Abstract

Domain-general theories of autism rest on evidence that the disorder impacts not only social communication skills but also nonsocial functions such as memory. Yet recognition memory deficits have been inconsistently documented, especially for stimuli other than faces and sentences. Here we tested school-aged children with high-functioning autism (ASD) and IQ and age-matched comparison children on a visual long-term memory task involving more than 100 photographs of objects, faces, cats, houses, and abstract stimuli. Children viewed each photograph for 2 seconds. After a 10-minutes filled delay we assessed recognition memory for object category as well as for specific exemplars. Data supported the presence of a high-capacity and high-precision visual memory in children with ASD. Both category memory and exemplar memory accuracies were above 90% for categories for which a single exemplar had been encoded. When more exemplars per category were encoded, category memory improved but exemplar memory declined. An exception was face memory, which remained highly accurate even after many faces had been encoded. Our study provided no evidence that visual memory in general, and face memory in particular, is impaired in children with ASD.

Keywords: autism, visual long-term memory, face memory

Introduction

Our world is organized into categories containing semantically coherent and visually similar objects. Critical for survival are the ability to make inferences based on an object’s category and the ability to differentiate individual objects within a category. Visual long-term memory is structured to support both general categorization and the recognition of specific exemplars. On the one hand, humans can rapidly extract the gist of a scene (Li, VanRullen, Koch, & Perona, 2002; Thorpe, Fize, & Marlot, 1996), categorize a briefly presented object (Grill-Spector & Kanwisher, 2005), and retain information about previously viewed pictures (Standing, 1973). On the other hand, we can also remember the visual details of individual objects (Brady, Konkle, Alvarez, & Oliva, 2008), faces (Bruck, Cavanagh, & Ceci, 1991), and scenes (Konkle, Brady, Alvarez, & Oliva, 2010b). The highly structured and detailed visual memory facilitates the adaptation to social and nonsocial environments. It aids the recognition of people and the recall of objects that belong to different people or context. Impairment in visual long-term memory is likely to cause difficulties in social interaction and daily functioning. In this study we characterized the capacity and precision of visual long-term memory in children with autism spectrum disorder (ASD, or “autism”). We asked two questions: (i) Is autism associated with reduced capacity and/or less precision in visual memory, and (ii) Is the memory impairment exacerbated by the social nature of the stimuli?
Autism spectrum disorder is a neurodevelopmental disorder that affects about 1 in every 88 children (Centers for Disease Control, 2012). Children with ASD show impaired social and communication skills and often have restricted interest or repetitive behaviors. Although primarily a disorder in social communication, ASD also frequently affects nonsocial functions. In the domain of memory, recent empirical work and comprehensive reviews have revealed deficits in some tasks and for some stimuli. First, memory for social stimuli is often, but not always, impaired. For example, some studies found poorer face memory in children with ASD relative to age- and IQ-matched typically developing children (for recent reviews, see Weigelt, Koldewyn, & Kanwisher, 2012). In some cases the face memory deficit is greater than that for nonface stimuli (e.g., Weigelt, Koldewyn, & Kanwisher, 2013), while in other cases it applies equally to nonface stimuli such as cars and inverted faces (e.g., Ewing, Pellicano, & Rhodes, 2013; Scherf, Behrmann, Minshew, & Luna, 2008). Second, memory for nonsocial stimuli is largely intact, though subtle impairment has been found in memory for sentences, semantically related words, and self-referencing activities (for a review, see Boucher, Mayes, & Bigham, 2012). When visual stimuli are used, children with ASD are often, though not always, impaired in visual working memory (Jiang, Capistrano, & Palm, 2014; Ozono & Strayer, 2001; Williams, Goldstein, Carpenter, & Minshew, 2005). Because working memory is considered the activated portion of long-term memory (Larocque, Lewis-Peacock, & Postle, 2014), the deficit may also extend to visual long-term memory. However, empirical findings on whether visual long-term memory for nonface stimuli is impaired have been mixed (for recent discussions, see Ewing et al., 2013; Weigelt et al., 2013).

This brief literature review identifies two gaps in previous research. First, few studies have systematically investigated visual long-term memory for nonface stimuli. When these stimuli were tested, their primary utility was to serve as a control for faces. Exceptions are noted in studies that showed intact recognition memory for doors (Ambery, Russell, Perry, Morris, & Murphy, 2006), line drawings (Renner, Klinger, & Klinger, 2000), or meaningless shapes (Bigham, Boucher, Mayes, & Anns, 2010). Yet these studies tested either impoverished stimuli (such as line drawings) or very few categories (e.g., doors). One study assessed recognition memory for faces, cats, horses, motorbikes, buildings, and leaves (Blair, Frith, Smith, Abell, & Cipolotti, 2002). Unfortunately, participants in that study performed different encoding tasks on different categories. Furthermore, the number of encoded exemplars varied across categories, making it difficult to directly compare memory performance across stimuli. Thus, it is unclear whether children with ASD show impaired visual long-term memory for a broad range of objects. A satisfactory answer to this question requires separate assessment of memory for object categories and memory for specific exemplars. This approach has only recently been undertaken in research on typical adults (Brady et al., 2008; Hollingworth, Williams, & Henderson, 2001; Konkle, Brady, Alvarez, & Oliva, 2010a; Koutstaal et al., 2003) and has not been applied to examining visual memory in people with autism (for a manipulation of category and item memory for words, see Gaigg, Gardiner, & Bowler, 2008).

A second issue in the literature surrounds the extent and specificity of a memory deficit for faces in individuals with ASD. Using the Cambridge Face Memory test (Duchaine & Nakayama, 2006), one study found that adults with ASD were just as impaired as patients with prosopagnosia (O’Hearn, Schroer, Minshew, & Luna, 2010). In other studies, the face recognition deficit was significant but small (e.g., Blair et al., 2002; Ewing et al., 2013), and still other studies did not find consistent face recognition impairments (for a review, see Jemel, Mottron, & Dawson, 2006). The highly diverse findings may partly be attributed to the use of different face stimuli and tasks. Most experimentally rigorous tests of face memory have used computer-manipulated gray-scale images, such as morphed faces, oval-shaped faces without external features, and faces presented within noise or after changes of pose. Yet the purity of experimental testing comes at a cost of ecological validity. Removing the inner or outer features of a face is known to impair face recognition in both children with ASD and typically developing children (Wilson, Pascalis, & Blades, 2007). To assess whether face memory deficits pose significant challenges in daily functioning, it is important to evaluate face memory using
photographs rather than computer-manipulated images. Although some previous studies have used photographs (e.g., Boucher, Lewis, & Collis, 1998), they have not contrasted faces with other stimuli.

To address the issues identified above, we tested visual long-term memory in children with ASD using photographs of a large number of stimuli. Children encoded a series of 128 photographs for a subsequent recognition test. Half of the photographs were common household objects and the other half were stimuli from special categories. First, to investigate the capacity and precision of memory for visual objects, we manipulated the number of exemplars per category that children had to remember, as well as the nature of the recognition test. In the encoding phase, children saw a varied number of exemplars (1, 2, 4, or 16) from a given category. In the testing phase, the foils could be an object from a new category or a new exemplar from the same category as an object encoded before (Brady et al., 2008; Konkle et al., 2010a, 2010b). These manipulations allowed us to test (a) visual recognition memory for object category versus specific exemplars, and (b) effects of number of exemplars on visual memory.

Second, to test the extent and specificity of recognition memory deficits for faces, we included 16 exemplars from each of five categories: faces, cats, houses, fractals, and household objects. These stimuli were included to test whether autism is associated with a deficit in recognition memory for social stimuli, and whether the deficit extends to animate categories of stimuli such as cats (Blair et al., 2002). Our special categories included fractals because they had no prior semantic associations and were difficult to name. If memory deficits in autism have a semantic origin, then the deficits should be reduced when fractals are tested. Importantly, to control for differences in attention across categories, we randomly intermixed all 128 stimuli during encoding and asked participants to perform the same task (i.e., preference rating) on all stimuli.

A general visual memory deficit hypothesis predicts impaired recognition performance on all stimuli in children with ASD. This is expected if children with ASD are less attentive during encoding, if encoding strategies or memory consolidation processes are impaired, or if children with ASD have a lower capacity for remembering a large number of stimuli (Boucher et al., 2012). Alternatively, the visual memory deficit may be specific to social stimuli such as faces (specific memory deficit hypothesis). Because the social nature of visual stimuli falls along a continuum (New et al., 2010), the deficit may extend to animate categories such as cats. Third, increased local processing and/or decreased global processing of visual input may result in enhanced memory for exemplars but impaired memory for category (weak central coherence hypothesis; Happé & Frith, 2006; Jemel et al., 2006). The opposite pattern is expected if information complexity exacerbates visual memory deficits (information complexity hypothesis; Minshew & Goldstein, 2001). In this case, children with ASD should have impaired recognition memory for exemplars but intact recognition memory for category.

**Method**

**Participants**

We tested 40 school-aged children (6-15 years old). Among these, 20 were typically developing and 20 were children with ASD. An additional 3 children with ASD were recruited but failed to understand the instructions. Twenty typically developing adults (18-35 years old) also completed the experiment. The adult group was matched with the child participants on gender (17 males and 3 females) and nonverbal IQ (mean 108, S.D. = 14, range: 80-128).

We obtained written consent from adult participants and parent of child participants and assent from the children. Participants were compensated for their time. The University of Minnesota IRB approved the study protocol.

**Age, gender, and IQ match in child participants**

The two groups of children were matched on age, gender and nonverbal IQ. Table 1 shows the characteristics from the child participants. Because the task involves visual processing and because children with ASD have known communication deficits, the two groups were matched on nonverbal IQ. For typically developing children and adults, nonverbal IQ was obtained from the block design
and matrices subtests of the Wechsler Abbreviated Scale of Intelligence. As part of their diagnosis children with ASD had already received prior IQ testing based on either the Wechsler Intelligence Scale for Children or Differential Ability Scales (including block design, matrices, recall of design, and sequential and quantitative reasoning; Elliott, 2007). IQ testing was not repeated. All IQ instruments are standardized to the same mean (100) and standard deviation (15) and have equivalence scores.

Table 1. Characteristics of children tested in this study. Standard deviation (S.D.) and range are shown in the parenthesis.

<table>
<thead>
<tr>
<th></th>
<th>Children with ASD</th>
<th>Typically developing children</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>18 boys, 2 girls</td>
<td>17 boys, 3 girls</td>
<td></td>
</tr>
<tr>
<td>Age (in years)</td>
<td>11.4 (S.D.=2.5; range: 7-15)</td>
<td>11.5 (S.D.=2.3; range: 7-15)</td>
<td>P &gt; .80</td>
</tr>
<tr>
<td>Nonverbal IQ</td>
<td>106.6 (S.D.=19; [82-149])</td>
<td>108.5 (S.D.=12; [87-128])</td>
<td>P &gt; .70</td>
</tr>
<tr>
<td>SCQ1 (lifetime)</td>
<td>20.5 (S.D.=6; [11-32])</td>
<td>3.2 (S.D.=4; [0-12])</td>
<td>P &lt; .001</td>
</tr>
<tr>
<td>SNAP-IV ADHD2</td>
<td>1.2 (S.D.=0.5; [0.3-2.2])</td>
<td>0.4 (S.D.=0.4;[0-1.5])</td>
<td>P &lt; .001</td>
</tr>
<tr>
<td>ADOS calibrated severity score3</td>
<td>7.5 (S.D.=1.4; [5-10])</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

Table Notes:
1SCQ: The higher the SCQ scores the more severe the social communication deficits (Rutter et al., 2003). The cutoff score for raising concerns about autism is 15.
2SNAP-IV ADHD: The higher the SNAP-IV scores the more severe the ADHD features. Scores could range from 0 (not at all) to 3 (very much).
3ADOS calibrated severity score was obtained from 19 children with ASD. Scores could range from 1 to 10 (1-3: nonspectrum; 4-5: ASD; 6-10: autistic disorder; Gotham, Pickles, & Lord, 2009).

Clinical assessment

Children with ASD were recruited from the University of Minnesota Autism Spectrum and Neurodevelopment Disorders Clinic (AS/NDD Clinic; N=16) or from the community (N=4). Children from the AS/NDD clinic received comprehensive diagnostic evaluations by licensed psychologists with established research reliability, including a diagnostic interview (Autism Diagnostic Interview-Revised or DSM-based diagnosis), the Autism Diagnostic Observation Schedule (ADOS-2; Lord et al., 2000), cognitive tests, and review of medical history. Children from the community received comprehensive evaluations by licensed psychologists from their school districts, including parental interviews, behavior observations, and a team review. Their diagnoses were additionally confirmed with ADOS administered by a psychologist affiliated with the University of Minnesota AS/NDD Clinic, or in one case by medical records provided by an autism treatment clinic. Typically developing children were recruited from the local community. They did not have a history of psychiatric or neurological conditions, as assessed through phone interviews with parents and with the parent questionnaires, see next.

Parent questionnaires

Parents of all children filled out two questionnaires: the Swanson, Nolan, and Pelham-IV (SNAP-IV, Zolotor, Mayer, & Hill, 2004) and the Social Communication Questionnaire (SCQ, (Rutter, Bailey, & Lord, 2003). Because the memory task requires active attention, we obtained a combined ADHD score from SNAP-IV items 1-9 and 11-19 (this included inattention and hyperactivity/impulsivity). The SCQ was used to ensure that the typically developing children did not meet cutoff for ASD (i.e., typical scores should be under 15). Table 1 lists the SNAP-IV and SCQ scores.

Equipment and materials
Participants were tested individually on a 15” MacBook laptop in a quiet room either at their home or in our lab. Viewing distance was unconstrained but was approximately 40 cm. The experiment was programmed with Psychtoolbox (Brainard, 1997; Pelli, 1997) implemented in MATLAB (www.mathworks.com). An eye chart verified that all participants had normal or corrected-to-normal visual acuity.

Photographs of household objects were taken primarily from Tim Brady and Aude Oliva’s database (http://cvc.mit.edu/MM/stimuli.html). Object categories that had faces or animals (e.g., dolls, dogs) were removed, leaving us with a database of 184 distinct categories with at least 17 exemplars per category. Additional exemplars were supplemented through online searches. Photographs of special categories were taken from various sources. Faces were colored photographs of children’s faces (same as those used in Fischer, Koldewyn, Jiang, & Kanwisher, 2014). Cats were colored photographs from Brady and Oliva’s database or online sources. They had faces but also differed in pose, fur color, and other features. Houses came from online searches and depicted typical American single-family houses. Fractals were from online searches and contained visually distinct but abstract patterns. A separate set of 20 pictures was used for practice. All images were resized to 500x500 pixels.

Procedure

Practice. Participants were told that they would see a lot of pictures and that their memory would be tested later. To gain familiarity with the procedure, they first completed a practice session that included a series of 5 pictures during the encoding phase, followed immediately by 5 test trials. The procedure used for practice was identical to that of the main experiment (see next).

Memory encoding. Participants clicked on a “Go” button (1.6°x1.6°) in the middle of the screen to start each encoding trial. Upon their mouse click a picture (16°x16°) was presented for 2 seconds and then replaced by a preference rating display (Figure 1). Participants were asked to remember the picture and rate “whether they liked the picture” by clicking one of three response buttons labeled “YES” (in green), “OK” (in yellow), and “NO” (in red) on the rating display (each button subtended 4.7°x3.6°). To ensure that children could not mindlessly click through the experiment, the response buttons were displayed at the bottom of the monitor (6.5° below the center of the display) whereas the “Go” button was in the middle. The mouse clicks were registered only when they fell in the proper locations. The encoding phase contained 128 trials. Each picture was presented once.

Recognition test. The recognition test started approximately 10 minutes after all 128 pictures were encoded. During the delay all participants performed the matrices task from the Wechsler Abbreviated Scale of Intelligence. Mean delay interval did not differ among groups (F < 1). After the delay, participants initiated each of the 128 recognition trials by clicking on the “Go” button in the middle of the screen. They were then shown four pictures, one in each visual quadrant (Figure 1; each picture subtended 8.9°x8.9°). The four pictures included a picture they saw in the encoding phase (old image), a new exemplar from the same category as the old image (the within-category foil), and two exemplars from a new category not shown in the encoding phase (the between-category foils). The locations of the four choices were randomized. Participants were asked to click on the picture they had seen earlier in the experiment. Note that we included the within-category and between-category foils on the same test trial. Compared with testing the within- and between- category foils on separate trials (e.g., Brady et al., 2008), our procedure was more efficient because it assessed both types of memory errors for every encoding image. Once participants made the mouse click, the display was erased and a tone indicated whether they had chosen the right picture. We also displayed the total number of correct responses. The 128 trials of the recognition test corresponded to the 128 stimuli that children encoded earlier, although the order of trials was randomized. No images were shown more than once in the testing phase.
Visual memory and autism

Jiang et al.

6

Figure 1. A schematic illustration of stimuli and procedures used in this study. Left: Encoding phase. Participants clicked the “Go” button to initiate a trial. An object was shown for 2 s and then replaced by the preference rating display. Right: The recognition test was administered 10 min after all 128 images had been encoded. Each trial of the recognition test included an old object, a new object from the same category as the old object, and two objects from a new category not shown before.

Design

The 128 images used in the encoding phase came from the following conditions. First, for household objects (henceforth “objects”; 64 images), we manipulated the number of encoded exemplars from a given category. The exemplar numbers were 1, 2, 4, or 16. The 16-exemplars-per-category condition contained 16 stimuli from a single category (e.g., 16 backpacks), the 4-exemplars-per-category condition contained 16 stimuli from four separate categories (e.g., 4 audio speakers, 4 suitcases), the 2-exemplars-per-category condition contained 16 stimuli from eight separate categories (e.g., 2 hourglass, 2 cell phones), and the 1-exemplar-per-category condition contained 16 stimuli from 16 different categories (e.g., 1 wind chime, 1 headphone). The household objects allowed us to estimate memory capacity and precision for categories of stimuli that varied in the number of exemplars presented to children.

Second, for special categories (64 images), we included 16 exemplars from each of the following categories: faces, cats, houses, and fractals. A comparison of these categories, along with the household objects (the 16-exemplars-per-category condition), allowed us to test recognition memory for different categories while holding the number of encoded exemplars constant.

To ensure that any differences between participant groups and any individual differences could not be attributed to differences in an item’s memorability, we randomly selected stimuli from the larger stimulus database (see materials). The same randomization was used for all participants, ensuring that all participants were exposed to the same stimuli and conditions (i.e., the different numbers of exemplars were not rotated across different sub-categories). However, the order of trials and the locations of the test choices were randomized for each participant individually. Appendix Table 1 lists the object categories used in our study.

Results

Our main question is how autism affects visual memory. To address this question, we first present data from the two groups of child participants (typical and ASD). Data from typically developing adults are included in a later section on “developmental change.”

1. Visual long-term memory for objects
between group (ASD or TD) and exemplar disruptiveness was largely similar between ASD and TD children, 25% of error in the condition and 32% of error in the category, indicating superior category memory. Notably, error type interacted with exemplar number, \( F(4, 114) = 34.07, p < .001, \eta^2_p = .47 \). An analysis of within-category errors showed that error rates increased when more exemplars were encoded, \( F(3, 114) = 37.12, p < .001, \eta^2_p = .49 \). The opposite was true for between-category errors: the more exemplars were encoded the better the category memory, \( F(3, 114) = 5.60, p < .001, \eta^2_p = .13 \). (These statistical significance levels survived correction for multiple comparisons). The pattern of rising within-category error and falling between-category error was seen in both groups of children, yielding a lack of three-way interaction between exemplar number, error type, and group, \( F < 1 \). Performance between the two groups of children differed in one regard: exemplar number interacted significantly with group, \( F(3, 114) = 3.31, p = .023, \eta^2_p = .08 \). This interaction was attributed to a steeper increase in error rates with increasing exemplar number in typically developing children than in children with ASD. When both types of errors were summed together, typically developing children committed 11% of error in the one-exemplar-per-category condition and 32% of error in the 16-exemplars-per-category condition. This was a 22% increase in error. In contrast, children with ASD committed 16% of error in the one-exemplar-per-category condition and 25% of error in the 16-exemplars-per-category condition, an increase of just 9%. Thus, although performance was largely similar between ASD and TD children, increasing exemplar number was less disruptive to performance in children with ASD, \( F(1, 38) = 4.75, p < .06, \eta^2_p = .11 \) on the interaction between group (ASD or TD) and exemplar number (1 or 16).

Figure 2 plots recognition memory errors for household objects, separately for the two groups of children and for within-category and between-category memory errors. The two groups of children showed remarkably similar levels of performance. They rarely selected a foil from a new category not seen before ("between-category errors"). Within-category errors were also low, especially for object categories for which a single exemplar had been shown.

To quantify memory performance, we entered exemplar number (1, 2, 4, or 16 exemplars per category) and memory error type (within-category or between-category errors) as within-subject factors, and group (ASD or TD children) as a between-subject factor in a repeated-measures ANOVA. The main effect of group was not significant, \( F < 1 \), indicating that the two groups of children had comparable levels of performance. The main effect of exemplar number was significant, \( F(3, 114) = 18.28, p < .001, \eta^2_p = .33 \), accompanied by a significant linear trend of exemplar number, \( F(1, 38) = 27.76, p < .001, \eta^2_p = .42 \). This finding indicated that memory errors increased when more exemplars were within a category rather than between categories, \( F(1, 38) = 60.70, p < .001, \eta^2_p = .62 \). Even though each display presented two between-category foils and just one within-category foil, participants were less likely to choose the between-category foils than the within-category foil, demonstrating superior category memory. Notably, error type interacted with exemplar number, \( F(3, 114) = 34.07, p < .001, \eta^2_p = .47 \). An analysis of within-category errors showed that error rates increased when more exemplars were encoded, \( F(3, 114) = 37.12, p < .001, \eta^2_p = .49 \). The opposite was true for between-category errors: the more exemplars were encoded the better the category memory, \( F(3, 114) = 5.60, p < .001, \eta^2_p = .13 \). (These statistical significance levels survived correction for multiple comparisons). The pattern of rising within-category error and falling between-category error was seen in both groups of children, yielding a lack of three-way interaction between exemplar number, error type, and group, \( F < 1 \). Performance between the two groups of children differed in one regard: exemplar number interacted significantly with group, \( F(3, 114) = 3.31, p = .023, \eta^2_p = .08 \). This interaction was attributed to a steeper increase in error rates with increasing exemplar number in typically developing children than in children with ASD. When both types of errors were summed together, typically developing children committed 11% of error in the one-exemplar-per-category condition and 32% of error in the 16-exemplars-per-category condition. This was a 22% increase in error. In contrast, children with ASD committed 16% of error in the one-exemplar-per-category condition and 25% of error in the 16-exemplars-per-category condition, an increase of just 9%. Thus, although performance was largely similar between ASD and TD children, increasing exemplar number was less disruptive to performance in children with ASD, \( F(1, 38) = 4.75, p < .06, \eta^2_p = .11 \) on the interaction between group (ASD or TD) and exemplar number (1 or 16).
Performance in the two groups of children was remarkably high considering the nature of the recognition test. Between-category errors were low; even in the one-exemplar-per-category condition participants rarely chose the wrong category. This finding showed that a single exposure to a visual object was sufficient to yield highly accurate memory for its category. Even more impressive was the surprisingly low within-category error rates. The one-exemplar-per-category condition yielded lower than 10% of within-category errors, comparable to results from typical adults (see “Developmental change”).

Increasing the number of exemplars per category strengthened category memory but impaired exemplar memory. This finding was consistent with a previous study on typical adults (Konkle et al., 2010a). Yet effects of exemplar number were by no means catastrophic. After seeing 16 objects from a specific category, children were able to select an old exemplar at 70-80% accuracy. These data provided strong evidence for the presence of a highly precise and detailed visual long-term memory in typically developing children as well as in children with ASD.

2. Visual long-term memory for special categories

An important feature of our experimental design was the control for the number of encoded exemplars for the five special categories tested. Sixteen exemplars were encoded in each of five categories: faces, cats, objects, houses, and fractals. Differences in exemplar number, therefore, could not account for differential recognition memory impairments in children with ASD (unlike Blair et al., 2002).

As shown in Figure 3, however, our data did not reveal any recognition memory impairment in children with ASD. An ANOVA on category (faces, cats, objects, houses, and fractals), error type (within-category or between-category errors), and group (ASD or TD) showed no effects of group, $F < 1$, group by category interaction, $F(4, 152) = 1.19, p > .30$, group by error type interaction, $F < 1$, or group by category by error type interaction, $F(4, 152) = 1.04, p > .35$. The main effect of category was significant, $F(4, 152) = 13.53, p < .001, \eta^2 = .26$, indicating that recognition performance varied across different categories. The main effect of error type was significant, driven by much higher within-category than between-category errors, $F(1, 38) = 169.61, p < .001, \eta^2 = .82$. A significant category by error type interaction $[F(4, 152) = 13.87, p < .001, \eta^2 = .27]$ was driven by uniformly low between-category errors across all five categories, $F < 1$, but variable within-category errors, $F(4, 152) = 16.04, p < .001, \eta^2 = .30$.

To evaluate recognition memory differences among the five categories, we performed seven planned $t$-tests on within-category errors. A Bonferroni correction for multiple comparisons was used to hold the critical alpha level at $p < .007$. First, using household objects as the baseline, we performed four $t$-tests comparing objects with each of the other four categories. This analysis showed just one
significant difference: Face memory was significantly better than object memory, $t(39) = 7.31$, $p < .001$. Cats, fractals, and houses produced comparable memory to objects, all $ps > .03$. Next, using faces as the baseline, we performed three $t$-tests to examine whether face memory was superior to that of cats, fractals, and houses. This was indeed the case, all $ts(39) > 6.08$, $ps < .001$. Note that this analysis combined data across all children (as justified by a lack of interaction with group). Separately analyzing TD and ASD children revealed the same pattern of statistical results. Thus, for both groups of children, recognition memory was better for faces than for all other categories. In addition, the memory advantage for faces did not extend to another animate category (i.e., cats), another special domain (i.e., houses), or visually distinct but abstract patterns (i.e., fractals).

In a further analysis, we showed that within-category errors for cats, houses, and fractals were all higher than that for the one-exemplar-per-category objects ($p < .001$ in TD, ASD, or both groups combined). This was likely driven by the difference in exemplar number. Notably, although the number of exemplars was large for faces, face memory was as good as memory for the one-exemplar-per-category objects, $t(39) = 1.23$, $p > .20$. Thus, face memory appeared to be resistant to the detrimental effect of increasing exemplar number. This was true for both typically developing children ($p > .20$ comparing faces with one-object-per-category objects) and children with ASD ($p > .50$).

3. Effects of IQ, autism severity, and ADHD characteristics

Owing to the large number of dependent variables but the lack of specific hypotheses, we refrained from selectively reporting correlations among variables. We do note that there was a generally positive correlation across all measures of memory. Children who performed better in one measure (e.g., face exemplar memory) also tended to perform better in other measures (e.g., object/fractal/cat exemplar memory). This was true for both groups of children. Neither age nor IQ accounted for a significant amount of variance, largest $F(1, 36) = 1.98$, $p > .15$.

We also examined the correlation between the calibrated autism severity (Gotham, Pickles, & Lord, 2009) and the various memory measures. Though higher severity generally corresponded to poorer performance, this relationship did not survive correction for multiple comparisons. Additionally, its interpretation was compounded by the significant correlation between autism severity and ADHD severity ($r = 0.47$, $p < .05$). Thus, our study did not clarify the relationship between autism severity and visual memory performance. Future research is needed to address this question by (i) testing a much larger sample, and (ii) examining the correlation in a hypothesis-driven manner.

4. Developmental change

Developmental change in memory, particularly face memory, has been a controversial issue in recent literature. Although many studies have shown an improvement in performance on face recognition tasks from children to adults, it is unclear whether this reflects developmental change in the face recognition system or an improvement in general cognitive capacity (Crookes & McKone, 2009; de Heering, Rossion, & Maurer, 2012). In our study, owing to the small number of children at each age span, we did not observe statistically significant changes in memory performance as a function of age. Splitting the child participants into younger (mean 9.5 years) and older (mean 13.5 years) groups revealed similar patterns of results in the two age groups. The lack of an age effect reflects a limitation in our sample size and the relatively tight age range. Greater statistical power is gained by comparing data from typical adults with typically developing children. Figure 4 shows results from the adults.
The essential statistical results comparing typical adults with typical children are as follows. (1) Category memory was comparable between the two groups. When between-category errors were analyzed, there were no main effect or interactions involving group, all ps > .20. (2) Exemplar memory was better in adults than children. When within-category errors were analyzed, adults showed significantly better performance than children (p < .05). In addition, age group interacted with exemplar number (p < .001). Increasing exemplar number was more detrimental to children's than to adults' memory. Every doubling of exemplar number led to an increase of 2.8% within-category errors in adults, and an increase of 5.9% within-category errors in typical children. Within-category errors also interacted with category (p < .001). Adults had better exemplar memory than children on objects (p < .003) and houses (p < .02), but similar performance to children on faces, cats, and fractals (all ps > .15). These data showed that development in visual long-term memory comes primarily from a greater ability to maintain high memory precision in the face of increasing exemplars. This factor may contribute to the developmental gain in face recognition observed in previous studies (Crookes & McKone, 2009; de Heering et al., 2012). Future studies that test a larger number of children spanning a wider age spectrum are needed to chart out the developmental change in visual memory.

**Discussion**

A major question in autism research is whether autism primarily affects an individual’s ability to communicate and interact with others, or whether it also affects nonsocial functions such as memory. A recent review on high-functioning autism and memory has revealed some, though often subtle, deficits in remembering nonsocial stimuli, such as sentences, semantically related words, or self-referencing activities (Boucher et al., 2012). Evidence for a deficit in face memory is stronger, though inconsistencies exist (Ewing et al., 2013; Jemel et al., 2006; Weigelt et al., 2012). An understanding of memory, particularly visual memory, has been hampered by a lack of systematic investigations into memory for a broad range of objects. Here we tested high-functioning children with ASD and typical children using a large number of photographs covering a wide range of household objects, faces, cats, houses, and abstract images. Importantly, we assessed category memory and exemplar memory in a novel 4-alternative recognition task. Our data showed that children with ASD had a phenomenal recognition memory for a large number of photographs viewed for just 2 seconds each. In addition, we found no evidence that face memory was impaired even when the delay was long (10 minutes) and the number of exemplars encoded was high (16 faces). In what follows we summarize the empirical findings and discuss their implications.

First, with regard to pictures of common objects, children with ASD possessed a visual long-term memory that was both high in capacity and detailed in precision. Recognition memory accuracy for category was above 90% even for categories for which just a single exemplar had been shown. As the number of encoded exemplars per category increased, category memory became more accurate, rising at a rate of about 1-2% for every doubling of exemplars and reaching ceiling after seeing about 4 exemplars per category. The excellent category memory and the sensitivity to increasing exemplar...
number indicated that high-functioning children with ASD had preserved gist (semantic) memory for visual objects. The category memory was supplemented by a highly detailed visual memory for specific exemplars. When presented with two exemplars from an encoded category, one of which was shown before, children with ASD were highly accurate in selecting the old exemplar. Accuracy was higher than 90% for categories for which only a single exemplar had been shown. Every doubling of encoded exemplars for a given category led to an approximately 4% increase in within-category errors in children with ASD. However, even after seeing 16 exemplars they remained accurate in selecting the exemplar(s) that they had seen.

Second, a comparison of face memory with memory for nonface stimuli revealed no evidence that face memory was specifically impaired in autism. Children with ASD performed at comparable levels to typically developing children in recognition tests of faces, cats, houses, objects, and abstract fractals. Though this was a statistically “null” result, it was accompanied by statistically significant differences between faces and other categories of stimuli. Owing to the large number of exemplars (16) per category, children made nearly no between-category errors. Yet, within-category errors rose to about 20-30% for all categories except faces. Although increasing exemplar number was detrimental to exemplar-level memory for other categories, face memory appeared to have escaped this detrimental effect. This was true for typically developing children as well as children with high-functioning ASD.

Both of our main findings appear to contradict some previously reported results. Regarding object memory, one previous study had tested children with ASD on a large number of stimuli, including faces, cats, horses, motorbikes, leaves, and buildings (Blair et al., 2002). The authors concluded that children with ASD had impaired face memory and that this deficit extended to other animate categories. However, as noted earlier participants in Blair et al.’s study engaged in different encoding tasks and were exposed to different numbers of exemplars for the six categories. In addition, conclusions were drawn on the basis of multiple, one-tailed t-tests comparing ASD and comparison groups for each category. Applying a correction for multiple comparisons would have rendered all of the differences between autism and IQ-matched controls statistically insignificant. Like Blair et al.’s study, we also tested recognition memory for a large number of categories. However, we equated the number of exemplars and the encoding task for all categories of stimuli. In addition, the manipulation of exemplar number and the measurement of both category and exemplar memory allowed us to draw conclusions on the basis of statistically significant effects. For example, our data showed convincingly that children with ASD performed at very high levels in recognition tests of objects. In addition, memory precision declined with increasing exemplar number, but this decline was no steeper (in fact, significantly shallower) than that seen in typically developing children. Our study therefore presents one of the most rigorous tests of visual long-term memory for a wide range of nonface stimuli in children with ASD. The empirical data provided compelling evidence for the presence of a high-capacity and high-precision visual long-term memory.

The lack of a face memory deficit may seem inconsistent with prior findings summarized in authoritative reviews (e.g., Weigelt et al., 2012). Yet most previous studies had used gray-scale, computer-manipulated faces. Our study showed that children with ASD had normal face memory for color photographs of children’s faces. Although this finding does not imply that face memory is normal when tested using more arduous stimuli and tasks, it suggests that in children’s daily experience the extent of their face recognition deficit is less than previously thought (O’Hearn et al., 2010). Might the intact performance be attributed to an atypical pattern of face processing? For example, perhaps children with ASD relied on external features such as hairstyle whereas typically developing children relied on holistic processing of internal features. Without experimental testing this possibility cannot be ruled out. However, it seems unlikely that children with ASD succeeded simply by relying on their nonface processing system. This is because if they had done so, then the large number of exemplars in the face category should have led to poor exemplar memory, much like memory for objects, cats, houses, and fractals. Yet face memory was excellent compared with memory for the other categories, suggesting that faces were not processed the same way as the other stimuli.
This discussion underscores the need to further investigate visual memory for a broad range of face stimuli, from gray-scale face morphs to color photographs and videos and even real faces. Such research would need to control for motivation, attention, and interest. Although our finding may be atypical relative to the larger literature, it provides an existence proof that when natural face images are used, children with ASD demonstrate intact memory for faces.

Of the four hypotheses laid out in the introduction, our data did not support the presence of a general deficit in visual long-term memory (Boucher et al., 2012; Ewing et al., 2013). Neither did we find evidence for a specific memory deficit for social stimuli (Blair et al., 2002; Weigelt et al., 2012). Because exemplar memory was not disproportionately more impaired than category memory, the information complexity account also did not apply (Minshew & Goldstein, 2001). Children with ASD appeared less affected by an increase in exemplar number. Relative to typically developing children, their memory was slightly worse for categories with a single exemplar and slightly better for categories with 16 exemplars. The greater tolerance for increasing exemplar number was in line with the weak central coherence hypothesis (Happé & Frith, 2006). However, among all categories with 16 exemplars performance was comparable between ASD and TD children, so the overall pattern of data did not imply a strong bias toward remembering object exemplars.

A domain-general view of autism would predict an across-the-board impairment in cognitive functioning, extending the core social and communication deficits to nonsocial domains. Such broad deficits may be expected on accounts such as a broad disruption of long-range connectivity in the brain. The preservation of a high-capacity, high-precision visual long-term memory does not support these accounts. Instead, our study shows that deficits in social communication coexist with excellent visual memory.

**Study limitations and future directions**

Our study leaves open many questions about autism and visual long-term memory. First, our autism sample included just high-functioning children. Yet memory impairment is more severe in lower functioning children with ASD, not fully accounted for by a reduction in IQ (see Boucher et al.’s excellent review contrasting high- and medium/low functioning ASD). It is important to extend the current study to a wider spectrum of children with ASD.

Second, the presence of an excellent memory in our study does not imply that this memory is properly utilized in the children’s daily functioning. Our study requires intentional encoding of the visual stimuli. Findings may differ if incidental memory is probed. The use of recognition test means that memory performance may be driven by a sense of familiarity rather than recollection. In addition, the preference-rating task may orient participants toward encoding individual photographs rather than relations among categorically similar stimuli (see Gaigg et al., 2008 for a discussion on item-specific and relational memory for words). Results may change if source memory is tested or if the task requires spontaneous processing of the relation among items (Gaigg et al., 2008). It is also important to note that relatively intact memory performance may be achieved by atypical brain processes (Massand, Bowler, Mottron, Hosein, & Jemel, 2013). Finally, we only tested memory for the exact image encoded earlier. Changes in pose, viewpoint, and context may cause difficulties not detected here. In sum, because our study is limited to just one type of memory test using behavioral measures, it will be important in future research to extend our findings to other tests of memory and other forms of memory.

Our emphasis on ecological validity has led us to test visual memory of photographs of real-world objects, faces, and other stimuli. This approach has its own limitation: when comparing across stimulus categories it is not possible to match the intrinsic memorability of stimuli. Superior memory for one category (e.g., faces) relative to another (e.g., fractals) could reflect the presence of special mechanisms that boost memory for the first category, or greater stimulus discriminability for stimuli in the first category. To fully control for stimulus discriminability, it is necessary to use computer-manipulated images such as face morphs or stimuli presented among visual noise. Alternatively, finding evidence for a category by group interaction can alleviate concerns about stimulus
discriminability (e.g., between patients with prosopagnosia and typical controls). Thus, our findings need to be validated in the future using computer manipulated stimuli that match stimulus discriminability across categories.

**Conclusion**

We have tested visual long-term memory of children with high-functioning ASD using a wide range of object categories, faces, and other stimuli. We showed that children with ASD possessed a highly detailed, large capacity visual long-term memory. Increasing the number of encoded exemplars per category led to better category memory but worse exemplar memory. We found no evidence for a specific deficit in face memory in children with ASD. Our results show that autism is not an across-the-board deficit in cognitive functioning. Future research should test visual memory in children with medium or low functioning autism, examine whether other forms of memory are more severely impacted by autism, and compare visual memory for natural and computerized images.

**Acknowledgements**

We are grateful for all the children and their families for participating in this study. We thank Dr. Amy Esler, Dr. Robin Rumsey, Tegan Carr, and the University of Minnesota Autism Clinic for help recruiting the participants, Tegan Carr, Joe Enabnit, Bo-Yeong Won, and Hyejin Lee for their help with testing. Thanks also to Wilma Koutstaal and Sheri Stronach for discussion and Dr. Jeff Epstein and two anonymous reviewers for comments. Correspondence should be directed to Yuhong Jiang, Department of Psychology, University of Minnesota, 75 East River Road, S251 Elliott Hall, Minneapolis, MN 55455. Email: jiang166@umn.edu.

**References**


Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and


Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face recognition deficits in autism spectrum disorders are both domain specific and process specific. *PloS One, 8*(9), e74541. doi:10.1371/journal.pone.0074541


**Appendix**

**A. Categories of objects used in this study (the names were not actually presented).**

From Brady and Oliva’s database (http://cvcl.mit.edu/MM/stimuli.html) and online search we randomly assigned object categories to the different object conditions used in this study. The final randomized assignment of stimuli was as follows. **One-exemplar-per-category:** abacus, bongo, bowl, bread loaf, coin, cook pot, donut, handheld game, head phone, lantern, lock, rock, sofa, stapler, suit, and wind chime. **Two-exemplars-per-category:** cellphone, decorative screen, fan, frame, hourglass, keyboard, mushroom, and water gun. **Four-exemplars-per-category:** lipstick, globe, loudspeaker, and trunk. **Sixteen-exemplars-per-category:** backpack.

Foils used in the testing phase are: apple, baby carrier, bagel, ball, beanbag chair, bed, beer mug, bell, bench, bike, bill, binoculars, bonsai, boots, bowtie, broom, bucket, button, cake, calculator, camcorder, camera, candle, car, ceiling fan, chair, cheese, cheese grater, chessboard, Christmas stockings, Christmas tree ornament, cigarette, click, coat rack, coffee mug, collar, compass, computer,
cookie, cooking pan, cup saucer, cushion, desk, dollhouse, domino, doorknob, dresser, dumbbell, Easter egg, exercise equipment, fish hook, flag, flashlight, Frisbee, garage trash, gift box, gloves, goggles, grill, guitar, hairbrush, hammer, handbag, hanger, hat, head band, helmet, jack-o-lantern, jacket, kayak, lamp, lawnmower, lounge chair, makeup compact, meat, microscope, microwave, motorcycle, mp3 player, nail polish, necklace, necktie, nunchaku, pants, pen, phone, pipe, pitcher, poker cards, radio, razor, record player, ring, ring binder, road sign, roller skates, rosary, rug, saddle salt and pepper shaker, sandwich, scissors, scrunchie, seashell, shoe, spoon, stool, suitcase, swimming suit, teddy bear, table, tape, telescope, tent, toilet seat, tongs, tree, tricycle, trophy, TV, umbrella, vase, video game controller, watch, wig, wineglass, and yarn.