This is an author produced version of a paper published in *Neuropsychology*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/42915

**Published paper**


http://dx.doi.org/10.1016/j.neuropsychologia.2010.11.012
The influence of visual saliency on fixation patterns in individuals with Autism Spectrum Disorders

Megan Freeth*, Tom Foulshamb, Peter Chapmanc

a University of Sheffield – Psychology Department
Western Bank, Sheffield. S10 2TP
United Kingdom
*m.freeth@sheffield.ac.uk

b University of British Columbia – Department of Psychology
2136 West Mall, Vancouver, B.C. V6T 1Z4
Canada
tfoulsham@psych.ubc.ca

c University of Nottingham – School of Psychology
University Park, Nottingham. NG7 2RD
United Kingdom
Peter.Chapman@nottingham.ac.uk

*corresponding author
Abstract

It is widely reported that individuals with Autism Spectrum Disorders (ASD) direct their attention in an atypical manner. When viewing complex scenes, typically developing individuals look at social aspects of scenes more rapidly than individuals with ASD. In the absence of a strong drive to extract social information, is something else capturing attention in these initial fixations, such as visually salient features? Twenty four high-functioning adolescents with ASD and 24 typically developing matched control participants viewed a series of indoor and outdoor scenes while their eye movements were tracked. Participants in both groups were more likely to fixate on salient regions in the first five fixations than later in viewing. Peak saliency at fixation occurred at fixation two for the typically developing participants but at fixation three for ASD participants. This difference was driven by typically developing participants looking at heads earlier than ASD participants – which are often visually salient. No differences between groups were observed for images in which the heads were not salient. We can therefore conclude that visual saliency impacts fixation location in a similar manner in individuals with ASD and those with typical development. It was found that social features in scenes (heads) captured attention much more than visually salient features, even in individuals with ASD.

Keywords
Autism; visual saliency; eye tracking; social attention
1. Introduction

The world around us contains a vast array of complex information. We must selectively attend to the important and relevant components of this array to avoid overwhelming our sensory and cognitive capacities. Our attention system is directed using a combination of top-down and bottom-up information processing. Top-down processing is the influence of our pre-conceived knowledge of the world. This involves our ability to recognise items from previous experience and to identify them as being targets worthy of fixation. Bottom-up information comprises image-dependant features such as changes in colour, intensity and orientation.

Many aspects of attention are reported to be a-typical in individuals with Autism Spectrum Disorders (ASD) (see Ames & Fletcher-Watson, 2010 for a review). When viewing complex stimuli, individuals with ASD do not attend to all top-down information in the same way as their typically developing peers (Klin, Jones, Schultz & Volkmar, 2003). Studies using neuropsychological and neuroimaging techniques have provided great insight into the brain structures and neural processes that potentially contribute to these atypicalities of attention (see Sanders, Johnson, Garavan & Gallagher, 2008 for a review). Abnormal neural connectivity is also hypothesised to be a contributing factor (Belmonte, Allen, Beckel-Mitchener, Boulanger, Carper & Webb, 2004). It has been argued that social aspects of complex stimuli do not strongly capture attention in individuals with ASD (Jones & Klin, 2008), and many studies have found that individuals with ASD make fewer fixations on faces when viewing videos and complex scenes (e.g. Pierce, Conant, Hazin, Stoner & Desmond, in press; Klin, Jones, Schultz, Volkmar & Cohen, 2002; Riby & Hancock, 2008, Sasson, Tsuchiya, Hurley, Couture, Penn, Adolphs & Piven, 2007).
A possible contributing factor to reduced social attention could be abnormal capture of attention by low-level, bottom-up visual properties of complex stimuli. A now well established model of how low-level visual properties may guide attention was proposed by Itti and Koch (2000). Natural images are analysed in terms of contrast, intensity, colour and orientation. These features are then combined to provide an estimate of the saliency at each point in the image: the extent to which scene regions stand out from their background. This model has previously been shown to be a useful predictor of where people fixate when viewing scenes (e.g. Parkhurst, Law & Neibur, 2002; Foulsham & Underwood, 2008) especially in visual agnosia patients, who have limited access to top-down information (Foulsham, Barton, Kingstone, Dewhurst & Underwood, 2009; Mannan, Kennard & Husain, 2009). On the other hand, many studies in healthy individuals have claimed that effects of saliency are small or non-existent, especially when there are strong top-down constraints, such as during a search task (Foulsham & Underwood, 2007) or in the presence of social stimuli (Birmingham, Bischoff & Kingstone, 2009). At a neural level, it has been proposed that regions involved in recognising meaningful objects (such as inferotemporal cortex, which might be compromised in some visual agnosics) or social stimuli (such as the superior temporal sulcus or the amygdala) modulate early visual areas and feed into the control of spatial attention and eye movements in prefrontal cortex (see Treue, 2003, for a review). It has also been suggested that the attention of young children with ASD may be captured by different visual features, when viewing sparsely revealed face stimuli (Neumann, Spezio, Piven & Adolphs, 2006) and watching video clips (Shic, Scassellati & Chawarska, 2007), and that these differences might reflect impaired neural functioning. However, at present the influence of visual saliency on fixation patterns in older individuals with ASD is not well established.
It has recently been shown that high-functioning adolescents and adults with ASD do attend to people when presented with complex static scenes, but initial fixations are still less likely to be on a faces (Fletcher-Watson, Leekam, Benson, Frank & Findlay, 2009; Freeth, Chapman, Ropar & Mitchell, 2010a). It is important to understand what captures and directs attention in these initial fixations. This will provide insight into what is perceived and the way in which information about the world is pieced together.

The aim of the experiment reported in this paper was to investigate whether high-functioning adolescents with ASD are influenced by visual saliency when viewing natural scenes in a similar or different manner to matched typically developing control participants in the early fixations of image viewing. Fletcher-Watson et al. (2009) found no overall difference in the mean saliency at fixation between high functioning adolescents and adults with ASD and controls when viewing naturalistic photographic scenes. However, at present the influence of saliency at early fixations in individuals with ASD is not known. It is particularly important to know whether there are any differences in the influence of visual saliency at these early fixations as differences in social attention are strongest at this point. Processing stimuli differently at this early stage could be hugely influential to the way stimuli are processed overall. The importance of subtle differences here may be magnified when attempting to process the rapidly changing environment in everyday life. In typically developing individuals it has been argued that saliency has a greater influence on initial fixations (Parkhurst et al., 2002). Is this also the case for individuals with ASD? Is this effect stronger in ASD? Answering these questions is necessary for determining whether early attention is operating in a different way in individuals with ASD and whether this could be contributing to their social attention abnormalities.
2. Materials and Methods

2.1 Participants

Twenty four 11-16 year old high-functioning (Full-Scale IQ >70) adolescents (21 males, 3 females) with an autism spectrum disorder (ASD) and 24 age, gender and Full-Scale IQ matched typically developing adolescents participated in the study. All of the participants with ASD had been formally diagnosed by a mental health professional according to DSM-IV criteria (American Psychiatric Association, 1994) and as a result had a statement of Special Educational Needs for Autism or ASD. An Autism Spectrum Screening Questionnaire (Ehlers, Gillberg & Wing, 1999) was completed by a teacher or parent of each participant giving an indication of current level of autistic features. See Table 1 for further details of participant information.

(Insert Table 1 about here)

2.2 Stimuli

Sixteen photographs (640x320 pixels) of everyday indoor and outdoor scenes were presented, each containing at least one person (see Figure 1 for examples). The images did not vary systematically on person size or number of people. These photographs were taken by the first author at various locations around the University of Nottingham campus and were presented in full colour.

(Insert Figure 1 about here)

2.3 Apparatus

Eye movements were recorded using a remote Tobii 1750 eye-tracker system. The eye-tracker sampled at 50Hz and was accurate to 1° of visual angle. The images were displayed on a 19” colour LCD monitor at a distance of approximately 60cms and subtended...
a visual angle of approximately $32^\circ$ horizontally and $24^\circ$ vertically. A fixation was recorded if eye-tracking points were within $1.5^\circ$ of visual angle for 80ms or more.

2.4 Procedure

Following a calibration on the eye-tracker and a head position check to ensure that both eyes were in range for recording, participants were presented with a series of photographs for 5 seconds each. Participants were told that all they needed to do was to look at the photos. Participants performed this task as part of a larger testing battery reported elsewhere (Freeth, Chapman, Ropar and Mitchell, 2010a; Freeth, Ropar, Chapman and Mitchell, 2010b).

3. Results

Participants in both groups made a very similar mean number of fixations per image: typically developing participants = 30.3; ASD participants = 31.0. An independent samples $t$-test confirmed that there was no difference between groups, $t(46)=0.40$, $p=0.69$, indicating that both groups were similarly attentive to the images. Visual saliency maps were created using the Saliency Toolbox (Walther & Koch, 2006) which uses the parameters originally identified by Itti and Koch (2000). The Saliency Toolbox produced a low-resolution (40 x 20 pixels) map of saliency values corresponding to each image. For example images and resulting visual saliency maps, see Figure 1. The mean saliency value of all map values was subtracted from the saliency at each fixation to take into account the average saliency in each image. Therefore, adjusted saliency values indicated whether saliency at fixation was above that expected by fixation at a random point on the image. Two single-sample $t$-tests demonstrated that the mean saliency at fixation for both the typically developing group and the ASD group were significantly above that of the map mean (typically developing group,
Saliency at each of the first five fixations was calculated (and adjusted by subtracting the map mean from each value). A 2 x 2 mixed-measures ANOVA (first five fixations vs. mean of all fixations x group) found that saliency at fixation was higher over the first five fixations than the mean saliency value for all fixations, $F(1,46)=337.5, p<.001$. There was no difference between groups, ($F<1$), and no interaction between factors ($F<1$) indicating that like the typically developing participants, the participants with ASD fixated on more salient regions over the first five fixations than later in viewing. Mean saliency over each of the first fixations was then considered separately. A 5 x 2 mixed-measures ANOVA (fixation number x diagnosis) found a main effect of fixation number, $F(4,184)=30.71, p<.001$. This effect was found to have a highly significant linear component, $F(1,46)=41.1, p<.001$, saliency values were higher at later fixations than earlier fixations, but there was also a highly significant quadratic component, $F(4,184)=69.1, p<.001$. As can be seen in figure 2, peak saliency at fixation was at around fixations 2 and 3 with lower values before and after this point. There was no main effect of diagnosis ($F<1$) but there was a significant fixation number x diagnosis interaction, $F(4,184)=2.39, p=.05$, indicating that the groups were showing a subtly different pattern.

(Insert Figure 2 about here)

In order to investigate the potential reasons for the fixation x diagnosis interaction, we considered whether the social content of the scenes could be modifying performance. All 16 images were re-analysed using an algorithm from the Saliency Toolbox to predict the first 10 fixation locations on the scenes in order of decreasing saliency. In four images, none of the first 10 fixations predicted by the model were on a person’s head (See Figure 1 Panel B as
examples). These images were classified as being “Low Social Saliency” images (mean adjusted saliency value for the head regions = -0.054). In four other images, a person’s head was identified as being the most salient region (see Figure 1 Panel A as examples). These images were classified as being “High Social Saliency” images (mean adjusted saliency value for the head regions = 1.107). The socially-relevant regions in the two groups of images were thus markedly different in their saliency, with the heads being selected either immediately (high) or very late (low) by the model. The remaining eight images fell somewhere between the two classifications and were excluded from the current analysis. Over the first five fixations, there was a very large effect of Social Saliency, \( F(1,46)=212.6, p<.001 \) on the saliency at fixation values. Saliency at fixation was much higher for images in which a head was the most salient region. Surprisingly this was true for both the typically developing participants, \( F(1,23)=148.9, p<.001 \), and the participants with ASD, \( F(1,23)=83.8, p<.001 \) and there was no significant overall difference in mean saliency at fixation between groups for these High Social Saliency images, \( F(1,46)=0.01, p=.92 \). However, a 5 x 2 x 2 (fixation number x social saliency x diagnosis) mixed-measures ANOVA revealed a 3-way interaction, \( F(4,184)=3.46, p=.009 \) (see Figure 3). Individuals with autism looked at images with High Social Saliency differently to controls over the first five fixations, as demonstrated by a significant fixation number x diagnosis interaction for “High Social Saliency” images, \( F(4,184)=3.72, p=.006 \); whereas the images with Low Social Saliency were looked at in a similar manner, \( F(4,184)=0.37, p=.83 \). Figure 3 shows that typically developing individuals tended to fixate the most salient regions in the High Social Saliency images by fixation 2 - 3 whereas the ASD participants did not fixate these regions until approximately one fixation later, thus displaying a different pattern of fixations for these images. In order to check that the between group interactions were likely to be driven by the ASD participants being slower to fixate the heads in the scenes, a region of interest (ROI) analysis was conducted. The head
regions in all images were identified, and fixations within this area were classed as on the head. The mean number of fixations taken before one of the heads was fixated was compared between groups. The ASD participants fixated the heads significantly later than the typically developing participants, $t(46)=2.25, p=.03$ (mean number of fixations, ASD=4.6, typical participants=3.8). In addition, ROI analyses on the images with High and Low Social Saliency were conducted. For the High Social Saliency images, the mean number of fixations before the heads were fixated were lower but the trend for typical participants to fixate the heads sooner than the ASD participants was still present, $t(46)=1.72, p=.09$, (mean number of fixations ASD=4.0, typical participants=2.9). There was no difference between groups for the Low Social Saliency images, $t(46)=0.12, p=.91$, but this is likely due to the heads in these images being very small – and not likely to be fixated at all – or centrally presented and so difficult to avoid (see Figure 1).

(Insert Figure 3 about here)

4. Discussion

The aim of the reported experiment was to investigate whether low-level visual properties of scenes capture the attention of high-functioning individuals with ASD in a similar or different way to their typically developing peers. In both groups, saliency at fixation was higher than the mean saliency value of the maps indicating that participants were in general fixating on regions that were of above average visual saliency. Overall, there was no difference in mean saliency at fixation between groups. This replicates the findings of Fletcher-Watson et al. (2009).

The novel findings of this paper related to the patterns exhibited over the first five fixations and how these relate to visual saliency maps associated with the images. It has previously been shown that saliency at fixation tends to be higher for early fixations than later fixations (e.g. Parkhurst et al. 2002). This finding was replicated in the reported
experiment. This pattern was found in both the typically developing participants and the ASD participants, thus suggesting that both groups were fixating on regions that were more visually salient in the first 5 fixations than later in viewing. We hypothesised that visually salient regions of images might capture the attention of individuals with ASD early in viewing and that this might be a reason for previous studies finding that individuals with ASD do not attend to social aspect of scenes early in viewing (Fletcher-Watson et al. 2009; Freeth et al. 2010a). However, this hypothesis was not supported. Individuals with ASD were not drawn to fixate on salient regions early in viewing more so than controls. The reason for their delay in fixating social regions must be due to other factors. On the contrary, the data suggest that social aspects of the scenes actually captured attention a great deal more than properties related to visual salience. In scenes where faces were particularly salient, fixations were much more likely to fall on salient areas. In support of findings reported by Birmingham et al. (2009), social information captured attention much more than low-level visual properties. This was true for both the typically developing participants and the participants with ASD and this effect was very large in both groups.

The reported data also support the finding that individuals with ASD tend to be drawn to faces significantly later than typically developing individuals (Fletcher-Watson et al. 2009; Freeth et al., 2010a), supporting the hypothesis that individuals with ASD lack the strong drive to rapidly attend to social information when presented with complex stimuli. Participants with ASD did not generally fixate on the people’s heads until fixation 4-5 whereas the typically developing controls fixated these regions by fixation 3-4. Although this difference seems small, in the rapidly changing world this difference may be having marked effects on the information that is processed. The reason for this absence of rapid attention to social aspects of stimuli is still unclear but the current study demonstrates that attention capture by low-level visually salient properties of stimuli is not a causal factor. Any
differences in fixation of salient regions in this study were explained by an atypical response to socially-relevant objects, i.e. heads. It is possible that the fast-track mechanism to attend to other people is merely absent in individuals with ASD.

In the future it will be important to discover whether there are any stimuli that capture the attention of individuals with ASD in a similar manner to the way that social stimuli clearly capture the attention of typically developing individuals. Individuals with ASD are known to take a special interest in certain classes of items, such as cars and computers (Sasson, Turner-Brown, Holtzclaw, Lam, & Bodfish, 2008). Items of special interest vary from person to person though commonalities in special interest in the ASD population do exist (South, Ozonoff, & McMahon, 2005). It would be interesting to discover whether these types of items rapidly capture the attention of individuals with ASD in a similar manner to the way in which social stimuli rapidly capture the attention of typically developing individuals when viewing complex scenes.

In summary, high-functioning adolescents are influenced by low-level visually salient properties of stimuli to a similar extent as their typically developing peers. Salient regions are fixated more early in viewing than later in viewing. In the current study, peak saliency at fixation occurred at fixation two for typically developing individuals but at fixation three for individuals with ASD. This effect was driven by the typically developing participants looking at salient heads earlier than the ASD participants. No differences in saliency at fixation were observed for images that did not contain salient heads. We can therefore conclude that the absence of a strong drive to rapidly attend to heads in individuals with ASD is not explained by low-level visually salient features capturing attention early in viewing. Overall, social features of scenes (heads) capture attention much more than visually salient features, even in individuals with ASD.
Acknowledgements

This research would have been impossible without the kind cooperation of staff, students, and parents involved in the studies. Participating schools: Aldercar Community Language College, Bigwood, Bracken Hill, Bramcote Park and Enterprise School, Cruckton Hall, Foxwood, The Long Eaton School, Ysgol Plas Brondyffryn.

This research was funded by a PhD studentship from the University of Nottingham awarded to the first author and Economic and Social Research Council grant number PTA-026-27-2283.
References


### Table 1

Participant characteristics – Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>ASD participants</th>
<th>Typically developing participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Age (years; months)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>13;10</td>
<td>14;0</td>
</tr>
<tr>
<td>SD</td>
<td>1.37</td>
<td>1.37</td>
</tr>
<tr>
<td>Range</td>
<td>11;6 – 16;8</td>
<td>11;2-16;4</td>
</tr>
<tr>
<td>Full-scale IQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>97.0</td>
<td>95.5</td>
</tr>
<tr>
<td>SD</td>
<td>13.6</td>
<td>9.53</td>
</tr>
<tr>
<td>Range</td>
<td>74-129</td>
<td>79-112</td>
</tr>
<tr>
<td>ASSQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td><strong>18.4</strong></td>
<td><strong>3.13</strong></td>
</tr>
<tr>
<td>SD</td>
<td>12.8</td>
<td>4.04</td>
</tr>
<tr>
<td>Range</td>
<td>0-42</td>
<td>0-14</td>
</tr>
</tbody>
</table>

** p<.001: Participants with ASD scored significantly higher on the ASSQ than typically developing participants.
**Figure Caption Sheet**

*Figure 1.* Example stimuli and their corresponding visual saliency maps (with warmer regions indicating higher saliency at that location). The first 10 fixations predicted by the saliency model are indicated (1 = 1st fixation location). Panel A - examples of High social saliency stimuli. Panel B - examples of Low social saliency stimuli.

*Figure 2.* Mean saliency at fixation – all images. Panel A – first 5 fixations; Panel B – mean of all fixations.

*Figure 3.* Mean saliency at fixation – 4 High social saliency images; 4 Low social saliency images.