Typical integration of emotion cues from bodies and faces in Autism Spectrum Disorder

Rebecca Brewer¹, Federica Biotti², Geoffrey Bird³,⁴,⁵, & Richard Cook²

¹Department of Psychology; Royal Holloway, University of London
²Department of Psychology; City, University of London
³Institute of Cognitive Neuroscience; University College London
⁴Experimental Psychology Department; University of Oxford
⁵MRC Social, Genetic & Developmental Psychiatry Centre; King’s College London

*Corresponding author:
Rebecca.Brewer@rhul.ac.uk
Department of Psychology
Royal Holloway, University of London
Egham Hill, Egham,
Surrey, UK, TW20 0EX

Key words: Autism Spectrum Disorder; Social Perception; Contextual Modulation; Emotion recognition; Body posture
Highlights

- Cues to another’s emotion are automatically integrated across the face and body
- We tested integration of face and body emotion cues in Autism Spectrum Disorder
- Integration of face and body emotion cues in individuals with autism was typical
- Reliance on body cues was greatest in those with poor facial emotion classification
Abstract
Contextual cues derived from body postures bias how typical observers categorize facial emotion; the same facial expression may be perceived as anger or disgust when aligned with angry and disgusted body postures. Individuals with Autism Spectrum Disorder (ASD) are thought to have difficulties integrating information from disparate visual regions to form unitary percepts, and may be less susceptible to visual illusions induced by context. The current study investigated whether individuals with ASD exhibit diminished integration of emotion cues extracted from faces and bodies. Individuals with and without ASD completed a binary expression classification task, categorizing facial emotion as ‘Disgust’ or ‘Anger’. Facial stimuli were drawn from a morph continuum blending facial disgust and anger, and presented in isolation, or accompanied by an angry or disgusted body posture. Participants were explicitly instructed to disregard the body context. Contextual modulation was inferred from a shift in the resulting psychometric functions. Contrary to prediction, observers with ASD showed typical integration of emotion cues from the face and body. Correlation analyses suggested a relationship between the ability to categorize emotion from isolated faces, and susceptibility to contextual influence within the ASD sample; individuals with imprecise facial emotion classification were influenced more by body posture cues.
1. Introduction

The facial expressions of others are a rich source of social information, conveying cues to affective and mental states. Correct interpretation of facial expressions is therefore important for fluent social interaction and wider socio-cognitive development (Adolphs, 2002; Frith, 2009). Previous research indicates that facial emotion perception is affected by the context in which a facial expression is encountered, suggesting that interpretations are informed by our knowledge and experience (de Gelder et al., 2006; Feldman-Barrett, Mesquita, & Gendron, 2011). Categorization of morphed facial expressions, for example, is biased by the concurrent presentation of social interactants (Gray, Barber, Murphy, & Cook, 2017) and other non-interacting faces (Masuda et al., 2008). Perceived facial expression can also be influenced by affective vocal cues (de Gelder & Vroomen, 2000; Massaro & Egan, 1996), situational stories (Carroll & Russell, 1996), and visual scenes (Righart & De Gelder, 2008a, 2008b).

A particularly strong form of contextual influence is exerted by body postures (Hassin, Aviezer, & Bentin, 2013). The same facial expression can vary in appearance when presented with different bodily expressions; for example, a facial expression may be classified as angry when presented in the context of a body expressing anger, but disgusted when presented in the context of a body expressing disgust (Aviezer et al., 2008; Aviezer, Trope, & Todorov, 2012b). These findings imply that the attribution of affective states involves the integration of emotion cues from across the face and body. The influence of posture contexts is often automatic; it occurs despite explicit instructions to disregard non-face information (Aviezer, Bentin, Dudarev, & Hassin, 2011), and modulates early neurophysiological markers of visual person processing (Meeren, van Heijnsbergen, & de Gelder, 2005).

The present study sought to examine whether observers with Autism Spectrum Disorder (ASD) integrate emotion cues from body posture contexts when interpreting others’ facial emotion. ASD is a neurodevelopmental condition associated with social and communication difficulties, repetitive behaviors and restricted routines (American Psychiatric Association, 2013). There has been considerable interest in the visual perception of individual with ASD (Dakin & Frith, 2005; Simmons et al., 2009). Observers with ASD may exhibit a local processing style that hinders their ability to form unified global percepts (Behrmann, Thomas, & Humphreys, 2006; Happé, 1999; Happé & Frith, 2006). For example, ASD is associated with good detection of embedded figures, which requires observers to disregard extraneous
information present within a complex pattern or scene, to locate a target element (Ropar & Mitchell, 2001; Shah & Frith, 1983). Those with ASD may also be less susceptible to context-induced visual illusions than typical individuals (Happé, 1996; Shah, Bird, & Cook, 2016; but see Manning, Morgan, Allen, & Pellicano, 2017) and show reduced global-to-local interference when responding to (“Navon”) compound letter arrays (Behrmann, Avidan et al., 2006). Similarly, individuals with ASD may rely less than typical individuals on contextual cues to distinguish homographs when reading (Frith & Snowling, 1983; López & Leekam, 2003).

Should observers with ASD exhibit a local processing style, their judgements of facial emotion may be less susceptible to modulation by body posture contexts. Where observed, aberrant integration of emotion cues from bodies and faces may contribute to difficulties attributing affective states sometimes seen in this population (Gaigg, 2014). Observers with ASD and matched control participants were required to classify expressions drawn from a morph continuum as either ‘Anger’ or ‘Disgust’. Target expressions were either judged in a no-context baseline condition, in the presence of a task-irrelevant disgusted posture, or a task-irrelevant angry posture.

2. Method

2.1 Participants

Nineteen individuals with a clinical diagnosis of ASD (three female; $M_{\text{age}} = 34.84$ years), and 27 individuals with no current or previous clinical diagnosis (eight female; $M_{\text{age}} = 33.85$ years) took part in the current study. All participants were aged between 18 and 65 years. Typical participants were recruited from local participant pools populated by university students and members of the general public. ASD participants were recruited from a database maintained by the authors. Individuals with ASD were diagnosed by an independent clinician, and the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) was used to assess current severity. Autistic traits were also measured in all participants using the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). Higher AQ scores, indicative of more ASD traits, were seen in the ASD group than in the typical group [$t(44) = 7.28, p < .001$]. Alexithymia, a trait associated with difficulties identifying and describing one’s own emotions (Brewer, Cook, & Bird, 2016a, 2016b), measured by the Toronto Alexithymia Scale (TAS-20; Bagby, Taylor, & Parker, 1994), was
also more severe in the ASD group ($M = 59.16$, $SD = 15.81$) than the typical group ($M = 46.58$, $SD = 12.77$) [$t(43) = 2.95$, $p = .005$]. However, the ASD and typical groups did not differ significantly in their age [$t(44) = .27$, $p = .787$], proportion of female participants [$X^2 = .13$, $p = .248$], or IQ [$t(43) = 1.032$, $p = .308$]. Detailed diagnostic information is provided in Table 1.

Table 1

2.2 Stimuli
Facial stimuli (Figure 1a) were static images drawn from a morph continuum created by blending two images of the same actor expressing disgust and anger (images taken from Ekman & Friesen, 1975) using Morpheus Photo Morpher Version 3.11 (Morpheus Software, Inc). The continuum parametrically manipulated the actor’s expression between disgust and anger in seven equidistant steps of 10%. The body contexts depicted the same actor posing angry and disgusted postures (Figure 1b). Where these posture contexts have been used previously (e.g., Aviezer et al., 2008; Aviezer et al., 2012b), the disgusted body posture has been shown gripping a disgusting object. In the present study, this object was removed to ensure that perceptual bias, where observed, was attributable to integration of face and body cues, and not additional semantic information. The morphed facial expressions were presented within a dark grey oval intended to resemble a hood. The relative location of the facial target did not vary as a function of expression intensity. In the baseline no-context condition, observers saw the expressions presented within the oval, but in the absence of a body posture. We opted for this baseline condition in light of concerns about the value of the neutral emotion construct (e.g., Lee, Kang, Park, Kim, & An, 2008); for example, a supposedly “neutral” body posture, where an actor’s arms are clenched by their side, may be perceived as angry. In all three conditions, facial stimuli subtended $3.5^\circ$ vertically when viewed at 57 cm.

Figure 1

2.3 Procedure
Each trial in the experimental procedure began with a central fixation point (1000 ms), followed by presentation of a facial target drawn from the morph continuum (1200 ms). The
facial target was always presented centrally. In the baseline no-context condition, the facial target was presented in isolation. In the two context conditions, the facial target was accompanied by an angry or disgusted body posture. Following stimulus offset, participants were prompted to categorize the facial emotion as either ‘Disgust’ or ‘Anger’ using a key press. Participants were explicitly instructed to disregard the body context. The procedure consisted of 420 trials (7 facial stimuli × 20 presentation × 3 context conditions) and was presented on an LCD display. Stimuli were presented in a randomized order, with the three context conditions interleaved. The experimental program was written in MATLAB (The MathWorks, Inc) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

For each observer, we fitted separate psychometric functions for the three context conditions, each modelling how the probability of a disgust response varied as a function of the strength of the disgust signal in the stimulus. Cumulative Gaussian functions were fitted using the Palamedes toolbox (Prins & Kingdom, 2009). Each function estimated two key parameters: Decision noise and the point of subjective equality (PSE). Decision noise is a measure of the precision with which stimuli are categorized, defined as the standard deviation of the symmetric Gaussian distribution underlying each cumulative Gaussian function. Lower noise estimates indicate that observers can perceive subtle differences in stimulus strength and vary their responses accordingly. Greater noise estimates reveal that participants’ responses are relatively invariant to changes in stimulus strength. Noise estimates are inversely related to the slope of the psychometric function; steep and shallow slopes are associated with low and high noise estimates, respectively. The PSE is a measure of bias that represents the hypothetical emotion intensity equally likely to be judged as ‘Disgust’ and ‘Anger’. Observers’ susceptibility to the contextual modulation was inferred from the difference between PSE of the anger function and the PSE of the disgust function (Figure 1c).

3. Results

One individual in the ASD group (participant 19) produced psychometric functions that could not be modeled so was excluded from all analyses. Goodness-of-fit was evaluated by calculating deviance scores for each function (McCullagh & Nelder, 1989). In general, fits were good in both groups; deviance scores (whereby lower deviance indicates better fit) for those with ($M = 6.97$, $SD = 8.55$) and without ($M = 8.85$, $SD = 7.44$) ASD did not differ significantly in the baseline no-context condition [$t(43) = .781$, $p = .439$].
First, we examined the decision noise of the two groups (Figure 2a). The ASD (M = 15.0%, SD = 13.9%) and typical (M = 11.7%, SD = 5.94%) groups did not differ significantly with respect to decision noise in the baseline no-context condition [t(43) = .968, p = .344], suggestive of broadly comparable emotion recognition. However, the ASD group exhibited greater variability in decision noise than the typical group (Levene’s F = 9.97, p = .003). Next, we examined the groups’ susceptibility to the contextual modulation (Figure 2b). Significant contextual modulation (shifts > 0%) was observed in both the ASD [t(17) = 2.50, p = .023] and typical [t(26) = 3.43, p = .002] groups. Furthermore, the ASD (M = 7.73%, SD = 12.9%) and typical (M = 3.47%, SD = 5.26%) groups did not differ in their susceptibility to the contextual modulation [t(43) = 1.30, p = .207], although modulation was numerically greater in the ASD group. The ASD and typical groups were also comparable in terms of their baseline decision noise [t(36) = 1.24, p = .234] and their susceptibility to the contextual modulation [t(36) = 1.17, p = .261] when participants of 50 years-of-age or over were excluded (four participants in each group).

Correlation analyses (Figure 2c) conducted on the combined sample of 45 observers revealed a significant positive association between decision noise in the baseline no-context condition, and susceptibility to the contextual modulation [r = .63, p < .001]. Participants who exhibited imprecise classification of isolated facial expressions exhibited greater contextual modulation. When the ASD and typical groups were separated, this association remained significant in the ASD group [r = .73, p < .001], even when an outlier with a large context effect (participant 17 of the ASD sample) was removed [r = .74, p = .001]. This correlation did not reach significance in the typical group [r = .19, p = .333]. A Fisher r-to-z test indicated that the strength of the correlation was significantly greater in the ASD than the typical group [z = 2.22, p = .026].

ADOS score was not significantly correlated with either contextual modulation (r = .03, p = .901) or decision noise in the baseline no-context condition (r = -.01, p = .983). In the sample as a whole, AQ scores correlated with context induced PSE shift (r = .32, p = .035), whereby presence of more ASD traits was associated with greater PSE shifts, but this correlation did
not remain significant when participant 17 was removed ($r = .18, p = .237$). AQ scores were not associated with decision noise in the baseline no-context condition ($r = .15, p = .322$). TAS-20 scores were not significantly correlated with contextual modulation ($r = .03, p = .856$) or baseline decision noise ($r = -.02, p = .904$).

4. Discussion

The current study investigated whether individuals with and without ASD differ in the extent to which body posture contexts influence interpretation of emotional facial expressions. Despite explicit instructions to disregard the context, the presence of the body postures biased observers’ categorization of facial emotion; expressions were more likely to be judged as angry and disgusted when aligned with angry and disgusted bodies, respectively. Contrary to prediction, however, participants with and without ASD were influenced to a similar degree by emotional cues present in the body contexts.

Previous studies indicate that individuals with ASD are less susceptible to contextual effects in a range of domains (e.g., Behrmann, Avidan et al., 2006; Happé, 1996; López & Leekam, 2003; Ropar & Mitchell, 2001), leading to the view that ASD is associated with difficulties forming integrated percepts (Behrmann, Thomas et al., 2006; Happé, 1999; Happé & Frith, 2006) and problems using context to inform perceptual inference (Lawson, Rees, & Friston, 2014; Palmer, Lawson, & Hohwy, 2017). Our finding that body contexts bias perception of facial emotion in observers with ASD is not consistent with this characterisation of ASD. The reason for this apparent disparity warrants further investigation. In typical observers, integration of bodily and facial cues likely emerges in response to covariation of facial and bodily expressions. It is possible that adults with ASD have sufficient visual experience of these contingencies to develop typical integration mechanisms, despite a local processing style. Reduced attention to faces (Dawson, Webb, & McPartland, 2005; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Nakano et al., 2010) may also promote greater reliance on bodily cues in this population.

We observed striking variation in expression categorization ability in our ASD sample. While some members of the ASD group exhibited categorization performance comparable with the best typical participants, three performed at least two standard deviations below the typical mean. This adds further weight to the emerging view that deficits of facial expression
perception in ASD are weak and inconsistent at the group level (Harms, Martin, & Wallace, 2010; Lozier, Vanmeter, & Marsh, 2014; Uljarevic & Hamilton, 2013). Indeed, one influential review concluded that “behavioral studies are only slightly more likely to find facial emotion recognition deficits in autism than not” (Harms et al., 2010, p317). While is it undeniable that some individuals within the ASD population experience difficulties recognizing facial emotion, these deficits do not appear to be a universal feature of this condition, and may instead reflect co-occurring conditions such as alexithymia (Bird & Cook, 2013; Cook, Brewer, Shah, & Bird, 2013).

Interestingly, the susceptibility of those with ASD to contextual modulation was related to their performance when judging the facial expressions in isolation; observers with imprecise expression categorization in the no-context condition were influenced more by the body posture contexts when present. A similar trend was observed in the typical group, however this correlation did not reach significance, possibly due the limited variability in baseline decision noise in this sample (e.g., Howitt & Cramer, 2009). While correlations inferred from small samples must be treated with caution (Schönbrodt & Perugini, 2013), the suggestion that stimulus ambiguity and contextual influence are closely linked, accords with current thinking about the role of context in human vision (Bar, 2004; Friston, 2005; Gilbert & Li, 2013; Gregory, 1997). When target stimuli are unmistakable, observers may distinguish between different perceptual hypotheses without reference to the context. When the physical attributes of a target stimulus do not clearly distinguish between different perceptual hypotheses, however, contextual information may be weighted more strongly to resolve the ambiguity. This view also fits well with the finding that physically similar facial expressions associated with intense positive and negative emotions (Aviezer, Trope, & Todorov, 2012a), or created by image morphing (Van den Stock, Righart, & De Gelder, 2007), are particularly receptive to contextual influence from body postures.

Some readers may be concerned that participants with poor facial emotion recognition chose to ignore the facial cues entirely, and responded to bodily emotion instead. Crucially, any attempt to treat the current procedure as a bodily expression recognition task would yield responses that failed to vary as a function of facial expression intensity. All but one participant (who was excluded from all analyses), however, produced monotonic classification functions that could be modelled using a cumulative Gaussian.
In conclusion, the current findings suggest that individuals with ASD and typical observers show broadly comparable integration of emotion cues from faces and bodies. Consistent with previous studies, individuals with ASD differed widely in their baseline ability to classify facial expressions. Correlation analyses suggested a relationship between ability to categorise emotion when faces were presented in isolation, and susceptibility to contextual influence; those observers with ASD who exhibited imprecise categorization in the baseline condition were influenced more by body posture contexts. This finding accords with the view that the visual system weights contextual information strongly when interpreting ambiguous stimuli.
References


Figures

Figure 1

(a) Facial expressions were drawn from a morph continuum that blended anger and disgust emotions in 10% increments. (b) The facial expressions were either presented in isolation (no context) or aligned with angry and disgusted body postures. (c) An observer’s susceptibility to the contextual modulation was inferred from the difference between the PSE of their disgusted function and the PSE of their angry function.
Figure 2. (a) Decision noise and (b) PSE estimates for the ASDs and the typical observers in the three conditions. (c) Scatterplot of the correlation between baseline decision noise and susceptibility to the contextual modulation.
**Table 1.** Autism Diagnostic Observational Schedule (ADOS) classification and total score, Autism Quotient Score, and Full-Scale IQ score for all participants in the ASD group. Mean AQ, IQ, and TAS-20 scores for typical participants are shown below. Note: one typical individual did not complete the TAS-20, and one did not complete the IQ assessment.

<table>
<thead>
<tr>
<th>Participant</th>
<th>ADOS Classification</th>
<th>ADOS</th>
<th>Gender</th>
<th>Age</th>
<th>AQ</th>
<th>Full-Scale IQ</th>
<th>TAS-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Autism Spectrum</td>
<td>7</td>
<td>Male</td>
<td>61</td>
<td>45</td>
<td>132</td>
<td>61</td>
</tr>
<tr>
<td>2</td>
<td>Autism</td>
<td>10</td>
<td>Male</td>
<td>35</td>
<td>46</td>
<td>112</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>Autism</td>
<td>10</td>
<td>Female</td>
<td>22</td>
<td>20</td>
<td>121</td>
<td>44</td>
</tr>
<tr>
<td>4</td>
<td>Autism</td>
<td>11</td>
<td>Male</td>
<td>31</td>
<td>37</td>
<td>118</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>Autism Spectrum</td>
<td>8</td>
<td>Male</td>
<td>18</td>
<td>21</td>
<td>107</td>
<td>73</td>
</tr>
<tr>
<td>6</td>
<td>Autism Spectrum</td>
<td>7</td>
<td>Male</td>
<td>38</td>
<td>40</td>
<td>116</td>
<td>71</td>
</tr>
<tr>
<td>7</td>
<td>Autism Spectrum</td>
<td>7</td>
<td>Male</td>
<td>38</td>
<td>17</td>
<td>125</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>Autism Spectrum</td>
<td>8</td>
<td>Male</td>
<td>35</td>
<td>36</td>
<td>108</td>
<td>72</td>
</tr>
<tr>
<td>9</td>
<td>Autism</td>
<td>14</td>
<td>Male</td>
<td>50</td>
<td>50</td>
<td>121</td>
<td>84</td>
</tr>
<tr>
<td>10</td>
<td>Autism Spectrum</td>
<td>9</td>
<td>Male</td>
<td>19</td>
<td>31</td>
<td>92</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>Autism</td>
<td>12</td>
<td>Male</td>
<td>31</td>
<td>48</td>
<td>107</td>
<td>64</td>
</tr>
<tr>
<td>12</td>
<td>Autism</td>
<td>15</td>
<td>Female</td>
<td>52</td>
<td>45</td>
<td>116</td>
<td>69</td>
</tr>
<tr>
<td>13</td>
<td>Autism</td>
<td>15</td>
<td>Male</td>
<td>52</td>
<td>27</td>
<td>118</td>
<td>41</td>
</tr>
<tr>
<td>14</td>
<td>Autism</td>
<td>10</td>
<td>Male</td>
<td>21</td>
<td>35</td>
<td>107</td>
<td>72</td>
</tr>
<tr>
<td>15</td>
<td>Autism</td>
<td>11</td>
<td>Male</td>
<td>42</td>
<td>33</td>
<td>117</td>
<td>36</td>
</tr>
<tr>
<td>16</td>
<td>Autism Spectrum</td>
<td>9</td>
<td>Male</td>
<td>25</td>
<td>39</td>
<td>133</td>
<td>61</td>
</tr>
<tr>
<td>17</td>
<td>Autism Spectrum</td>
<td>8</td>
<td>Female</td>
<td>21</td>
<td>46</td>
<td>93</td>
<td>81</td>
</tr>
<tr>
<td>18</td>
<td>Autism Spectrum</td>
<td>9</td>
<td>Male</td>
<td>33</td>
<td>35</td>
<td>128</td>
<td>60</td>
</tr>
<tr>
<td>19</td>
<td>Autism</td>
<td>14</td>
<td>Male</td>
<td>38</td>
<td>23</td>
<td>78</td>
<td>56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>ASD mean</th>
<th>AQ</th>
<th>Full-Scale IQ</th>
<th>TAS-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.84</td>
<td>35.47</td>
<td>113.11</td>
<td>59.16</td>
<td></td>
</tr>
<tr>
<td>(12.46)</td>
<td>(10.14)</td>
<td>(14.03)</td>
<td>(15.81)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Typical mean</th>
<th>AQ</th>
<th>Full-Scale IQ</th>
<th>TAS-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33.85</td>
<td>16.93</td>
<td>108.73</td>
<td>46.58</td>
<td></td>
</tr>
<tr>
<td>(11.92)</td>
<td>(5.42)</td>
<td>(14.06)</td>
<td>(12.77)</td>
<td></td>
</tr>
</tbody>
</table>