Enhanced Visual Statistical Learning in Adults With Autism

Matthew E. Roser
Plymouth University

Richard N. Aslin
University of Rochester

Rebecca McKenzie and Daniel Zahra
Plymouth University

József Fiser
Central European University

Objective: Individuals with autism spectrum disorder (ASD) are often characterized as having social engagement and language deficiencies, but a sparing of visuospatial processing and short-term memory (STM), with some evidence of supranormal levels of performance in these domains. The present study expanded on this evidence by investigating the observational learning of visuospatial concepts from patterns of covariation across multiple exemplars. Method: Child and adult participants with ASD, and age-matched control participants, viewed multishape arrays composed from a random combination of pairs of shapes that were each positioned in a fixed spatial arrangement. Results: After this passive exposure phase, a posttest revealed that all participant groups could discriminate pairs of shapes with high covariation from randomly paired shapes with low covariation. Moreover, learning these shape-pairs with high covariation was superior in adults with ASD than in age-matched controls, whereas performance in children with ASD was no different than controls. Conclusions: These results extend previous observations of visuospatial enhancement in ASD into the domain of learning, and suggest that enhanced visual statistical learning may have arisen from a sustained bias to attend to local details in complex arrays of visual features.

Keywords: Autism Spectrum Disorder, Asperger’s, statistical learning, visual features

Sensitivity to structure and regularity in the environment is necessary for successful navigation and understanding of a complex world. In the visual domain, extracting invariant features from a cluttered visual array is essential for the recognition of objects (Biederman, 1987; Marr, 1982). For example, the features of a cup consist of a bowl and a handle, and the features of a face consist of the eyes, nose, mouth, and hair. However, the visual features that define many objects are not, as in the case of a face, immediately obvious to a naïve observer. For example, a circular-shaped clock and a wheel share a feature (“roundness”), but it is a defining feature only for the wheel. Thus, in addition to features that are provided by low-level perceptual mechanisms, there must be a learning mechanism that extracts from multiple exemplars embedded in cluttered visual scenes the defining features of objects.

A substantial body of research over the past decade has documented that such a learning mechanism is not only present in adults, but it also operates without feedback about which features are relevant (i.e., unsupervised learning by mere exposure). Moreover, this form of implicit learning (a) operates on input from the auditory, visual, and tactile modalities, (b) acquires features defined in either the temporal or spatial domain, and (c) is robustly present in adults, children, infants, and nonhuman animals. The key property of this learning mechanism is that it establishes associative links between two or more elements in the input (e.g., either across time or space) into meaningful chunks that in turn serve as new features in building higher-order structures. This process has been shown with speech streams in adults (Saffran, Newport, & Aslin, 1996) and infants (Saffran, Aslin, & Newport, 1996), in visually presented objects in adults (Fiser & Aslin, 2001, 2002a, 2005; Fiser, Scholl, & Aslin, 2007), infants (Fiser & Aslin, 2002b; Kirkham, Slemmer, & Johnson, 2002), and in other modalities (Conway & Christiansen, 2005). Together, these findings of sensitivity to co-occurrence statistics between elements are referred to as examples of statistical learning (Aslin & Newport, 2012; Perruchet & Pacton, 2006).

The learning of spatial statistics from visually presented stimuli has been investigated (Fiser & Aslin, 2001, 2002a, 2002b, 2005) by presenting, participants with a family of scenes, each of which contained multiple shapes arranged in a grid. In each scene the spatial arrangement of the entire array of shapes was changed, but the spatial arrangement between particular pairs of shapes (base-
pairs) was always preserved. After passive observation of these statistically constrained arrays of shapes without engaging any specific task, observers could distinguish base-pairs that have high co-occurrence statistics from shape-pairs that have lower co-occurrence statistics, demonstrating learning of spatial relations among the shape-pairs that were used to compose the scenes. This type of extraction of shape co-occurrence statistics, also termed visual statistical learning seems to depend, at least initially, upon sophisticated perceptual analyzers, lateralized to the right hemisphere, since Roser and colleagues found that the isolated right hemisphere of a callosotomy (split-brain) patient, but not his left hemisphere could perform the learning task (Roser, Fiser, Aslin, & Gazzaniga, 2011).

Visual statistical learning of shape-pair associations suggests a mechanism of human learning that occurs with little explicit top-down input from higher cognitive processes (Fiser, 2005; Fiser, Berkes, Orbán, & Lengyel, 2010) even though top-down processes can certainly influence this type of learning (Turk-Browne, Jungé, & Scholl, 2005). However, once such associations are learned by for example, a visual chunk learning mechanism (Orbán, Fiser, Aslin, & Lengyel, 2008), they become available as implicit top-down information leading to a variety of midlevel perceptual effects, and they can also influence the outcome of subsequent learning or even to emergence of explicit rules by constraining the statistics that are computed (MacKenzie & Fiser, 2008; Saffran, Pollak, Seibel, & Shkolnik, 2007).

A wealth of evidence suggests that the statistical learning process is intimately integrated with perceptual processes in the brain as a variety of constraints bias its outcome. In the auditory-temporal domain, the learning of co-occurrence statistics is limited to adjacent elements, unless nonadjacent elements share some perceptual property, for example, falling within the same musical octave (Cree, Newport, & Aslin, 2004), or the same speech category of vowel or consonant (Newport & Aslin, 2004). In the visual-temporal domain, the perception of streaming (continuous motion) or bouncing interactions between visual shapes affected the statistical learning of shape associations, suggesting that statistical learning is constrained at the level where representations of objects as spatiotemporal entities are formed (Fiser et al., 2007). In addition, overt attention to a subset of shapes (cued by color) limited the extraction of statistical information as the events unfolded in time (Turk-Browne et al., 2005). In the visual-spatial domain, a physical connectionedness, a cue to perceptual grouping, modulated statistical learning of shape associations (Baker, Olson, & Behrmann, 2004). Thus, both explicit and implicit mechanisms play a role in statistical learning under conditions of passive observation without feedback (Perruchet & Pacton, 2006), consistent with a process of statistical learning that begins with raw sensory input, continuously constrained by midlevel perceptual and attentional mechanisms, and then forms internal representations that themselves become constraints on the interpretation of novel inputs.

Given these demonstrations that statistical learning can be manipulated experimentally by various perceptual and attentional factors, it is important to establish how individual differences in perceptual and attentional processing influence visual statistical learning. Here we capitalize on the fact that the perceptual and learning capabilities of individuals with Autism Spectrum Disorder (ASD) differ substantially from those of non-ASD controls. For example, in ASD there is an avoidance of direct eye-contact, an insensitivity to social cues for conversational turn-taking, and a focus on nonsocial objects that manifests itself as high-level expertise in a limited domain (e.g., memorizing thousands of technical terms). On the other hand, both adults and children with High-Functioning Autism (HFA) or Asperger Syndrome often show a characteristic dissociation of elements of IQ tests that draw on visuospatial and linguistic processes (Bölte, Dziobek, & Poustka, 2009; Chalmers et al., 2011; Dawson, Soulières, Gernsbacher, & Mottron, 2007; Morsanyi & Holyoak, 2010; Soulières, Dawson, Gernsbacher, & Mottron, 2011). It is not clear how the above biases cause the characteristic pattern of competence that is often found in ASD, namely, that visuospatial ability is relatively preserved while linguistic ability is impaired. In fact, some studies have found superior visuospatial performance in ASD and Asperger Syndrome compared with controls in tasks involving matrix completion (Hayashi, Kato, Igarashi, & Kashima, 2008; Soulières et al., 2009) and block design (Caron, Mottron, Berthiaume, & Dawson, 2006).

One possible explanation for the spared or superior visuospatial performance by individuals with ASD is a bias toward the processing of local elements over global configurations of elements (e.g., in faces). This preferential processing of local details over global properties, as described by the theory of weak central coherence (Frith, 1989; Happé & Frith, 2006), has received considerable support for the preserved or superior visuospatial abilities in ASD (Rondan & Deruelle, 2007). A local preference in ASD has been found in tasks involving detecting figures within larger designs (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983), visual search for targets among distractors (Plaisted, O’Riordan, & Baron-Cohen, 1998), block design (Shah & Frith, 1993), and identifying inverted faces (Langdell, 1978; Rondan & Deruelle, 2004). In studies presenting hierarchical “Navon” stimuli (Navon, 1977), people with autism do indeed show a global processing advantage, as do control subjects. This is seen in high (Ozonoff, Strayer, McMahon, & Filloux, 1994) and low (Deruelle, Rondan, Gepner, & Fagot, 2006) functioning children with autism, high-functioning adolescents with autism (Mottron, Burack, Stauder, & Robaey, 1999), and adults with autism (Rondan & Deruelle, 2007). Often, however, the degree of global advantage is less in ASD than in typical controls (Gross, 2005). Furthermore, a global advantage in a task of selective attention shifted to a local advantage in a task of divided attention with children with autism (Plaisted, Swettenham, & Rees, 1999). In addition, when the visuospatial configuration of elements was manipulated, the evidence for a relative bias in ASD toward local elements increased. Both children (Deruelle et al., 2006) and adults (Rondan & Deruelle, 2007) showed a preference for matching geometric patterns or schematic-face stimuli according to local element identity, rather than according to the spatial configuration of elements across the entire stimulus. The above evidence suggests that adults with ASD have a preference for local processing and an aversion to global, configurational processing.

At face value the foregoing results, and other findings of impairment of configural processing in autism, might predict impairment in visuospatial statistical learning compared with non-ASD controls. Configuration, however, is typically defined as encompassing the entire complex stimulus—such as between the mouth and the rest of a schematic face (Deruelle et al., 2006; Rondan &
The elements of a real face (Tanaka, Kay, Grinell, Stansfield, & Szecht, 1998), or a pattern of colored blocks (Shah & Frith, 1993). In the stimuli used by Fiser and colleagues in their studies of visual statistical learning, however, critical configurational relationships exist only between local elements. That is, the base-pairs can be positioned in any location within the larger scene and with any other base-pair, thereby creating a myriad of between-pair configurations. A bias, therefore, toward attending to local statistics defined over adjacent elements may lead to an increase in the processing of relationships between immediate-neighbors, thereby increasing the likelihood of encoding the base-pairs. In this regard, one might predict superior performance in a visuospatial statistical learning task, consistent with the good performance by individuals with autism on the block design task, which is attributed to the disruption of configural processing as attention is directed toward the arrangement of local elements and not toward the configuration of elements across the entire stimulus (block pattern) (Deruelle et al., 2006; Shah & Frith, 1993). Indeed, in an earlier implicit Alternating Serial Reaction Time learning study, Nemeth et al. (2010) found that children with autism performed no worse than age matched controls in a standard implicit skill-learning task.

Not all studies of implicit learning find preserved function in ASD, with mixed results across a variety of paradigms. Some studies of serial reaction time (RT) (Gordon & Stark, 2007), motor-sequence learning (Gidley Larson & Mostofsky, 2008), and prototype formation (Klinger & Dawson, 2001) have found impaired performance in ASD, whereas other studies (Barnes et al., 2008; Molesworth, Bowler, & Hampton, 2005; Travers, Klinger, Mussey, & Klinger, 2010) have found intact performance in these tasks. Thus, it is not yet established whether a deficit in implicit learning in ASD exists. More closely related to the current study, investigations (Mayo & Eigsti, 2012; Scott-Van Zeeland et al., 2010) of nonvisual statistical language learning in ASD found no group differences in behavioral indices of statistical learning, although the Scott-Van Zealand study found group differences in brain function. The Mayo and Eigsti (2012) study found evidence of learning in both groups (ASD and Control) and the Scott-Van Zeeland et al. (2010) study found no sensitivity to element covariance in either group. Most pertinent, Jeste and colleagues (2014) investigated visual statistical learning by presenting sequences of shapes singly in a continuous stream. The transitional probabilities of pairs of shapes were manipulated so that shapes were paired over time by virtue of high transitional probability within pair and low transitional probability between pair. This visual-temporal structure makes the paradigm similar to the bulk of studies of nonvisual statistical language learning in which stimuli are presented in sequence, but different from the paradigm used in the current study in which visuospatial covariance between elements presented simultaneously was manipulated. Jeste et al. (2014) did not present behavioral data indexing learning, so it is not possible to determine whether visual statistical learning in the temporal domain was impaired or otherwise in ASD. The authors did, however, find evidence for electrophysiological effects consistent with learning in the Control group and did not find these effects in the ASD group. Thus there was some evidence for a disruption to normal processing of visual statistics in ASD.

The goal of our study was to investigate the observational learning of visuospatial linkages between shapes from patterns of covariation both in adult and child ASD populations. Given that the successful statistical learning of shape-pair associations depends on the extraction of constituent neighboring elements of larger displays, we predicted that the documented local advantage in ASD will facilitate the identification of shape pairs and result in a comparable or better performance in individuals with autism compared with healthy controls.

Method

Participants: Child

Twenty-eight children (CASD) with a diagnosis either of HFA, Asperger Syndrome, or ASD and 22 age-matched control children (CCONT) with no diagnosed learning or psychological disorder were tested. All CASD participants had a definitive clinical diagnosis meeting either Diagnostic and Statistical Manual for Mental Disorders-Fourth Edition (DSM–IV) (Association & DSM–IV, 1994) or ICD-10 (Organization, 1993) criteria. Diagnoses were carried out by either a pediatrician or child and adolescent psychiatrist after multidisciplinary assessment. Participants were recruited from two primary (elementary) schools (named in the Acknowledgments) local to the institution at which the research was conducted. Recruitment of individuals who met the criteria for inclusion in the study was handled by the Special Educational Needs Coordinators of the two schools. No details of psychological assessments were available. The CASD group comprised four females and 24 males with a mean age of 13 years (SD = 1.64). The CCONT group comprised 12 females and 10 males with a mean age of 13 years (SD = 1.62). Data on any other medical or cognitive condition, or data on IQ and other aspects of performance were not made available to the researchers but all participants had normal or corrected-to-normal vision. Constraints on the available testing time for child participants precluded further collection of data on these factors. The two groups were, therefore, distinguished on the basis of the presence of HFA, Asperger Syndrome, or ASD. All child participants were educated in mainstream schools.

Participants: Adult

Ten adults (AASD) with a diagnosis either of HFA, Asperger Syndrome, or ASD and 10 control participants (ACONT) were tested. Participants were recruited through a support group for adults with autism, affiliated with the National Autistic Society (NAS United Kingdom). Diagnoses were made by a variety of health, educational and social-services bodies. The stage of life (during education, midlife) at which diagnoses were made, and the period since diagnosis, varied between participants. Descriptive statistics for age, intelligence (IQ), handedness, and autism quotient, split by group and sex, are shown in Table 1. There were four women in the AASD group and six women in the ACONT group. Groups were matched for age, IQ, and handedness. IQ was assessed with The Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). IQ data were unavailable for two of the female AASD participants. Handedness was assessed with the Edinburgh Handedness Inventory (EHI) (Oldfield, 1971) in which positive score denote a preference for use of the right hand. The AASD group ranged from strongly right to strongly left-handed, as

Participants: Adult
Table 1
Descriptive Statistics and Means for Intelligence, Handedness Assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and Autism-Spectrum Quotient (Baron-Cohen et al., 2001) Scales

<table>
<thead>
<tr>
<th>Group</th>
<th>Sex</th>
<th>N</th>
<th>Age</th>
<th>IQ</th>
<th>EHI</th>
<th>AQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>AASD</td>
<td>M</td>
<td>8</td>
<td>37</td>
<td>117</td>
<td>−12</td>
<td>40</td>
</tr>
<tr>
<td>AASD</td>
<td>F</td>
<td>4</td>
<td>45</td>
<td>113</td>
<td>12</td>
<td>42</td>
</tr>
<tr>
<td>ACONT</td>
<td>M</td>
<td>6</td>
<td>41</td>
<td>123</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>ACONT</td>
<td>F</td>
<td>6</td>
<td>32</td>
<td>117</td>
<td>1</td>
<td>13</td>
</tr>
</tbody>
</table>

Stimuli

The stimuli were the same as those used by Fiser and Aslin (2001). Twelve complex two-dimensional shapes were created. Shapes were black on a white background and were displayed within a 3 × 3 grid using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). The grids were displayed centrally from ~57 cm, subtending ~12 degree visual angle with a shape size of ~1.2 degrees.

Procedure

The experiment consisted of Familiarization and Test phases. Unbeknownst to the participants, the multishape scenes were randomly composed from six base pairs, each pair comprising two of the shapes in an invariant spatial relationship. Figure 1 shows three base pairs (horizontal, oblique, and vertical) combined to form one multishape scene.

These six base pairs were then further divided into two sets of three pairs, and the scenes for the Familiarization phase were created by configuring these three base-pairs within various positions in the 3 × 3 grid. For instance, the three base pairs shown in Figure 1 were also combined so that the horizontal pair was displayed in the same location as in Figure 1, but the other two pairs were moved, yielding a different multishape scene (see Figure 2) in which base pairings were preserved. This was done separately for the two sets. The three base pairs in each grid always comprised one pair grouped vertically, one pair grouped horizontally, and one pair grouped obliquely. Each base pair had four possible locations within the grid and each pair always neighbored at least one other pair (i.e., no gaps between adjacent shapes). This arrangement resulted in a total of 36 scenes, 18 constructed from the first set of pairs and 18 from the second set. Over all Familiarization stimuli, the probability of appearance of any given element, and of any given base pair, was 0.5. The probability of appearance of a nonbase pair was 0.02.

Before beginning the Familiarization phase participants were instructed to keep their gaze on the screen. Task instructions (see Appendix) were made intentionally vague to avoid informing participants as to the nature of the underlying spatial structure of the base pairs. Participants were told simply to watch the series of grids with shapes in them. This vagueness extended to the description of the research project as “observing visual displays.” No mention of familiarization or of any test or response requirements was made at this point. These instructions, therefore, were appropriate for the investigation of learning in a relatively undirected context. Participants were then presented with eight blocks of the 36 scenes in a random sequence. Stimuli were displayed for ~3,000 ms after a blank screen lasting 1,000 ms. A short rest was taken after each block of 36 trials. When viewed sequentially the stimuli appear as a complex array of shapes, frequently changing the specific shapes and their positions as each scene was presented.

After the Familiarization phase, participants were asked this series of questions verbatim.

1. Can you tell me anything about the things you have just seen?

2. Did you notice anything about the arrangement of the shapes in the grid?

3. Did you notice any regularity in the arrangement of the shapes?

4. Did you notice any shapes that were associated in any way with any other shapes?

5. Did you notice that shapes always appeared in pairs in the same spatial arrangement?

These questions were designed to systematically probe explicit awareness of stimulus-set structure without revealing it to participants. The progression from very general to more specific questions was designed to allow participants to provide descriptions after little prompting. To our knowledge this is the first use of a hierarchical questionnaire probing explicit knowledge in a study of visual statistical learning.

Participants were then informed (see Appendix) that certain shapes always appeared together in a particular arrangement, and this was demonstrated visually by the experimenter using their hands to illustrate the horizontal, vertical, and oblique arrangements. They were then instructed to identify pairs of shapes that appeared together in a particular spatial arrangement in the Familiarization phase of the experiment.

The Test phase immediately followed familiarization and debriefing, and consisted of two blocks of 48 trials. All six base-pair stimuli and six of the many nonbase-pair stimuli, randomly selected from the full set, were paired in a two-alternative task during the Test phase. On each test trial two grids (see Figure 3) were presented simultaneously for 3,000 ms, one to the left of center and one to the right. One grid contained a base pair and the other contained a nonbase pair. Each base-pair was presented eight times in each of two blocks of 48 trials, randomly ordered. Each presentation of a base pair was accompanied by a nonbase pair (randomly assigned). In each block of 48 trials, each base pair was presented four times on the left and four times on the right. Although elements in nonbase pairs did appear at test in grid locations where they had not appeared during familiarization, this was also true of elements in base pairs. Learning absolute element positions could not, therefore, be relied upon to distinguish between base pairs and nonbase pairs in the current experiment. The stimuli remained on the screen until a response was made. Participants responded by pressing one of two buttons (z or m) to indicate the pair on the left or on the right. Trials were separated by a blank screen of 1,000 ms.

**Results**

Figure 4 shows mean percent correct in the Test phase for all four groups. Participants in all groups clearly chose the base-pairs in preference to the random pairs. Separate one-sample *t* tests found that each group’s mean percent correct was significantly higher than chance (CASD: *t*(27) = 2.64, *p* = .014, *r* = .45; CCONT *t*(21) = 3.00, *p* = .007, *r* = .55; AASD: *t*(11) = 7.25, *p* < .001).

![Figure 2](image1.png)

*Figure 2.* A second grid array for the Familiarization blocks constructed by rearranging the three base pairs.

![Figure 3](image2.png)

*Figure 3.* Simultaneous presentation of two grids containing a False and a True pair for the Test blocks.

![Figure 4](image3.png)

*Figure 4.* Percent correct for each group. Error bars show SD.
.001, $r = .91$; ACONT: $t(11) = 3.55, p < .01, r = .49$). An analysis of variance (ANOVA) of data from child participants found no effect of group or sex, or interaction between these factors. Although effects of sex on visual statistical learning have not been documented, the factor was included here to check that the lack of control for the sex of participants was irrelevant. An ANOVA of data from adults found a significant effect of group ($F(1, 20) = 9.13, p < .01$, partial $\eta^2 = 0.31$), no effect of sex, $F(1, 20) = 1.42, ns$ and no interaction between these factors. The mean percent correct was, as predicted, significantly higher for the AASD group ($M = 85, SD = 16.7$) compared with the ACONT group ($M = 65, SD = 14.4$).

**Qualitative Responses: Child**

When asked the first open question about the stimuli no child participants offered any explanation, suggesting that shapes were associated implicitly, except for two Control children who identified that two sets of shapes were used to construct the grid arrays. One of these participants, when describing the two sets, identified a base pair, including the spatial arrangement of the shapes. This participant responded negatively to all subsequent questions, suggesting they had no explicit knowledge of the details, or existence of, pairwise association. The other participant responded to subsequent questions that the same two shapes had always appeared next to each other and correctly described two of the base pairs without mentioning spatial arrangement. Three other children, one CASD and two CCONT, responded to the second question that directly probed shape arrangement by describing one base pair correctly but did not generalize that knowledge to all shapes being paired. The CASD participant later identified a second base pair, including its spatial arrangement, when asked Question 4. One additional CASD participant, who had not identified any relevant structure in answers to Questions 1 and 2, stated in response to Question 3 that two shapes were “bonded” together and correctly identified a base pair. In response to Question 4, this participant identified another base pair but in neither answer were any details of spatial arrangement given. Seventeen children (10 CASD, 7 CCONT) responded in the affirmative when asked Question 5 (“Was there a pair structure?”), including a majority who had given no indication of awareness of the underlying base-pair arrangement when asked less revealing questions. In summary, the pattern of qualitative responses suggests that only a small proportion of child participants, both CASD and CCONT, gained some awareness of regularity in the stimuli, with pairwise appearance being more apparent than pairwise spatial arrangement. This evidence must be weighed against the observation that self-report by children can be inaccurate (Burkhart, Dunbar-Jacob, & Rohay, 2001), leaving open the possibility that explicit awareness was greater than that suggested by the postexposure questionnaire. Nevertheless, there is little reason to attribute above-chance performance in both groups to explicit awareness.

**Qualitative Responses: Adult**

Eight participants (6 AASD, 2 ACONT) identified that there were two groups of shapes when asked Question 1. One AASD participant showed good explicit knowledge of underlying structure. After Question 1, the AASD participant stated that shapes were arranged in three pairs, one vertical, one horizontal and one oblique. No further details of pair element identity (which shapes were paired) were provided. One further AASD participant and two further ACONT participants volunteered that shapes were paired or always appeared next to another when asked Question 1. Neither of these participants provided further details of the shapes that were paired when asked subsequent questions. When asked Question 2, one AASD participant answered that “some patterns clicked together” and referred again to this answer when asked Question 4. This participant, when asked Question 5, correctly identified two base pairs but did not provide information on their spatial arrangement. When asked Question 3, one ACONT participant answered that the six shapes in each grid comprised three pairs. No further details were elicited by subsequent questions. Together, these results suggest that only a small number of adult participants were aware of the simultaneous appearance of some shapes but not of the unchanging spatial arrangement of pairs. Only one AASD participant achieved this level of awareness.

**Discussion**

The present study is the first test of whether children and adults with autism are sensitive to the visuospatial statistics defined by shape co-occurrences in multishape scenes. After undirected observation of complex arrays, above-chance discrimination of shape pairs varying in their level of statistical coherence was found in both ASD groups, replicating this finding in adults without ASD (Fiser & Aslin, 2001, 2002a) and in typically developing infants (Fiser & Aslin, 2002b). This finding suggests that an unsupervised learning mechanism that links two or more elements in the visual array into meaningful chunks is intact in autism. Furthermore, as predicted, it was found that adults with autism were significantly better than age- and IQ-matched controls in identifying emergent visuospatial features defined by co-occurrence statistics, whereas 13-year-old children with autism performed no better than their age-matched controls. This finding from adults extends evidence for superior visuospatial processing in autism beyond performance in figure detection, block design, and visual search (Jolliffe & Baron-Cohen, 1997; Plaisted et al., 1998; Shah & Frith, 1983, 1993) and into the domain of learning. It also confirms earlier findings of equally good implicit learning abilities of ASD children compared with controls (Nemeth et al., 2010).

A preference in ASD for attending to local elements provides a plausible explanation for our finding of ASD superiority in this statistical learning paradigm. The critical statistics to be learnt exist between local elements, immediately adjacent, and not across the entire visuospatial array. A bias toward attending to local elements, as is apparent in ASD, may have facilitated the encoding of relationships between immediate-neighbor elements as fewer resources were allocated to perceiving the global array of shapes, which not only appears random but does not contain learnable statistics. Thus, our result may parallel the superior performance by individuals with autism on the block design task, attributed to a focus on the local arrangement of shapes (Deruelle et al., 2006; Rondan & Deruelle, 2007; Shah & Frith, 1993). Future research should investigate the learning of nonadjacent statistics in multishape visual scenes and the mediating role of Gestalt cues in linking these nonadjacent shapes (Baker et al., 2004).
Also of interest is whether the visuospatial properties of the stimulus can impede visual statistical learning in ASD populations. Here we have demonstrated that the restriction of systematic covariance to the local level is associated with enhanced learning in ASD. Many other visuospatial manipulations are available, however, and it has been demonstrated that visual statistical learning is constrained by the kinematic properties of moving objects perceived as either exhibiting continuous motion or collision (Fiser et al., 2007). Some visual processes, such as detection of motion from contrast (Bertone, Mottron, Jelenic, & Faubert, 2003, 2005) and the perception of depth (Kaplan, 2005), are found to be affected by ASD. It remains to be determined whether these factors constrain visual statistical learning in the same way that a preference for local details does. Further investigations of visual statistical learning in ASD might well profit from considering individual variability in visual processing and the heterogeneity of impairment in ASD. A recent study found a positive correlation between nonverbal IQ and an electrophysiological index of statistical learning (Jeste et al., 2014) in children with ASD. The current study, although demonstrating a significant effect in the predicted direction with modest sample sizes, does not allow for the full heterogeneity of the ASD spectrum to be represented, nor does it allow for the role of individual differences to be assessed as they were in Jeste et al. (2014). A sensible next step would be to investigate the association of individual differences in local bias with performance in learning locally defined visual statistics across a large and heterogeneous sample of people with ASD.

Our previous results (Fiser & Aslin, 2001, 2002a, 2002b, 2005; Roser et al., 2011) suggest the existence of a mechanism of human learning that can operate with little explicit input from higher cognitive processes, although it does incorporate iterative top-down input that builds on processes arising at a low level (Fiser et al., 2010). The results of the current study showing little explicit awareness of the co-occurrence statistics underlying above-chance performance are congruent with this model. Given that explicit knowledge has been found to be ineffective in guiding perceptual processing in children with autism (Ropar & Mitchell, 2002), it is likely that the preserved performance in our study seen in children with autism, and the superior performance seen in adults with autism, depends not on explicit testing of hypotheses but on an implicit system for encoding regularity in the environment in accordance with the finding of Nemeth et al. (2010). As suggested by previous studies involving callosotomy patients (Roser et al., 2011) and functional magnetic resonance imaging (fMRI) (Turk-Browne, Scholl, Chun, & Johnson, 2009), this is likely to involve visual-processing areas, predominantly in the right hemisphere.

The current study also testifies to the potential educational benefits of leveraging stimulus preferences associated with enhanced performance in ASD. Although an enhancement of visual statistical learning was only evident in the adult group, it is possible that the performance of ASD children, which was comparable to age-matched Controls, was supported by the use of locally defined statistics of a visual-configural nature. This result stands in contrast to the equivocal results from comparisons of individuals with autism and controls from previous studies of sequential visual statistical learning (Jeste et al., 2014), auditory statistical learning (Mayo & Eigsti, 2012; Scott-Van Zeeland et al., 2010) and implicit learning more generally (Barnes et al., 2008; Gidley Larson & Mostofsky, 2008; Gordon & Stark, 2007; Klinger & Dawson, 2001; Molesworth et al., 2005; Travers et al., 2010). This suggests that learning may be facilitated by presenting material in a way that maximally engages processes of relative strength in ASD. The most effective approach would likely involve tailoring materials to an individual’s profile of strengths. Confirmation of this speculation will require additional research.

The evidence for implicit learning among our four groups of subjects argues against an alternative explanation for superior performance in autism. Systemizing is a predilection for constructing or analyzing stimulus and response relationships to explain or predict behavior, technical outcomes, or environmental changes and is found to be higher than normal in ASD (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003). Systemizing involves looking for patterns and explicitly noting associations in an active attempt to identify rules (Baron-Cohen, Ashwin, Ashwin, Tavassoli, & Chakrabarti, 2009). A heightened tendency to systemize in our AASD group, relative to ACONT, may dispose individuals to seek structure in the seemingly random arrays of shapes with which they were presented. This process of active search for evidence to support or refute an explicit hypothesis is, however, at odds with the qualitative feedback from our participants, none of whom reported actively searching for structure or engaging any systematic testing of hypotheses and who showed little awareness of the statistics that later testing verified as well above chance. In a recent study (Pellicano et al., 2011) of foraging behavior, children with ASD were found to be less systematic than control participants. The authors speculated that the children with ASD may have failed to adopt a coherent representation of a large-scale visuospatial environment, which required explicit processes of orientation and memory for searched locations, to search efficiently. The systematicity hypothesis is also at odds with much previous research on statistical learning in an observational paradigm (Fiser et al., 2010), where subjects who utilize explicit hypothesis-testing strategies tend to do very poorly and sometimes perform below chance because they focus on spurious coincidences in the familiarization stimuli. Thus, a tendency to systemize is unlikely to account for the observed results.

A final consideration is the difference in the pattern of performance between children and adults. Although both age groups showed evidence for observational learning of visual statistics, only the adults with autism showed superior performance to age-matched controls. One possibility is that children, in comparison to adults, are less able to sustain the general attentional mechanisms that are required to perform well on statistical learning tasks (Allen & Courchesne, 2001). We know that general attention, even in the absence of task instructions, plays a facilitative role in statistical learning (Toro, Sinnett, & Soto-Faraco, 2005). Children may require more exposure to the learning materials than adults to overcome their less efficient sampling of the relevant statistical structures embedded in the familiarization stimuli. Thus, perhaps our familiarization phase resulted in a “floor effect” for both groups of children, and if the familiarization had been extended, ASD children might have exceeded the performance of the control children. Arguing against this explanation is evidence that typically developed adults and children do not differ in their performance on a...
statistical learning task given exactly the same familiarization exposure (Saffran, Newport, Aslin, Tunick, & Barrreco, 1997). Alternatively, impairment in ASD children in sustaining attention to externally imposed tasks (Christakou et al., 2013; Garrettson, Fein, & Waterhouse, 1990) may have ameliorated any advantage gained by a bias toward attending to local elements.

Another possibility for the absence of superior performance by the ASD children is that, like typically developing children and adults, they may have a bias to sample local (i.e., spatially adjacent) co-occurrence statistics, but not to the exclusion of other potentially relevant statistics (i.e., higher order relations). That is, efficient learners should allocate at least some of their resources to “explore” novel structures in the environment. In contrast, adults with ASD appear to be more likely to “exploit” familiar structures and, therefore, learn them better than typically developed adults. Thus, ASD children may not yet have progressed to this stage of exploiting local visuospatial statistics to the same degree as adults with ASD. Whether this exploitation process is driven by maturation or learning—and, if the latter, by a reward (or mastery) signal—is a topic for further investigation.

A final possibility for explaining these findings is that although implicit learning based on local feature co-occurrence may be intact in both children and adults with autism, by adulthood, individuals with autism may develop effective compensatory strategies based on these abilities to counterbalance their relative disadvantage in recognizing global patterns. These compensatory strategies would rely heavily on an excellent ability to process local statistical information. In contrast, a recent study found a marked decrement in the ability of healthy children around the age of 12–14 with no impairment to learn implicitly the statistical structure of a task (Janacsek, Fiser, & Nemeth, 2012). The authors proposed that this relative drop in sensitivity to raw statistical information is paralleled by a shift to a more internal-model-based interpretation of the incoming signal allowing a more successful learning of highly complex, configural structures in the input at the expense of less sensitivity to basic statistical measures (Janacsek et al., 2012). Thus, in light of this argument, and given the equally good performance by children with autism and their age-matched controls found in the present study, it is conceivable that while in the adults control population interpretation of the incoming signals is dominated by higher-order implicit pattern interpretations obscuring the small local relations between shape pairs, the local feature-based strategies used by adults with autism is more than adequate to perform well in the present task. A prediction of this conjecture is that when the building blocks of the stimuli in the visual statistical learning task become larger (comprising more elements) and more complex so that, similar to the arrangements used in Orbán et al. (2008), purely local measures could not help in choosing the correct configurations in the test, the relative advantage of adults with autism should vanish. In fact, an impairment of visual statistical learning might arise, paralleling the decrease in visual search performance seen in ASD participants with a transition from table-top to large-scale environments (Pellicano et al., 2011).

In summary, we found that children with autism performed equally well and adults significantly better than their age matched controls in a standard visuospatial implicit-learning task. These results provide an intriguing way to gain insights about both the typical development of visual pattern learning and the compensatory strategies individuals with autism use to circumvent their impairments in such tasks.

References


Christakou, A., Murphy, C., Chantiluke, K., Cubillo, A., Smith, A., Giampietro, V. . . . Rubia, K. (2013). Disorder-specific functional abnormalities during sustained attention in youth with Attention Deficit Hy-
peractivity Disorder (ADHD) and with Autism. Molecular Psychiatry, 18, 236–244. doi:10.1038/mp.2011.185
Pellicano, E., Smith, A. D., Cristiano, F., Hood, B. M., Briscoe, J., & Gilchrist, I. D. (2011). Children with autism are neither systematic nor


---

**Appendix**

**Instructions to Participants**

Instructions before the experiment:
This study will involve viewing pictures on a computer screen. Your task is to simply view some scenes of a grid with shapes in it. We will run 8 blocks, each of 36 presentations. Each presentation will last a few seconds. There will be a short break between each block.

Instructions after participant questioning:
Now there is a short test based on what you have seen. We call the first part of this study the “Familiarisation phase.” In this experiment we asked you to simply observe complex displays of stimuli; the shapes in the grid. The grids were constructed by combining pairs of shapes, for example Shape A might have always appeared directly above Shape B (example arrangement demonstrated visually by experimenter using hands). During the Familiarisation phase you might have noticed that certain shapes always appeared together in a particular arrangement. Now we will test you to see if you can identify pairs of shapes that were paired with each other during the Familiarisation phase. This time on each trial you will see two grids; one presented on the left and the other on the right. Each grid will have only TWO shapes in it. One of the pairs of shapes always appeared together in that arrangement when they were included in the larger displays. The other pair is false and did not appear together in that arrangement during the Familiarisation phase. You must decide which pair of shapes appeared together. If it was the pair on the LEFT hit the Z key. If it was the pair on the RIGHT hit the M key.

Received December 9, 2013
Revision received July 21, 2014
Accepted July 21, 2014