
This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/24832/

Link to published version: http://dx.doi.org/10.1002/aur.2387

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.
Short-Term Memory Span and Cross-Modality Integration in Younger and Older Adults With and Without Autism Spectrum Disorder

Melanie Ring ©, Bérengère Guillery-Girard, Peggy Quinette, Sebastian B. Gaigg ©, and Dermot M. Bowler ©

This study tested whether adults with autism spectrum disorder (ASD) show the same pattern of difficulties and absence of age-related differences in short-term memory (STM) as those that have been reported in episodic long-term memory (LTM). Fifty-three adults with ASD (age range: 25–65 years) were compared to 52 age-, biological sex-, and intelligence-matched typically developing (TD; age range: 21–67 years) adults on three STM span tasks, which tested STM performance for letters (Verbal), grid locations (Visuospatial), or letters in grid locations (Multimodal). A subsample of 34 TD and 33 ASD participants ranging in age from 25 to 64 years completed a fourth Multimodal Integration task. We also administered the Color Trails Test as a measure of executive function. ASD participants’ accuracy was lower than that of the TD participants on the three span tasks (Cohen’s $d$: 0.26–0.50). The Integration task difference was marginally significant ($p = .07$) but had a moderate effect size (Cohen’s $d = 0.50$). Regression analyses confirmed reduced STM performance only for older TD participants. Analyses also indicated that executive processes played a greater role in the ASD group’s performance. The demonstration of similar difficulties and age-related patterning of STM in ASD to those documented for LTM and the greater recruitment of executive processes by older ASD participants on the Integration task suggest a compensatory role of frontal processes both as a means of achieving undiminished task performance and as a possible protection against older age cognitive decline in ASD. Longitudinal research is needed to confirm this. Autism Res 2020; 00: 1–15. © 2020 The Authors. Autism Research published by International Society for Autism Research and Wiley Periodicals LLC.

Lay Summary: Little is known about short-term memory (STM) in younger and older adults with autism spectrum disorder (ASD). This study tested different kinds of STM and showed that ASD adults remembered shorter sequences of letters, crosses, or letters in grid cells less well than matched participants with typical development. However, older ASD individuals performed similarly to younger ASD individuals, nor showing the reduction in performance usually seen with older age. The data suggest that ASD individuals use different underlying mechanisms when performing the tasks and that this might help protect their memory as they grow older.

Keywords: autism spectrum disorder; short-term memory; span; binding; integration

Introduction

Research into long-term memory (LTM) in individuals with autism spectrum disorders (ASD) is converging on a picture that emphasizes difficulties in the processing of complex information [Minshew & Goldstein, 1998; Minshew, Johnson, & Luna, 2000; Williams, Goldstein, & Minshew, 2006a, 2006b] and more precisely, the flexible binding and rebinding of elements of experience that define particular episodes in memory [Bowler, Gaigg, & Lind, 2011]. The empirical basis of these conclusions ranges from demonstrations of difficulties with free recall compared to cued recall or recognition [see Ben Shalom, 2003; Boucher, Mayes, & Bigham, 2012; Desarnay et al., 2020], both in target and in source memory [Bowler, Gardiner, & Berthollier, 2004], difficulties in episodically recollecting the personally experienced past [Bowler, Gardiner, & Grice, 2000; Cooper, Plaisted-Grant, Baron-Cohen, & Simons, 2017; Maister, Simons, & Plaisted-Grant, 2013; Souchay, Wojcik, Williams,
Klinger, 2017 for more mixed results. The aging analogy, 2018 but see Geurts & Vissers, 2012; Powell, Klinger, & Geurts, 2015; Ring, Gaigg, & Bowler, 2016; Roestorf, 2014) suggested that ASD is characterized by difficulties in relational binding, which gives rise not only to the patterning of memory difficulties described above, but also to other difficulties experienced by autistic people, such as difficulty in mentalizing and problems in utilizing meaning in support of recall [Bowler, Matthews, & Gardiner, 1997; Cooper & Simons, 2019]. Bowler et al. (2011) also drew on current neural models of relational binding in episodic memory [Davachi, 2006; Opitz, 2010] to suggest a possible fronto-hippocampal basis for these difficulties [see Hogeveen, Krug, Geddert, Ragland, & Solomon, 2020; Cooper et al., 2017; Gaigg, Bowler, Ecker, Calvo-Merino, & Murphy, 2015 for supporting evidence].

The striking similarity between memory difficulties in ASD and those seen in healthy aging adults [Craik & Anderson, 1999; Craik & Salthouse, 2000; De Beni et al., 2013; Klencklen, Lavenex, Bradner, & Lavenex, 2017; Naveh-Benjamin, 2000; Naveh-Benjamin & Mayr, 2018; Tse, Crabtree, Islam, & Stott, 2019], led Bowler and colleagues [Bowler, Gardiner, & Gaigg, 2007; Bowler et al., 2004] to propose an "aging analogy" for memory in autism. The analogy has been supported by cross-sectional studies showing fewer age-related differences in memory between older and younger ASD adults compared to neurotypical controls [Lever & Geurts, 2016; Lever, Werkle-Bergner, Branmaier, Ridderinkhof, & Geurts, 2015; Ring, Gaigg, & Bowler, 2016; Roestorf, 2018 but see Geurts & Vissers, 2012; Powell, Klinger, & Klinger, 2017 for more mixed results]. The aging analogy has also been supported by the similar difficulty experienced by adults with ASD as that of healthy older neurotypical participants when asked to recognize previously studied episodically defined combinations of object features [Bowler, Gaigg, & Gardiner, 2014] accompanied by intact recognition of individual features. Bowler et al. [2014] interpreted these findings as supporting Bowler et al.’s [2011] argument for a relational binding deficit in ASD mentioned above and which echoes accounts of the memory difficulties faced by healthy older neurotypical individuals. Prominent among these is Naveh-Benjamin’s associative deficit hypothesis (ADH) [Naveh-Benjamin, 2000; Peterson & Naveh-Benjamin, 2016; Naveh-Benjamin & Mayr, 2018], which sees age-related declines in declarative episodic memory as the results of a diminished capacity to bind together the defining elements of an episode. Although considerable research supports the ADH [see Old & Naveh-Benjamin, 2008 for review; Bastin et al., 2013; Craik, Luo, & Sakuta, 2010; Wang, Dew, & Giovanello, 2010], Kimmsse, Zimmer, and Ecker [2018] note that age-related associative difficulties seem to be less in evidence in studies that use working memory (WM) [Baddeley, 2012] or short-term memory (STM) paradigms possibly because of procedural differences between methods used to test WM/STM and LTM [see also Allen, 2015]. Reduced age-related difficulties in WM as opposed to LTM in ASD may also result from different underlying neural and cognitive mechanisms. Whereas LTM is thought to depend upon the hippocampus [Opitz, 2010], WM is supported by fronto-parietal brain regions [Chai, Abd Hamid, & Abdullah, 2018].

Although the conclusions that could be drawn from very early studies documenting diminished STM and immediate memory in ASD were limited by methodological issues [Poirier, Martin, Gaigg, & Bowler, 2011], more recent, better-controlled studies reveal ASD-specific difficulties, especially in serial recall [Poirier et al., 2011; Bowler, Poirier, Martin, & Gaigg, 2016], which revealed that ASD participants were less sensitive to the temporal and spatial context of items in memory. Recent meta-analyses of WM in ASD show that the majority of investigations demonstrate diminished performance in ASD compared to typically developing (TD) participants [see Desaunay et al., 2020; Habib, Harris, Pollick, & Mellelve, 2019; Wang et al., 2017]. Desaunay et al.’s meta-analysis also reports a similar patterning of STM performance—diminished recall and intact recognition—to that reported in the LTM, declarative memory literature described earlier, although the number of studies in that meta-analysis was small, hence the need for more investigations like the present one. Barendse et al. [2013] in a review of WM and its neuropsychological correlates also describe WM as enabling the “online processing of complex cognition” (p. 1) and conclude that non-intellectually disabled adolescents with ASD are often reported as having difficulties with spatial WM that increase when more complex information places greater demands on the WM system, echoing earlier, complexity-oriented accounts of LTM in ASD [Williams et al., 2006a, 2006b]. These observations and findings suggest that there may be common, memory-specific processes that coexist with executive difficulties across the STM and LTM systems.

The coexistence of an executive component alongside the short-term stores in the WM system [Baddeley, 2012; Baddeley & Hitch, 1974] has led some authors to argue that WM difficulties in ASD (as well as in aging) might reflect executive rather than specifically memory difficulties in this population [see Bowler et al., 2016; Kercood, Grskovic, Banda, & Begeske, 2014]. Executive function (EF) difficulties have long been known to be a feature of the clinical picture of ASD [Hill, 2004a, 2004b; Demetriou et al., 2018; Johnston, Murray, Spain, Walker, & Russell, 2019], although there is considerable overlap with other clinical conditions [see, e.g., Geurts, Verté, Oosterlaan, Roeyers, & Sergeant, 2004]. EF difficulties have also long been known to be a feature of older age [West, 2000] and many authors have argued that they contribute
significantly to other aspects of cognitive decline in older age [Salthouse, Atkinson, & Berish, 2003].

In an attempt to explore further the question of binding and EF in the STM of healthy aging individuals, Quinette and colleagues [Quinette et al., 2013; Lecouvey et al., 2015] designed four tasks, a Verbal Span, Visuospatial Span, a Multimodal Span, and a Multimodal Integration task to assess multimodal binding and maintenance of information over the short term (see “Methods” section below for further details of these procedures). The theoretical context of their work was that of the episodic buffer [Baddeley, 2000, 2010; Baddeley, Allen, & Hitch, 2011], a subsystem of WM, which holds together information from a range of sources as an interface between the other WM sub-systems and LTM. Groups of healthy adults ranging in age from 18 to 85 years showed an age-related decline or a negative correlation with age on a combined measure of performance on these four tasks [Lecouvey et al., 2015; Quinette et al., 2013]. In addition, Lecouvey et al. [2015] report that this decline was associated with behavioral measures of inhibition and processing speed as well as with an altered metabolism (as measured by Positron Emission Tomography) in frontal and cingulate cortices. These associations disappeared when behavioral measures of inhibition and processing speed were controlled, suggesting an important role for these executive processes for performance on the experimental tasks. These authors also noted the similarity between some of their behavioral and imaging findings and those reported in the ASD literature and drew on the existing literature on diminished EF and prefrontal activity to speculate that ASD individuals might show a similar pattern of functioning on their four tasks to the one they found for healthy aging individuals.

The first aim of the present study follows from this speculation by attempting to establish whether, at a behavioral level, individuals with ASD show a similar patterning of performance across the four tasks mentioned above to those found in healthy older participants. Our second aim is to determine whether or not the LTM difficulties seen in ASD, particularly diminished recall and poorer memory for more complex material are also mirrored in Lecouvey et al.’s set of STM tasks. Our prediction is that the ASD participants should show diminished performance on the Verbal, Visuospatial, and Multimodal span tasks and that the difference would be greater on the last two of these tasks compared to the first, principally because of existing findings of greater ASD-related difficulty on spatial WM tasks [Alloway, Seed, & Tewolde, 2016; Christ et al., 2017; Steele, Minshew, Luna, & Sweeney, 2007, but see Desaunay et al., 2020] but also because the last two tasks involve a greater number of cognitive operations [Steele et al., 2007; Williams, Goldstein, Carpenter, & Minshew, 2005]. Our third aim is to use Lecouvey et al.’s Multimodal Integration task to test whether ASD individuals show the same difficulties with the episodic buffer as they found with an older typical sample. Our prediction is that our ASD participants should show diminished performance on this task for three reasons. First, because the task involves the integration of a number of separate cognitive processes and as such can be called a complex task. Second, it involves “ternary” or three-way relations [Halford, 1992]. Bowler et al. [2011] argue that these pose particular difficulty for individuals on the autism spectrum because of the inherent complexity of ternary processing and because three-way relations emerge later in development [Halford, 1992] and thus may not develop completely in ASD. Furthermore, Bowler et al. [2011] suggest that ternary processing may also underlie false belief understanding and joint attention—two processes that are also difficult for individuals with ASD. Additionally, in light of Lecouvey et al.’s [2015] findings on healthy older adults, the aging analogy of memory in ASD [Bowler et al., 2004, 2007] would predict ASD-related diminished performance on the Multimodal Integration task. We compared the performance of a group of adults with a diagnosis of ASD with a group of matched neurotypical adults on the four tasks described by Lecouvey et al. [2015] and Quinette et al. [2013]. Our participants ranged in chronological age from 21 to 67 years, which also allowed us to assess the effect of older age on performance. Because of previous finding of reduced age-related memory differences between younger and older ASD as opposed to TD individuals in LTM [e.g., Lever & Geurts, 2016; Ring et al., 2016; Roestorf, 2018], we predicted only the TD participants would show age-related reduction in performance on all four tasks. Finally, in view of Lecouvey et al.’s [2015] documentation of the role of frontal processes in performance on the tasks used here, we were able to include data from the Color Trails Test (CTT) [D’Elia, Satz, Uchiyana, & White, 1996], which were available for a subsample of 52 ASD and 46 TD participants. The CTT was used because it is language and culture free, it is quick and easy to administer, it has been studied extensively and is being used frequently in studies assessing EFs. It enables comparison with previous studies, for example, Bowler et al. [2014], who used it in their study of relational memory. Our prediction is that the memory difficulties in the ASD group observed in the STM tasks used in this study go beyond the well documented EF difficulties and significant differences between ASD and TD individuals will remain present even when EF difficulties are controlled for.

Methods
Participants: Span Tasks
Fifty-three adults with ASD (42 men, $M_{\text{age}} = 43.66$ years, age range: 25–65 years) and 52 TD participants (38 men,
Table 1. Descriptive Statistics for Autism Spectrum Disorder (ASD) and Typically Developing (TD) Individuals for the Span Tasks at the Top and the Integration Task at the Bottom.

<table>
<thead>
<tr>
<th></th>
<th>ASD</th>
<th>TD</th>
<th>t (df)</th>
<th>p</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Span task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>43.66</td>
<td>42.40</td>
<td>0.51 (103)</td>
<td>.61</td>
<td>0.10</td>
</tr>
<tr>
<td>VIQ/VCI a</td>
<td>113</td>
<td>114</td>
<td>0.47 (103)</td>
<td>.64</td>
<td>0.09</td>
</tr>
<tr>
<td>PIQ/PRI b</td>
<td>117</td>
<td>108</td>
<td>0.37 (103)</td>
<td>.72</td>
<td>0.07</td>
</tr>
<tr>
<td>FIQ c</td>
<td>111</td>
<td>111</td>
<td>0.19 (92) d</td>
<td>.85</td>
<td>0.04</td>
</tr>
<tr>
<td>AQ e</td>
<td>35.13</td>
<td>35.96</td>
<td>16.04 (95.82)</td>
<td>.000</td>
<td>3.12</td>
</tr>
<tr>
<td>CTT Int f</td>
<td>0.11</td>
<td>0.12</td>
<td>0.24 (97) g</td>
<td>.81</td>
<td>0.05</td>
</tr>
<tr>
<td>ADOS-C h</td>
<td>2.49 (0–6)</td>
<td>1.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-RSI i</td>
<td>5.89 (1–13)</td>
<td>2.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-Total j</td>
<td>8.37 (3–17)</td>
<td>3.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-Im k</td>
<td>1.21 (0–2)</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-SB l</td>
<td>1.31 (0–5)</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Integration task subsample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>43.03</td>
<td>43.24</td>
<td>0.07 (65)</td>
<td>.94</td>
<td>0.03</td>
</tr>
<tr>
<td>VIQ/VCI a</td>
<td>113</td>
<td>116</td>
<td>0.63 (65)</td>
<td>.53</td>
<td>0.15</td>
</tr>
<tr>
<td>PIQ/PRI b</td>
<td>107</td>
<td>111</td>
<td>1.03 (65)</td>
<td>.31</td>
<td>0.25</td>
</tr>
<tr>
<td>FIQ c</td>
<td>111</td>
<td>112</td>
<td>0.19 (54) m</td>
<td>.85</td>
<td>0.05</td>
</tr>
<tr>
<td>AQ e</td>
<td>36.94</td>
<td>36.45</td>
<td>15.01 (58.40)</td>
<td>.000</td>
<td>3.65</td>
</tr>
<tr>
<td>CTT Int f</td>
<td>0.14</td>
<td>0.12</td>
<td>0.39 (60) n</td>
<td>.70</td>
<td>0.10</td>
</tr>
<tr>
<td>ADOS-C h</td>
<td>2.00 (0–4)</td>
<td>1.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-RSI i</td>
<td>5.88 (3–11)</td>
<td>2.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-Total j</td>
<td>7.88 (3–15)</td>
<td>3.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-Im k</td>
<td>1.00 (0–2)</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-SB l</td>
<td>1.18 (0–5)</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table Notes:**
- a Verbal IQ (WAIS-IIIUK) or Verbal Comprehension Index (WAIS-IVUK).
- b Performance IQ (WAIS-IIIUK) or Perceptual Reasoning Index (WAIS-IVUK).
- c Full-scale IQ (WAIS-IIIUK or WAIS-IVUK).
- d A FIQ score was only available for 94 (51 ASD, 43 TD) of the 105 (53 ASD, 52 TD) participants.
- e AQ—Autism-Spectrum Quotient.
- f CTT Int—Color Trails Test Interference Score.
- g A CTT Interference score was only available for 99 (52 ASD, 47 TD) of the 105 (53 ASD, 52 TD) participants.
- h ADOS Communication subscale.
- i ADOS Reciprocal Social Interaction subscale.
- j ADOS Total score—Communication + Reciprocal Social Interaction.
- k ADOS Imaginability/Creativity subscale.
- l ADOS Stereotyped Behaviors and Restricted Interests. For ADOS scores range in brackets.
- m A FIQ-score was only available for 56 (32 ASD, 24 TD) of the 67 (33 ASD, 34 TD) participants.
- n A CTT Interference score was only available for 62 (29 ASD, 33 TD) of the 67 (33 ASD, 34 TD) participants.

$M_{age} = 42.40$ years, age range: 21–67 years) individually matched on Verbal Intelligence Quotient (VIQ) or Verbal Comprehension Index (VCI), Performance IQ (PIQ), or Perceptual Reasoning Index (PRI) and Full-scale IQ (FIQ) as measured by the third or fourth edition of the Wechsler Adult Intelligence Scale (WAIS-IIIUK or WAIS-IVUK) [The Psychological Corporation, 2000, 2008] were tested. Groups were closely matched on biological sex, $X^2 = 0.55$, $p = .46$, and chronological age (see Table 1). Participants were recruited through a database of individuals with whom the Autism Research Group at City, University of London is in regular contact and, in addition, through newspaper advertisements, flyers, and word of mouth. All participants were native English speakers.

All ASD individuals had received a clinical diagnosis according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) criteria [American Psychiatric Association, 2000] prior study. As a means of sample description to enable better comparison with other studies in terms of ASD severity, 35 of the ASD participants were available to take part in the Autism Diagnostic Observation Schedule (ADOS) [Lord et al., 1989] administered by researchers trained to
research reliability standards on this instrument.² Nine of these individuals scored just below the total cut-off score of 7 for ASD. They were nevertheless included in the study since records confirmed that they all had a clinical diagnosis of an ASD, which was our main inclusion criterion.³ TD individuals were included if they did not report taking psychotropic medication or a personal or family history of a psychological or neurodevelopmental disorder including autism and considering first- and second-degree relatives. All participants filled in the Autism Spectrum Quotient (AQ) [Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001] with the ASD group presenting significantly higher values compared to the TD group (Table 1). Fifty-two of the ASD and 47 of the TD participants took part in the CTT.

Participants: Integration Task

A subset of the participants just described took part in the Multimodal Integration task described by Lecouvey et al. [2015]. The sample consisted of 33 individuals with ASD (27 men, M_age = 43.03 years, age range: 25–64 years) and 34 with TD (23 men M_age = 43.24 years, age range: 21–67 years). Both groups were individually matched on VIQ, PIQ, FIQ, and closely matched on biological sex, X² = 0.84, p = .36, and chronological age (see Table 1). Of the 18 participants who received the ADOS, 5 scored below the cut-off. All filled in the AQ and scored significantly higher than the TD group (see Table 1). Thirty-two of these ASD and 29 of these TD participants received the CTT.

All participants gave informed consent prior study and were reimbursed for their time and travel expenses according to standard university fees. This study was approved by the ethics committee of the Psychology Department of City, University of London and the procedures used in this study adhered to the guidelines set out by the British Psychological Society.

Materials and Procedure

Participants were tested individually, and testing took about 2 hours. The order of task presentation was counterbalanced across participants, with members of each matched pair (one ASD and one TD individual with similar IQ) receiving the same presentation order. Each of the tasks lasted about 10 min and they were separated by about 20 min. Each of these 20-min intervals was filled with an unrelated task such as parts of the WAIS-IV UK [The Psychological Corporation, 2000, 2008] or other neuropsychological tests.

Span Tasks

We used the three STM span tasks adapted from Quinette et al. [2006], examples of which are presented in Figure 2. Before every task, participants were given the instructions, asked whether they had any questions and were then given the first trial. The tasks presented randomly generated sequences of letters, sequences of crosses in cells of a 4 × 4 grid or sequences of letters in cells of a 4 × 4 grid. Items were presented at a speed of approximately one item per second interspaced with blank screens or blank grids. Item presentation ended with the presentation of a blank screen or a blank grid and a “beep” sound asking participants to reproduce the items in presentation order. The examiner noted down the answers on the answer sheet. After recall of every trial, the examiner pressed the space bar to continue the task.

In the Verbal span task, participants saw sequences of phonologically dissimilar consonants, which were presented as single capital letters in the center of the screen. After a “beep,” participants were asked to reproduce the letters in the order of presentation by speaking them out loud to the examiner. In the Visuospatial span task, participants were presented with sequences of capital letters X in the cells of a 4 × 4 grid. After a “beep,” participants were asked to reproduce the previously marked locations in the grid in the correct order by touching them on the computer screen. The Multimodal span task presented a combination of both of these tasks. Participants saw sequences of phonologically dissimilar consonants as single capital letters in the cells of a 4 × 4 grid. They were asked to remember the letters, the locations in which they were presented as well as the presentation order. After a “beep” sound, participants were asked to reproduce the letters and grid cells by naming the presented letters in the correct order and simultaneously touching the grid cells which the letters were presented in on the computer screen. Span length ranged from 3 to 11 items. Participants were presented with three trials per level. The task ended when a participant failed two out of the three trials of a certain level or when the third stage involving 11 items was reached. For each participant and task, we calculated the corrected accuracy score as the proportion of correctly recalled sequences less the proportion of incorrect sequences. The three tasks are illustrated in Figure 1.

Integration Task

Before or after the three STM span tasks, participants were presented with a STM Multimodal Integration task [Quinette et al., 2006] written in Microsoft Visual Basic 6. The examiner explained the task to the participant
using example screens followed by two self-paced and one computer-paced practice trial. Once it was clear that participants understood the procedure, task administration started. Participants were given 20 trials with the opportunity for a short break after 10 trials. On every trial, participants were shown a $4 \times 5$ grid containing four phonologically dissimilar consonants presented as capital letters in different colors (yellow, green, red, blue) in the center of the grid. Four crosses in the same colors as the letters were placed randomly in the cells of the grid.

---

**Figure 1.** Examples for Verbal (a), Visuospatial (b), and Multimodal (c) span tasks with three items each. Item presentation (letter, grid cells marked with X or letters in the cells of the grid) was interspaced by blank screens or grids. One trial finished with a blank screen or grid and a “beep” sound asking participants to recall the presented items and their order (and locations).

**Figure 2.** Procedure of the short-term memory integration task. Participants had to form letter-location associations memorizing the letter in the location with the cross of the same color. After a retention interval of 1 sec participants had to respond with yes if the small black letter corresponded to the letter-location association formed previously (target trial, top) and no if the small black letter did not match the previously studied letter-location association (lure trial, bottom).
Stimuli were shown for 5 sec each within which participants had to mentally replace the letters in the cells with the crosses of the same colors by forming letter-color-location associations. After a 1-sec presentation of a fixation cross, participants saw another 4 × 5 grid containing a small black letter in one of the grid cells. They then had to decide whether the location of the letter corresponded to one of the four letter-color-location associations they had formed earlier and were given 4 sec to indicate their response by pressing the correct key (yes or no) on the keyboard before the next trial started with the presentation of a fixation cross for 2 sec followed by another screen presenting another four letters and crosses. Figure 2 sets out the task procedure. Corrected accuracy scores were calculated as set out in Quinette et al. [2006, p. 2512].

**Color Trails Test**

A subsample of 52 of the ASD and 47 of the TD participants who took part in the Span tasks and of 32 ASD and 29 TD participants who took part in the Integration task was available to also take part in the CTT [D’Elia et al., 1996]. This measure was developed as a potentially less culturally biased version of the Trail Making Test [Reitan, 1971] with the advantage of not including an alphabetical component in the task that could potentially disadvantage participants. It consists of two trials. Trial 1 measures sustained attention and processing speed and requires participants to join up in numerical order circles containing the numbers 1-25 randomly distributed on the page. Trial 2 measures attentional shifting, inhibition, and sequencing and comprises two sets of 25 circles, one yellow and one pink, each set containing the numbers 1-25. The participant has to join the circles in numerical order alternating between pink and yellow circles. We used an interference score as a measure of EF, which was calculated by subtracting the standardized score for Trial 1 time from the standardized score for Trial 2 time and dividing the result by the standardized score for Trial 1 time. A higher score indicates more interference caused by the alternating demands in Trial 2 partialing out the effects of undivided attention and perceptual tracking as also measured by Trial 1.

**Results**

The data were analyzed using chi-squared tests for nominal data, independent samples t-tests, repeated measures analysis of variances (ANOVAs) and covariances (ANCOVAs), bivariate correlations, and linear regression analyses. In the case of significant differences, Bonferroni-corrected post hoc tests were used. Greenhouse Geisser correction was applied when the Sphericity assumption was violated. The level of significance was set to .05 and one-tailed tests were used in the case of direct tests of the directional predictions made in the Introduction. All other significance levels reported are two-tailed.

**Participant Characteristics**

Analysis of the participant data set out in Table 1 shows that the only significant between-group difference was on AQ with higher scores for ASD compared to TD adults, which had a large effect size. Effect sizes for the other variables were negligible to small.

**Span Tasks**

The data for the corrected accuracy scores for each of the three span tasks are set out in Figure 3 and were analyzed with a 2 (Group) × 3 (Task: Verbal, Visuospatial, Multimodal) repeated measures ANOVA.

As shown in Figure 3, the ASD group showed lower corrected accuracy scores than the TD group on all three span tasks. For both groups the Verbal task showed the highest performance and the Multimodal task the lowest. This pattern was confirmed by significant main effects for Group, F (1,103) = 6.44, p < .05, $\eta^2_g = 0.06$, and Task, $F$ (2,206) = 60.73, $p < .001$, $\eta^2_g = 0.37$, but no significant interaction, $F$ (2,206) = 0.53, $p = .59$, $\eta^2_g = 0.01$. Separate follow-up analyses on each task revealed significant main effects for Group only for the Visuospatial, $t$ (103) = 2.53, $p < .02$, $d = 0.49$, confidence interval (CI): 0.10, 0.88, and Multimodal tasks, $t$ (103) = 2.57, $p < .02$, $d = 0.50$, CI: 0.11, 0.89, but not for the Verbal task, $t$ (103) = 1.56, $p = .12$, $d = 0.31$, CI: −0.08, 0.69. Because of the

---

\[ \text{Corrected Accuracy Score} = \frac{\text{Hits} - (\text{omissions} + \text{false alarms})}{\text{Total number of responses}} \]

\[ \eta^2_g \] is the partial eta squared, a measure of effect size.

Figure 3. Corrected accuracy scores for Verbal, Visuospatial and Multimodal short-term memory and short-term memory Multimodal Integration for individuals with autism spectrum disorder (ASD) and typical development (TD). The data are presented as mean ± SEM.
directional hypotheses, all $p$ values are one-tailed. To test whether the effect sizes were significantly different from one another we calculated $z$ scores for the difference between Verbal and Visuospatial, $z = 0.66$; Verbal and Multimodal, $z = 0.68$; and Visuospatial and Multimodal, $z = 0.04$. These show that although the effect size difference is much smaller for Visuospatial versus Multimodal, all $z$ scores are below 1.96 and therefore are not statistically significant from one another.

Following Bowler et al. [2014], we examined the effect of CTT interference on the performance on the span tasks by including CTT interference as a covariate in the analysis reported above. The resulting 2 (Group) × 3 (Task: Verbal, Visuospatial, Multimodal) repeated measures ANCOVA left the results reported above unchanged. We found the significant main effects of Group, $F(1,96) = 5.25$, $p < .05$, $\eta^2_p = 0.05$, and Task, $F(2,192) = 38.49$, $p < .000$, $\eta^2_p = 0.29$, and no interaction, $F(2,192) = 0.78$, $p = .46$, $\eta^2_p = 0.02$. There was also no interaction between Task and CTT interference, $F(2,192) = 1.89$, $p = .15$, $\eta^2_p = 0.02$. Thus, the ASD group performed worse on all three STM tasks and this effect went beyond the influence of EF.

Integration Task

Corrected accuracy scores, set out in Figure 3, showed a higher performance for the TD ($M = 0.38$, $SD = 0.17$) compared to the ASD group ($M = 0.29$, $SD = 0.24$), a difference which was only marginally significant, $t = 1.79$, $p = .078$, one-tailed) with a small effect size (Cohen’s $d = 0.44$).$^6$

Effects of Chronological Age and Executive Functions on Span and Integration Accuracy

To investigate the effects of chronological age and CTT Interference on STM span and integration, we first calculated bivariate Pearson correlations between chronological age, CTT Interference, and corrected accuracy on the three STM span tasks as well as corrected accuracy on the Integration task. These data are set out in Table 2. There were significant positive correlations among all tasks for the sample as a whole as well as the TD and ASD groups separately. For the sample as a whole we found significant negative correlations between chronological age and corrected accuracy on the visuospatial and the multimodal tasks (those memory measures that could be argued to be more complex), indicating a decrease in visuospatial and multimodal STM with increasing chronological age. These correlations seem to have been driven by TD performance as they only remained for the TD but not the ASD group. Similarly, there were significant negative correlations between CTT Interference and

\[ \text{Table 2. Bivariate Correlations Among Corrected Accuracy on the Three Span Tasks (Verbal, Visuospatial, and Multimodal) and the Integration Task, Age, and the Color Trails Interference Score for the Group as a Whole As Well As individuals with Autism Spectrum Disorder (ASD) and Typical Development (TD) Separately.} \]

<table>
<thead>
<tr>
<th>TD</th>
<th>Age</th>
<th>CT</th>
<th>Verb</th>
<th>VS</th>
<th>MM</th>
<th>IG</th>
<th>Age</th>
<th>CT</th>
<th>Verb</th>
<th>VS</th>
<th>MM</th>
<th>IG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0.12</td>
<td>0.34</td>
<td>0.64</td>
<td>0.55</td>
<td>0.35</td>
<td>1</td>
<td>0.04</td>
<td>0.33</td>
<td>0.60</td>
<td>0.23</td>
<td>1</td>
</tr>
<tr>
<td>ASD</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.03</td>
<td>0.14</td>
<td>0.08</td>
<td>1</td>
<td>0.33</td>
<td>0.33</td>
<td>0.45</td>
<td>0.35</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.06</td>
<td>0.09</td>
<td>0.10</td>
<td>0.20</td>
<td>0.15</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
</tr>
</tbody>
</table>

$^*$ $p < .1; ^{**}p < .05; ^{***}p < .01$.

$^6$ Rerunning this analysis with CTT interference as a covariate left the direction of the effects unchanged.
corrected accuracy on the visuospatial and the multimodal tasks (again the memory measures that could be argued to be more complex), indicating a decrease in visuospatial and multimodal STM with increasing interference on the CTT. These correlations only held for the ASD but not the TD group when groups were analyzed separately. Chronological age did not correlate with CTT interference for either group.

Because of the lack of correlations between corrected accuracy on verbal span and integration task with chronological age and CTT Integration, the following analyses focus on visuospatial and multimodal span performance. To explore the correlations reported above in greater depth, we conducted two separate sets of stepwise linear regression analyses. We used corrected accuracy on the Visuospatial and Multimodal tests as dependent variables and entered Chronological age, Group, and a Chronological age × Group interaction term in the first set and CTT, Group, and a CTT × Group interaction term for the second set of analyses.

Regarding our first set of analyses, entering Chronological age, Group, and a Chronological age × Group interaction term to predict Visuospatial STM, the Chronological age × Group interaction term significantly explained 6.2% of the total variance, $R^2 = 0.062$, 95%, $F(1,104) = 6.83$, $p = .01$ and it remained as the only significant predictor of Visuospatial STM ($\beta = -0.25$, $p < .01$). Similar results were found for Multimodal STM. Chronological age × Group interaction significantly explained 6.2% of the total variance, $R^2 = 0.062$, 95%, $F(1,104) = 6.83$, $p = .01$ and it remained as the only significant predictor of Multimodal STM ($\beta = -0.25$, $p = .01$). These results indicate that age had a differential effect on both groups which is presented in Figure 4. Inspection of Figure 4 shows that there was a stronger age-related effect on corrected accuracy on Visuospatial and Multimodal STM in the TD compared to the ASD group.

Regarding our second set of analyses, entering CTT Interference, Group, and a CTT Interference × Group interaction term to predict Visuospatial STM, the CTT Interference × Group interaction term significantly explained 8.2% of the total variance, $R^2 = 0.062$, 95%, $F(1,98) = 6.86$, $p = .004$ and it remained as the only significant predictor of visuospatial STM ($\beta = -0.29$, $p < .01$). Similar results were found for Multimodal STM. CTT Interference × Group interaction significantly explained 14.9% of the total variance, $R^2 = 0.149$, 95%, $F(1,98) = 16.97$, $p = .000$ and it remained as the only significant predictor of multimodal STM ($\beta = -0.39$, $p = .000$). These results indicate that CTT Interference had a differential effect on both groups which is presented in Figure 5. Inspection of Figure 5 showed that higher CTT interference (i.e., poorer EF) was associated with a greater decrement in performance in the ASD than in the TD participants.

**Discussion**

Three aspects of the data reported above address the first aim of the study, which was whether ASD participants performed similarly to typical older adults on the four tasks used here. The significantly poorer performance by the ASD participants on the three Span tasks and the Integration task shown in Figure 3 confirm that they experienced greater difficulty than the TD group on these types of STM tasks. The results of the multiple regression analyses illustrated Figure 4 also show that the ASD participants at all ages tended to have similar Visuospatial and Multimodal corrected task accuracy to that of the older TD participants. These three observations confirm both Lecouvey et al.’s [2015] speculation that autistic individuals would perform similarly to healthy older participants, and Bowler et al.’s [2004] aging analogy, which makes a similar prediction. On all tasks, while for the TD group, older age was associated with diminished performance on all memory measures, this was not the case for the ASD participants. This mirrors findings from some previous studies [Lever & Geurts, 2016; Lever et al., 2015; Ring et al., 2016] but not others [Geurts & Vissers, 2012; Powell et al., 2017]. For example, Lever et al. [2015] report no age-related differences in an n-back WM task in ASD adults and Lever and Geurts [2016] found a smaller effect of age on visual memory in ASD compared to TD adults. Geurts and Vissers [2012], by contrast found greater age-related performance differences in ASD than in TD on immediate recall from visual memory but not in verbal memory leading them to suggest that since some cognitive difficulties (such as with verbal memory, planning and fluency) appear to reduce with age, growing older might be a protective factor in these domains in individuals with ASD. Although it is tempting to draw a similar conclusion from the findings of the present study, an alternative conclusion might be in terms of recruitment of compensatory neural mechanisms as was found, for example, by Baxter et al. [2019].

The lower performance of the ASD group on the three span tasks also replicates existing studies documenting diminished short-term recall in ASD [Alloway et al., 2016; Bowler et al., 2016; Christ et al., 2017; Poirier et al., 2011] and support the second aim of the present study, which was to provide further evidence that the well-documented ASD-related LTM difficulties [Desaunay et al., 2020; Boucher et al., 2012; Boucher & Bowler, 2008; Ben Shalom, 2003; Minshew & Goldstein, 1993] extend to STM span. However, the lack of a significant Group by Task interaction and the overlapping effect size confidence intervals for the three span tasks do not support the hypothesis that ASD individuals experience greater difficulty on more complex memory tasks [Minshew & Goldstein, 1998; Minshew et al., 2000; Williams et al., 2006a, 2006b]. Although the
S a m p l e s i z e w a s g u i d e d b y p r e v i o u s r e s e a r c h o n t h e t o p i c [Lecouvey et al., 2015; Quinette et al., 2006] and is much larger than in most memory studies, a power calculation (G*Power) [Faul, Erdfelder, Lang, & Buchner, 2007] showed that to detect a significant Group × Task interaction with an effect size of \( f = 0.07 \) and a statistical power of 0.90, a total sample size of 422 participants would be needed.

In addition to low statistical power, the conflicting conclusions and discrepant findings of earlier studies of memory in older autistic adults may also result from inherent heterogeneity in the ASD population [see Waterhouse, 2013]. In addition, meta-analyses often aggregate studies under headings such as “verbal” or “visuospatial” that have quite different methodological features which can present contrasting processing challenges to typical and atypical participants. For example, the discrepancy between Geurts and Vissers’ [2012] finding of no age-related difference in verbal memory in ASD and Powell et al.’s [2017] finding of an age-related difference may simply be the result of selecting small samples \( (n < 30) \) from a large, heterogeneous population that also ages in diverse ways. For example, in Howlin, Savage, Moss, Tempier, and Rutter’s [2014] longitudinal research, cognitive functions of most individuals remained stable, quite a large percentage of individuals (25%) showed a steep decline in cognitive functions. Other factors such as the age-range of participants might also be relevant. In the case of Geurts and Vissers [2012], the age range was rather small with 63.6 ±7.5 years compared to Powell et al.’s [2017] participants who ranged in age from 30 to 67 years or those of the current study (age range: 21–67 years). Furthermore, the age of diagnosis might also play a role in the differences in findings between different studies. Happé et al. [2016] suggested recently that in individuals diagnosed in later life, a life of trying to cope with autism might mitigate the effect of the symptoms. Unfortunately, we do not have reliable information concerning age of diagnosis for the ASD participants included in this study. However, this could be done in future studies to determine whether these findings may be a cohort effect.

The contrast between our finding of no association for the ASD group between age and the visuospatial span...
task and Geurts and Vissers’ [2012] finding of an age-related difference in visual memory for their ASD may have resulted from differences between two tasks. Geurts and Vissers’ [2012] task was the Visual Reproduction sub-test of the Wechsler Memory Scale [Wechsler, 1987] and involved participants having to study a set of geometrical shapes and reproduce them after varying delays, whereas the visuo-spatial task used here required participants to recall a series of strings of locations of the letter “x” in a grid. Although both tasks are tests of “visual memory” or “visuospatial memory” they place quite different demands on participants, which might be even more different for a neuropsychologically atypical group such as ASD. This point echoes Mottron, Dawson, and Soulières’ [2008] observation that we should be careful when analyzing autistic psychological functioning on the basis of categories developed in the context of neurotypical cognition.

Our hypothesis, which was to test whether autistic participants experienced the same difficulty on the Multimodal Integration task as did Lecouvey et al.’s healthy older adults received little support. This may have resulted from lower power and small sample size but in retrospect, and from a more theoretical perspective, although the Integration task can be thought of as a complex procedure, requiring different elements of experience to be held in memory for a short time, it can also be conceived as a series of single-trial, cued recall tests, thereby contrasting with the span tasks where on each trial, participants had to hold in memory sequences of items ranging in complexity from that of a single modality (letters and spatial locations) in the Verbal and Visuo-spatial tasks to combinations of these in the Multimodal task. In addition, the cued recall aspect of the Integration task makes the relatively good performance of the ASD participants not too surprising, since cued recall is an area of strength for this population [Boucher et al., 2012; Desaunay et al., 2020]. Yet the binding of letters to colors and locations makes the Integration resemble other binding tasks that both ASD and older typical participants find difficult [Chalfonte & Johnson, 1996; Bowler et al., 2014].
One difference between the Integration task and that of Bowler et al. [2014] and Chalfonte and Johnson [1996] is that whereas the latter task required participants to recognize a large set of previously studied items at test, the Integration task involved a single yes/no decision in response to a cue on each trial. Further studies manipulating the relational structure of the studied material and the inferences required in the test procedure should help to clarify the discrepancy between the findings reported here and those of earlier studies. Such studies should help to shed further light on which processes individuals on the autism spectrum find easy and which they find difficult. In recent years, the concept of binding in memory has undergone considerable refinement and elaboration [Allen, 2015] covering such phenomena as conjunctive (or intra-item) binding and relational (or inter-item) binding. Age-related decline in typical older individuals is generally only seen in relational binding [Kirmse et al., 2018]. Although investigations of both types of binding have been carried out with ASD individuals [Bowler et al., 2014; Loth, Gómez, & Happé, 2011; Massand & Bowler, 2015; Solomon, Frank, Smith, Ly, & Carter, 2011; Stevenson et al., 2019], often with mixed results, to the best of our knowledge, no studies have systematically explored different kinds of binding and their correlates with an autistic population.

Regarding the underlying mechanisms, our exploration of the effects of EF on performance on the visuospatial and multimodal tasks set out in Figure 5 show that although ASD individuals’ performance on the tasks, unlike that of the TD comparison participants, was not adversely affected by older age, it did appear to rely more on executive processes (see Fig. 5). In this respect, the ASD participants contrast with the TD participants, whose task performance in the visuospatial and multimodal tasks did not appear to rely on EFs. This compensation through reliance on EFs by the ASD participants echoes a long-standing speculation that people with autism perform tasks using different mechanisms from those used by neurotypical comparison participants [Hermelin & O’Connor, 1985; Bowler, 1992; Happé, 1995; Livingston & Happé, 2017]. However, small $R^2$ values and the fact that the differences in all three STM span tasks remained when covarying for CTT Interference scores suggest that difficulties with EF are not the only relevant factor. These speculations have been borne out by a recent study by Hogeveen et al. [2020] demonstrating atypical fronto-hippocampal neural mechanisms underlying undiminished relational processing in ASD. Disentagling the effects of different processes is a task for future research. One relevant factor is the relation between age and EF. In the current study, age did not correlate with CTT Interference, however, in a previous study, Abbott, Happé, and Charlton [2018] found a significant relation between CTT performance and age. Another relevant factor and direction for future research could be the differentiation between different types of errors on the STM span tasks whereby errors in the sequence could indicate executive dysfunction as opposed to intrusion errors in the letters which could indicate source memory difficulties.

To conclude, the current study is the first to systematically investigate STM relational binding and integration across different modalities in a sample of autistic adults with intelligence within the normal range. We found parallels with typical aging and with previous findings of difficulties in relational binding in LTM. In addition, different processes to tackle the task as well as differential reliance on EFs between the two groups seem to have played an important role in task performance. More research is needed to study the evolution of short-term relational binding across the life span as well as across the full range of verbal and intellectual disability.

Acknowledgments

We would like to thank all participants for taking part in this research. There are no conflicts of interest for any of the authors. Melanie Ring was supported by a Doctoral Research Studentship from City, University of London. Dermot Bowler was supported by a Chaire d’Excellence from the European Commission and the Conseil regional de Basse-Normandie. Open access funding enabled and organized by Projekt DEAL.

References


