things to show you. And somewhere amongst this lot, there will be a tube of Smarties! No Mars bar this time, but Smarties instead. What you have to do is try to remember if you see the objects *before* or *after* you see the Smarties. Are you ready?" The objects were then presented one at a time in a predetermined order at a rate of 1 every 5 s (3–4 s looking time), and placed in predetermined orientation on the table in front of the child, each being removed from sight before presentation of the next item. The Smarties were presented as item 10. There was no retention interval, other than the time taken to give the test instruction.

At test the before/after strip was placed centrally and the Smarties placed in the centre rectangle of the strip. The Tester said: "Let's see if you can remember which things you saw *before* the Smarties, and which you saw *after* the Smarties. Are you ready?" The 16 objects were then presented one at a time in a predetermined order different to that used at study, the Tester asking: "Did you see this before you saw the Smarties? Or did you see it after the Smarties?" (alternating 'before' and 'after' in successive items). Children were encouraged to place the item onto the before/after strip, but if unwilling to do this, pointing or a clear indication of a choice was accepted, and the Tester placed the object to the side indicated by the child. Each object remained on the side selected by the child until the all of the remaining objects had been presented. If a child hesitated, or seemed reluctant to make a choice, they were encouraged to make a response, the Tester saying: "Just guess, quick as you can." If a child perseverated (defined as choosing the before or after side 5 times consecutively), the Tester said: "It's not always this side" (indicating). "It could be this side sometimes" (indicating).

### *Results*

Mean scores for each of the 3 groups on Experiment 1 are shown in Table 2. Some ceiling effects (scores of 16) and

**Table 2** Results of experiment 1 by group: mean scores, sds, ranges, numbers tested, and numbers at ceiling and at floor

| | LFA | TD | ID |
|---|---|---|---|
| Experiment 1: TSM (Maximum score 16) $\mu$ (sd) | 10.4 (2.7) | 12.4 (2.3) | 12.4 (2.6) |
| Range | 3–15 | 5–16 | 7–16 |
| No. tested | 27 | 22 | 21 |
| No. at ceiling | 0 | 1 | 3 |
| No. at chance | 4 | 1 | 1 |

some floor effects (scores at chance, defined as 8 or below) occurred, and the numbers of instances for any group are also shown in Table 2. Despite some individuals scoring at chance, all 3 groups scored significantly above chance on this task.

None of the measures on which the participant groups were equated correlated significantly with performance on the TSM task. None of these measures were therefore entered as covariates, following the arguments of Miller and Chapman (2001).

A one-way analysis of variance was performed to investigate group differences (LFA, TD, ID) on the TSM task. The effect of group was significant: $F(2, 65) = 4.98$, $p < .05$, partial eta squared = .13. Post hoc Bonferroni comparisons revealed that the LFA group scored significantly lower than the TD and ID groups ($p < .05$), supporting predictions (1) and (2). The two control groups scored almost identically ($p = 1.00$), supporting prediction (3).

*Comments on Experiment 1*

The results of this experiment are of interest in showing not only that memory for temporal contextual information is impaired in individuals with LFA, something that has not previously been demonstrated; but also that impaired memory for temporal source is specific to individuals with LFA as opposed to individuals with intellectual disability without autism. The results of Experiment 1 are also consistent with the hypothesis that recollection is impaired in LFA.

However, there are various difficulties of interpretation when using this paradigm as an assessment of recollection. In the first place, in this experiment we did not explicitly test the possibility that impaired familiarity also contributes to the impaired performance of LFA participants relative to the two comparison groups. Thus it could be the case that the participants with LFA recognised the everyday objects less well than the comparison groups (i.e. they experienced a less strong sense of familiarity) when these were presented in the test phase of the experiment, and that this affected their performance. Evidence against this argument comes from the results of an experiment not reported here, in which mean scores on a forced 2-choice recognition test using a different set of 16 everyday objects were over 14 in all three groups, with most participants performing at or very close to ceiling. Despite our arguments against this particular criticism as it applies to Experiment 1, the need to control explicitly for any contribution from familiarity is important in principle. We therefore took account of it when developing a different test of recollection in Phase 2.

In the second place, an occasional child in the TD group used overt verbal mediation during the study phase (e.g., saying 'pencil-before', 'scissors-after'), and it may be inferred that at least some others rehearsed silently. There was no evidence that either of the other two groups utilised verbal mediation. However, when developing an alternative test of recollection in Phase 2, we also took care to minimise possibilities for verbal mediation.

A third reason for caution in interpreting findings from Experiment 1 as demonstrating impaired recollection in LFA is that other evidence suggests that individuals with ASDs may have a specific impairment of time-related thinking (Boucher et al. 2007). In connection with this, it is noteworthy that performance on the TSM task did not correlate with any of our baseline measures, as was also the case in the studies reported by Boucher, Pons et al. It has been suggested that time-related thinking is a specific ability independent of other abilities commonly used as indicators of intelligence (Pons and Montangero 1999). It could be the case, therefore, that our findings on Experiment 1—(and also Bennetto et al.'s 1996 finding of impaired temporal source memory in HFA)—reflect a specific impairment of temporal information processing, rather than—or in addition to—a specific impairment of recollection.

One further query might be raised concerning interpretation of findings on Experiment 1, which is that floor effects (most prevalent in the LFA group) might have resulted from failure to understand the task. We would argue, however, that this is unlikely because our practice procedures were specifically designed to ensure that all children included in the experiment understood the task requirements and were able to respond correctly on an easy version of the task.

## Phase 2

In Phase 2 we developed a pair of related tests for assessing familiarity and recollection separately that worked well with young children. The following methodological points derived specifically from our experiences during Phase 1. First: because of problems in calibrating difficulty levels for groups of very mixed ability we focussed only on young children with HFA, developing and piloting tests avoiding either floor or ceiling effects in this group. Second: we explicitly minimised any contribution of familiarity on our test of recollection, and any contribution from recollection on our test of familiarity. Third: we minimised possibilities for verbal mediation. Fourth: in the test of recollection the contextual information to be recalled was not only arbitrarily related to the recognised cue (as in Experiment 1), but was unrelated to temporal order. Fifth: we continued to use practice procedures designed to ensure that floor effects did not result from failure to understand the procedures.

The tests we developed during Phase 2 of the study are reported as Experiments 2 and 3, below.

## Experiment 2: A Test of Familiarity in Young Children with High-Functioning Autism

It has been argued by Holdstock et al. (2002) and by Migo et al. (2009) that familiarity can be accurately assessed in recognition tests in which somewhat similar non-meaningful stimuli are presented at study and recognition is tested with a 4-choice forced recognition task using foils similar to the target item. Holdstock et al. argue that to succeed on this kind of recognition test the participant must rely very largely on the feeling that one of the four items has been seen before; recollection of the study episode contributes very little because of the similarities amongst both to-be-remembered stimuli and the foils amongst which target items are presented at test. Experiment 2 was a recognition test of this kind, with difficulty level calibrated to avoid floor or ceiling effects in young HFA or TD children.

Numerous studies of individuals with HFA have reported unimpaired recognition when the usual kinds of experimental or clinical tasks are used. On the basis of these findings we predicted that familiarity in the HFA group would be entirely normal, consistent with the arguments of Joseph et al. (2005) and our own hypothesis (Boucher et al. 2008b). Although the prediction here is very strong, it has not been tested previously.

### Method

*Participants* Two groups of primary school age children were recruited, a group of 18 children with HFA and a group of 29 TD children. None of the children had taken part in Experiment 1 or in the development and piloting of Experiments 2 and 3. The groups were equated for verbal comprehension as assessed by the BPVS (Dunn et al. 1997) but they differed in terms of CA, the HFA group being somewhat older than the TD group. Nevertheless, verbal IQ was not significantly different in the two groups (see Table 3).

Children in the HFA group were recruited from a range of specialist ASD units within mainstream primary schools. All children had scores of 25 or more on the Asperger Syndrome Screening Questionnaire (ASSQ; Ehlers et al. 1999) as scored by their teachers (the recommended cut-off point is 22). This questionnaire was considered to be more appropriate for the assessment of autism in high functioning children than the CARS, regardless of whether or not a child had in the past been described as having 'Asperger syndrome', or 'high functioning autism'. Children in the TD group were recruited from a range of mainstream primary schools with a socially mixed catchment area.

*Materials and Procedure* General points of procedure were the same as those described for Experiment 1, including the methods used to obtain informed consent. Practice was given to ensure that all children fully understood what they were required to do. For practice, 4 non-representational shapes cut out of red card and stuck on pieces of white card were used as stimuli, with recognition tested using an easy 4-choice forced recognition task (foils were relatively unlike target stimuli). The procedure used during practice was identical to that described below for the test proper. Two practice runs were given, with

**Table 3** Phase 2 participant details

| Group (n) | HFA (18) | TD (29) | Group comparisons |
|---|---|---|---|
| Male/female | 10/8 | 21/8 | |
| CA in months $\mu$ (sd) | 109.9 (21.1) | 95.6 (14.7) | $p < .01$ |
| Range | (76–161) | (60–124) | |
| BPVS raw score $\mu$ (sd) | 94.0 (18.1) | 90.3 (15.2) | Equated |
| Age equivalent in months | $\mu$ 117 | $\mu$ 110 | |
| Range | (68–138) | (57–121) | |
| Standard score $\mu$ (sd) | 103.7 (10.6) | 109.4 (11.6) | Equated ($p = .098$) |
| Range | (93–138) | (90–138) | |

additional instruction and support provided on the first run, if needed to ensure correct responding. All participants recognised at least 3 shapes correctly on one or both of the practice runs, and none were excluded.

For the main test, a set of 16 non-representational shapes were cut out of blue card and stuck onto pieces of white card approximately 15 × 10 cms. There were also 16 A4 cards, each containing one target plus 3 foils. An example of a target card and target-foil card is shown in Fig. 1.

The Tester introduced the study phase by saying: "We'll play that game again, but it will be a bit harder this time as I have many more shapes and they are different to the ones that I have just shown you."—(i.e. the practice materials). "Try to remember exactly what each shape looks like. They all look a bit the same, so you need to look at each one carefully. Are you ready?" Each card with the target shape was presented in a predetermined order and orientation, at three-second intervals. When all of the target items had been presented the Tester said: "Let's see how many of the shapes you can remember. I'm going to show you four shapes and I want you to tell me which one is exactly the same as one of the ones you saw just now. Are you ready?" The time taken to move from the study phase to the test phase constituted the retention interval, which was timed to last for approximately 15 s.

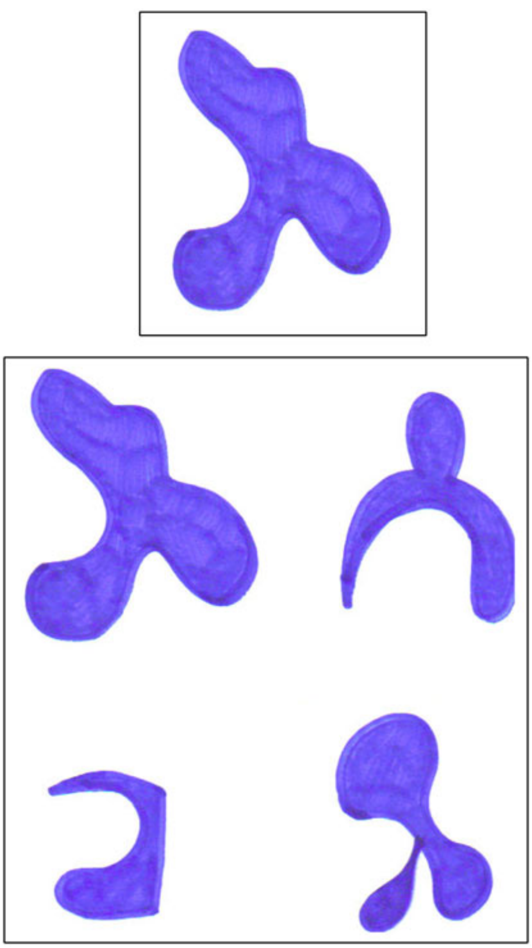

**Fig. 1** An Example of the target and target plus 3 foils used in experiment 2: shape recognition

During the test phase, the Tester presented the target-foil cards one at a time in a predetermined order, different to that used at study, and said, "Which one of these shapes have you seen before?" Children were required to point to the item of their choice. When the child had responded the target-foil card was removed from sight and the next one was presented. If the child was reluctant to respond the Tester elicited a response by saying "Just guess, quick as you can". There were no instances in which a child perseverated to a particular location.

Each correctly recognised target item was awarded one point. Incorrect responses were not awarded any points. Possible scores therefore ranged from 0 to 16.

### *Results*

Mean scores of the two groups on the shape recognition task in Experiment 2 are shown in Table 4. A one-way ANOVA showed that the effect of group was not significant, $F$ (1, 45) = .02, $p$ = .91. The two groups were equated for BPVS score, and this significantly correlated with performance on the recognition test ($r$ = .37, $n$ = 47, $p < .05$). However, entering BPVS score as a covariate did not affect the result.

### *Comments on Experiment 2*

In Experiment 2, young children with HFA and an ability-matched young TD group were given a recognition test designed to provide a relatively pure measure of familiarity. In this test, meaningless shapes were used as stimuli and a forced choice test utilising 3 foils not easily discriminable from the target stimulus. To perform well on this test participants must rely heavily on the feeling that the target stimulus is familiar. No floor effects occurred, indicating that all children understood the task requirements. Verbal mediation was essentially excluded by using meaningless shapes as stimuli. It might be argued that the significant correlation between recognition and BPVS scores could be taken as evidence that verbal mediation was not completely eliminated. However, it is possible that the correlation between recognition and BPVS scores derived from common dependence on some third factor, such as nonverbal intelligence, rather than reflecting use of verbal mediation. Moreover, in view of the fact that entering BPVS scores as a covariate did not alter the main finding, it is unlikely that even tighter control of possibilities for verbal mediation would produce a different result.

**Table 4** Results of experiments 2 and 3 by group: mean scores, sds, ranges, numbers tested, and numbers at ceiling and at floor

| | HFA | TD |
|---|---|---|
| Experiment 2: shape recognition (maximum score 16) $\mu$ (sd) | 12.7 (2.1) | 12.6 (2.3) |
| Range | 9–16 | 7–16 |
| No. tested | 18 | 29 |
| No. at ceiling | 1 | 1 |
| No. at floor | 0 | 0 |
| Experiment 3: action recall (number correct as a percentage of number correctly recognised) $\mu$ (sd) | 52.7 (.26) | 62.7 (.15) |
| Range | 0–80 | 30–90 |
| No. tested | 18 | 29 |
| No. at ceiling | 0 | 0 |
| No. at floor | 2 | 0 |

Finally, one child in each group performed at ceiling, suggesting that the difficulty level could be marginally increased in any future use of this paradigm with children of this age and ability level. This could be done either by increasing the number of target stimuli by one or two additions, or by decreasing the discriminability of foils very slightly.

In Experiment 2, children with HFA performed as well as ability-matched young TD children. Given that the method used in Experiment 2 was satisfactory, we conclude that familiarity is unimpaired in HFA, consistent with our prediction. The finding of unimpaired familiarity is not surprising in view of abundant evidence of unimpaired recognition in individuals with Asperger syndrome or high-functioning autism, combined with the fact that most recognition tests rely more on familiarity than on recollection. However, the experiment reported here is the first to assess familiarity excluding the usual contributions from recollection.

## Experiment 3: A Test of Recollection in Young Children with High-Functioning Autism

In Experiment 3 we assessed recollection in a source memory test in which recognised stimuli were some of the meaningless shapes used in Experiment 2, and the contextual information to be recalled consisted of a set of actions one of which was performed by the tester on each of these stimuli. The arbitrary nature of the links between a recognised shape and the action performed on it during the study episode minimises the role of familiarity at recall, forcing reliance on recollection of the study episode. We predicted that young children with HFA would be significantly impaired in their recall of actions, demonstrating a selective impairment of recollection, consistent with the arguments of Joseph et al. (2005) and our own hypothesis (Boucher et al. 2008b).

### *Methods*

*Participants* Participants were those tested in Experiment 2 (see Table 3).

*Materials and Procedure* General points of method were the same as those in the two previous experiments. For practice, 4 non-representational shapes (different to those used for practice in Experiment 2) cut out of red card and stuck on pieces of while card were used as stimuli, with recognition tested using an easy 2-choice forced recognition task. Each target shape was paired with a target-focused manual action, such as flicking it or moving it from side to side. The procedure used during practice was identical to that described below for the test proper. Two practice runs were given, with additional instruction and support on the first run, if needed to ensure correct responding. All participants recognized at least 3 shapes and recalled at least 2 actions on one or both of the practice runs, and none were excluded.

For the main test, 10 target shapes from Experiment 2 were used, and each of these was paired with only one foil (different to those used in Experiment 2) as shown in Fig. 2. Use of target shapes from Experiment 2, and use of a single foil, was designed to facilitate recognition, because we were aiming for 100% correct recognition in both groups to exclude any effects of impaired familiarity.

Each target shape was paired with a target-focused manual action, such as picking up the card and placing it face down or placing a clenched fist on the card. Actions involving bringing the card into relation with own body were not used because tests of imitation suggest that this kind of action may be difficult for children with autism to encode (Hobson and Lee 1999), whereas object-directed actions are imitated normally (Williams et al. 2004).

Experiment 3 immediately followed the shape recognition test in Experiment 2. Children were congratulated on their success in Experiment 2 and asked: “Do you want to

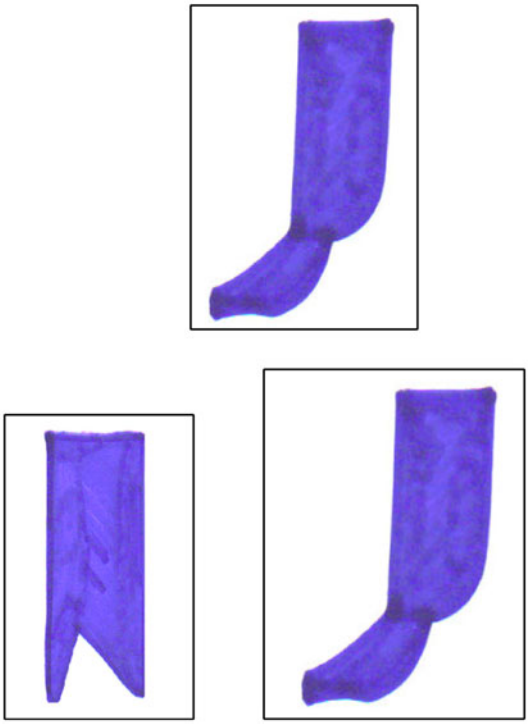

**Fig. 2** An example of the target and target-foil pair used in experiment 3: action recall

play another game? It is a bit harder but more fun?" Every child chose to continue. After the practice phase, the Tester said, "I've got some more shapes here,—different to what I showed you just now" (indicating the practice materials). "I'm going to do different actions with them, different to the ones that you just saw."—(i.e., the practice actions) "And I'm going to go a bit faster than before. Try to remember what each shape looks like and the action that goes with it. Are you ready?" The 10 target shapes were presented one at a time in a predetermined order and orientation, each being placed centrally in front of the child. After approximately 2 s, the Tester performed the action associated with the target item twice, quite slowly in succession, checking that the child was attending to the action. The target shape was then removed and the next shape-action pair was presented until all 10 shape-action pairs had been presented.

At test the Tester said: "Let's see how many of the shapes and actions you can remember. I'm going to show you two shapes and I want you to show me which one you just saw. Which one of these did you just see?"—placing the first target-foil pair centrally in front of the child, in prescribed orientation and with the position of the target stimulus to the left or right of the foil in predetermined order. The order in which target-foil pairs were presented was different to the order of presentation during study. If the child failed to respond or looked uncertain the Tester prompted by saying: "Have a guess. Which one did you see before, this one? Or this one?" If the child indicated the incorrect picture, that target-foil pair was removed and the next pair presented. However, if the child indicated the correct picture, the foil was removed, the target item moved to a central position in front of the child, and the Tester said: "What was the action that went with it?" 5-s were allowed for the child to reproduce the original action. The Tester prompted if necessary by encouraging the child to have a guess if uncertain.

Performance on the shape recognition pre-test was scored out of 10, each correctly recognized shape being awarded one point. Not all participants achieved 100% correct recognition, three children in the HFA group scoring 9 out of 10 and one child in the TD group scoring 8 out of 10. Scores for the main action recall test were therefore calculated as percentages of correct responses out of the number of action recall items given. 'Correct' recall was operationalized as an unambiguous attempt to reproduce the required action, not taking into account person-related detail such as the hand used, or the direction in which an action had been carried out (e.g. sliding the card to one or other side of the table).

### *Results*

Mean scores of the two groups on the action recall task are shown in Table 4.

Analysis of variance demonstrated a trend towards the predicted impairment in the HFA group: $F(1, 45) = 2.88$, $p = .096$. The two groups were equated on BPVS raw and standard scores, but only BPVS raw score significantly correlated with scores on the recall test ($r = .33$, $n = 47$, $p < .05$). Entering BPVS raw score as a co-variate strengthened the difference between these two groups for this task: $F\ (1, 44) = 4.36$, $p < 05$. On further investigation it was noted that the correlation between BPVS and recall scores was not significant when looking at the TD group alone ($p = .19$), but was significant for the HFA group ($r = .49$, $n = 18$, $p < .05$).

### *Comments on Experiment 3*

In Experiment 3 the same groups of children as those assessed in Experiment 2 were given an action-recall test designed to provide a relatively pure measure of recollection, operationalised as the cued recall of contextual information. In this test, cues were meaningless shapes, and the items of contextual information to be recalled were arbitrary actions. The use of meaningless shapes as cues minimised possibilities for verbal mediation of the shape-action links, and the arbitrary nature of the pairings between individual shape cues and the actions performed during the study phase minimised possibilities for successful guessing. We explicitly controlled for any contribution that familiarity might make to scores on this task by ensuring that recognition performance was at or near ceiling. In addition, we can be confident that familiarity did not contribute to task performance because recognition scores in Experiment 2 did not correlate with recall in Experiment 3 ($p = .56$). To perform well on the action recall test, therefore, participants must rely heavily on recollection of the study episode. Floor effects were largely avoided. However, two children in the HFA group performed at floor, suggesting that the difficulty level might be marginally decreased in any future use of this paradigm with children of this age and ability. This could be done either by reducing the number of actions for recall, or by increasing the salience of some actions. We do not consider that floor effects resulted from failure to understand the task requirements because children who failed the easy practice items had been excluded from the experiment.

It might be suggested that the HFA children's relatively poor scores on the action recall task resulted from impaired imitation. Not only does the small number of children at floor argue against this: the fact that all the children tested were able to produce remembered actions on the practice trials also argues against it. In addition, there was no evidence to suggest that any child encountered difficulty in the actual performance of a remembered action. Moreover, our criteria for correct imitation explicitly did not require the reproduction of person-related detail such as direction of movement or hand used, which is a source of particular difficulty for children with ASDs (Hobson and Lee 1999). It might also be suggested that the children with HFA might have performed better had they been required to imitate each action at study, rather than only observing the tester carrying out the actions. However, pilot work showed that imitating the actions at study significantly reduced typically developing children's memory performance (probably because it decreased attention to the shape-action association). In addition, we did not want to assess memory for self-performed actions, as memory for such actions may, again, present particular difficulty for individuals with ASDs (Russell and Jarrold 1999; Millward et al. 2000; but see Lind in press). Nevertheless, if using this paradigm in future work we would include as a precautionary measure a post-test of children's immediate imitation of the actions they had been asked to recall.

It might further be suggested that the results of an experiment reported by Hill and Russell (2002) throw doubt on our finding of impaired performance in an HFA group. These authors found no impairment of action recall in an LFA group relative to appropriate comparison groups. There were crucial methodological differences, however, between their experiment and ours. In Hill and Russell's experiment two everyday objects were shown together and an instruction for an action combining these objects (e.g. placing a ball on an upturned cup) was written on a card placed in front of the participant and read out by the Tester. Thus verbal mediation is not excluded but rather encouraged by the methodology. In addition, the possible combinations of any two objects were limited, increasing the likelihood of successful guessing. Not surprisingly in view of both these points, all three groups assessed in Hill and Russell's experiment performed well.

The contrast between Hill and Russell's finding and our own on Experiment 3 is methodologically instructive, however, in that it underlines the importance of minimising possibilities for both successful guessing and verbal mediation. In our experiment, successful guessing was almost certainly excluded by the novelty and arbitrariness of the actions performed on each object. The fact that entering BPVS scores as a co-variate increased group difference in our experiment suggests that we were less successful in completely excluding verbal mediation despite our efforts to do so. Most interestingly, however, the fact that correlation between BPVS scores and recall scores was driven by the HFA and not the TD group suggests that verbal mediation occurred mainly in the HFA group, almost certainly to compensate for impaired recollection.

In sum, we argue that it is safe to conclude that recollection is mildly impaired in HFA, consistent with our prediction. This conclusion is also consistent with Bowler et al.'s (2000) finding on adults with Asperger syndrome using the remember-know paradigm. However, as noted in the Introduction, some doubts have been raised about the accuracy of remember and know scores as measures of recollection and familiarity (Dunn 2004; Gardiner et al. 2006). With the minor modifications indicated above, the paradigm utilised in Experiment 3 may, therefore, offer a better method of obtaining a relatively pure measure of recollection in the most able individuals on the spectrum.

## Discussion

Memory in individuals with Asperger syndrome or high-functioning autism without language impairment, referred to collectively here as having HFA, has been quite extensively researched (Boucher and Bowler 2008; Boucher and Mayes 2010). Memory in low-functioning individuals has been much less researched over recent decades. Distinctions between memory profiles in HFA as opposed to LFA are rarely made, and it is common to conclude that memory in individuals across the spectrum is characterised by a mild impairment of episodic memory (sometimes referred to as relational memory) or of complex information processing, in combination with intact semantic memory (sometimes referred to as single-item memory) or simple information processing. Joseph et al. (2005) suggested that this kind of memory profile might be explained in terms of impaired recollection (which makes a major contribution to the recall of personally experienced episodes) in combination with intact sense of familiarity (which makes a major contribution to memory for decontextualized factual information).

We have argued that memory profiles in high- and low-functioning autism are different in ways that are important for understanding why some individuals with ASDs have additional language and learning impairments, whereas others do not (Boucher et al. 2008a, b). More specifically, we have hypothesised that individuals with HFA have a selective impairment of the process of recollection leaving sense of familiarity intact, whereas individuals with LFA have a combined impairment of both recollection and familiarity (Boucher et al. 2008a, b; Boucher and Mayes 2010). To test our hypotheses it is necessary to be able to assess recollection and familiarity separately using paradigms that are suitable for use with young or intellectually disabled individuals. However, currently no such paradigms exist, and the major aim of the study reported here was to develop such methods. Our primary aim was not, therefore, at this stage to test our hypotheses. However, we expected to obtain some preliminary data relating to these hypotheses.

The study reported here was undertaken in two stages, with most of the exploratory work carried out in Phase 1, and the development of methodologically tight procedures not achieved until Phase 2. Two major lessons were learned in Phase 1. First that it may not be possible to devise paradigms suitable for use with groups of able and learning disabled individuals simultaneously (even if equated for mental ages) whilst avoiding ceiling and floor effects in all groups. Second, certain methodological points must be stringently observed (these are fully discussed in the Comments following each of the reported experiments).

In Phase 1 we carried out one experiment (Experiment 1) that produced reportable and interesting results. This was a test of temporal source memory in teenagers with LFA. Although not fully satisfactory as a test of recollection (for reasons discussed in the relevant Comment section), Experiment 1 showed that temporal source memory is impaired in individuals with LFA and, moreover, that this impairment is specific to this group and does not extend to learning disabled individuals without autism. This finding is consistent with findings of impaired temporal recency in individuals with HFA (Bennetto et al. 1996), and of impaired time-related thinking in individuals with LFA (Boucher et al. 2007).

In Phase 2 we developed a pair of related tests to be carried out consecutively in a single session. These were a test of shape-recognition to assess familiarity (Experiment 2) and a test of action-recall to assess recollection (Experiment 3). We ran these tests with a group of young children with HFA and an ability-matched TD group and found familiarity to be intact but recollection to be mildly impaired in the HFA group, consistent with our predictions.

In our judgement, both paradigms used in Phase 2 are methodologically fit for purpose, although when using the action-recall test in future we would include a post-test of imitation ability. For future use with groups of individuals with LFA and intellectual disability without autism the difficulty level of both tasks will also need to be recalibrated to avoid floor effects. For the shape-recognition test this can be achieved by decreasing the number of target stimuli or increasing the discriminability between target stimuli and foils. For the action-recall test can be made easier by increasing the salience of the actions—something we experimented with during piloting with young TD children. It is not desirable to reduce the number of actions to be recalled because this would reduce test sensitivity unduly. Recalibration of the shape-recognition and action-recall tasks for use with less able groups is an essential but relatively easy next step in developing methods of testing our hypothesis of a combined impairment of recollection and familiarity in individuals with LFA.

One limitation of the paradigms developed in Phase 2 concerns the lack of comparability in test format between Experiments 2 and 3. Paradigms recently developed for use with adults by Migo et al. (2009)overcome this difficulty by utilising a pair of minimally but critically different forced choice recognition tests designed to discriminate between the contributions of familiarity and recollection. In future work the paradigms developed by Migo et al. together with the shape-recognition and action-recall tasks reported here, will be used to test the part of our hypothesis predicting that individuals with LFA differ from those with HFA in having impairments of both recollection *and* familiarity.

**Acknowledgment** This research was supported by an award from the Economic and Social Research Council of the UK (Ref. No.: RES-000-22-0989). We are grateful for the assistance of staff and students at the schools which took part in this study. We are also grateful to Tom Muskett for assistance in testing in Experiment 1, and to Heather Ball and Anneka Rees for initial piloting of the materials and procedures used in Experiments 2 and 3.

## References

Aggleton, J., & Brown, M. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences, 22*(3), 425–489.

Aggleton, J., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences, 10*, 455–463.

American Psychiatric Association. (1994). *DSM-IV: Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

Bennetto, L., Pennington, B., & Rogers, S. (1996). Intact and impaired memory functions in autism. *Child Development, 67*, 1816–1835.

BenShalom, D. (2003). Memory in autism: Review and synthesis. *Cortex, 39*, 1129–1138.

Boucher, J., Bigham, S., Mayes, A., & Muskett, T. (2008a). Recognition and language in low functioning autism. *Journal of Autism and Developmental Disorders, 38*, 1259–1269.

Boucher, J., & Bowler, D. M. (Eds.). (2008). *Memory in Autism*. Cambridge: Cambridge University Press.

Boucher, J., & Mayes, A. (2010). Memory. In D. Fein (Ed.), *The Neuropsychology of autism*. Oxford: Oxford University Press.

Boucher, J., Mayes, A., & Bigham, S. (2008b). Memory, language and Intellectual ability in low functioning autism. In J. Boucher & D. M. Bowler (Eds.), *Memory in Autism*. Cambridge: Cambridge University Press.

Boucher, J., Mayes, A., & Bigham, S. (2010). Memory in autistic spectrum disorders: Evidence, interpretations, and implications for behaviour. (under review).

Boucher, J., Pons, F., Lind, S., & Williams, D. (2007). Temporal cognition in children with autistic spectrum disorders: Tests of diachronic thinking. *Journal of Autism and Developmental Disorders, 37*(8), 1413–1429.

Bowler, D. M., & Gaigg, S. (2008). Memory in ASD: Enduring themes and future prospects. In J. Boucher & D. M. Bowler (Eds.), *Memory in Autism*. Cambridge: Cambridge University Press.

Bowler, D., Gardiner, J., & Grice, S. (2000). Episodic memory and remembering in adults with Asperger syndrome. *Journal of Autism and Developmental Disorders, 30*, 295–304.

Dunn, J. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*(2), 524–542.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale (BPVS)* (2nd edition ed.). Windsor: National Foundation for Educational Research.

Ehlers, S., Gillberg, C., & Wing, L. (1999). Screening questionnaire for Asperger syndrome. *Journal of Autism and Developmental Disorders, 29*, 129–141.

Gaigg, S. B., Gardiner, J. M., & Bowler, D. M. (2008). Free recall in autism spectrum disorder: The role of relational and item-specific encoding. *Neuropsychologia, 26*, 983–992.

Gardiner, J., Gregg, V., & Karayianni, I. (2006). Recognition memory and awareness: Occurrence of perceptual effects in remembering or in knowing depends on conscious resources at encoding, but not at retrieval. *Memory & Cognition, 34*(2), 227–239.

Gardiner, J. M., & Java, R. I. (1993). Recognising and remembering. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory*. Hove, UK: Psychology Press.

Hill, E., & Russell, J. (2002). Action memory and self-monitoring in children with autism: Self versus other. *Infant and Child Development, 11*(2), 159–170.

Hobson, P. R., & Lee, A. (1999). Imitation and identification in autism. *Journal of Child Psychology and Psychiatry, 40*, 649–659.

Holdstock, J., Mayes, A., Isaac, C., Gong, Q., & Roberts, N. (2002). Differential involvement of the hippocampus and temporal lobe cortices in rapid and slow learning of new semantic information. *Neuropsychologia, 40*(7), 748–768.

Joseph, R., Steele, S., Meyer, E., & Tager-Flusberg, H. (2005). Self-ordered pointing in children with autism: Failure to use verbal mediation in the service of working memory? *Neuropsychologia, 43*(10), 1400–1411.

Kirwan, C., Wixted, J., & Squire, L. (2008). Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. *Journal of Neuroscience, 28*(42), 10541–10548.

Lind, S. (in press). Memory and the self in autism: A review and theoretical framework. *Autism.*

Mayes, A. R., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences, 11*, 126–135.

Migo, E., Montaldi, D., Norman, K. A., Quamme, J., & Mayes, A. R. (2009). The contribution of familiarity to recognition memory is a function of test format when using similar foils. *Quarterly Journal of Experimental Psychology, 62*(6), 1198–1215.

Miller, G. A., & Chapman, P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40–48.

Millward, C., Powell, S., Messer, D., & Jordan, R. (2000). Recall for self and other in autism. *Journal of Autism and Developmental Disorders, 30*, 15–28.

Minshew, N., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry, 42*(8), 1095–1101.

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus, 16*, 504–520.

Pons, F., & Montangero, J. (1999). Is the diachronic thought a specific reasoning ability? *Swiss Journal of Psychology, 58*, 191–200.

Russell, J., & Jarrold, C. (1999). Memory for actions in children with autism: Self versus other. *Cognitive Neuropsychiatry, 4*, 303–331.

Salmond, C., Ashburner, J., Connelly, A., Friston, K., Gadian, D., & Vargha-Khadem, F. (2005). The role of the medial temporal lobe in autistic spectrum disorders. *European Journal of Neuroscience, 22*(3), 764–772.

Schopler, E., Reichler, R., & Renner, B. (1988). *The childhood Autism rating scale (CARS)*. Los Angeles: Western Psychological Services.

Toichi, M., & Kamio, Y. (2003). Long-term memory in high-functioning autism: Controversy on episodic memory in autism reconsidered. *Journal of Autism and Developmental Disorders, 33*, 151–161.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1–12.

Wechsler, D. (1999). *Wechsler abbreviated scales of intelligence*. San Antonio: Psychological Corporation.

Williams, D., Goldstein, G., & Minshew, N. (2006). The profile of memory function in children with autism. *Neuropsychology, 20*(1), 21–29.

Williams, J., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders, 34*, 285–300.