Autism and the Extraction of Emotion From Briefly Presented Facial Expressions: Stumbling at the First Step of Empathy

Tedra F. Clark
Midcontinent Research for Education and Learning, Denver, Colorado

Piotr Winkielman
University of California, San Diego

Daniel N. McIntosh
University of Denver

Identification of other people’s emotion from quickly presented stimuli, including facial expressions, is fundamental to many social processes, including rapid mimicry and empathy. This study examined extraction of valence from brief emotional expressions in adults with autism spectrum disorder (ASD), a condition characterized by impairments in understanding and sharing of emotions. Control participants were individuals with reading disability and typical individuals. Participants were shown images for durations in the range of microexpressions (15 ms and 30 ms), thus reducing the reliance on higher level cognitive skills. Participants detected whether (a) emotional faces were happy or angry, (b) neutral faces were male or female, and (c) neutral images were animals or objects. Individuals with ASD performed selectively worse on emotion extraction, with no group differences for gender or animal–object tasks. The emotion extraction deficit remains even when controlling for gender, verbal ability, and age and is not accounted for by speed–accuracy tradeoffs. The deficit in rapid emotional processing may contribute to ASD difficulties in mimicry, empathy, and related processes. The results highlight the role of rapid early emotion processing in adaptive social–emotional functioning.

Keywords: autism, emotion, faces, empathy

The ability to identify other people’s emotions, including their facial expressions, is fundamental to many social processes (Ekman, 1984). Among such processes is empathy, broadly considered a complex combination of affective and cognitive processes in which a person understands or experiences the emotions of another (Davis, 1994). Here, we focus on a simple but crucial component of empathy processes. Empathy begins with another’s emotional condition (Eisenberg & Fabes, 1990), so an early step involves at least rudimentary identification of another person’s affective state (Hoffman, 1977; Preston & de Waal, 2002). In particular, many basic empathy processes require that emotion be quickly extracted from briefly presented stimuli. For example, online perception of quickly changing emotional states (Ickes, 2003), detection of subtle social cues (Ekman, 1984), or rapid mimicry of emotional expression (Dimberg, Thunberg, & Elmenhed, 2000) may involve detection of expressions lasting around one-thirtieth of a second (Bartlett et al., 2005). Stumbling at this first step of emotion processing could have important consequences for empathy and other forms of social–emotional functioning. For example, in children an impairment in understanding and sharing of affective states may impair development of intersubjectivity (i.e., creating shared experience and understandings) with their caregivers, which can have implications for development of cognitive, language, and emotional abilities (Yirmiya, Sigman, Kasari, & Mundy, 1992).

Our understanding of the basic mechanisms of empathy can be informed by studying individuals with known difficulties in social–emotional functioning. One such group is individuals with autism spectrum disorder (ASD). These individuals show deficits in several empathy-related processes, including emotional contagion, intersubjectivity, and rapid and spontaneous mimicry of emotional expression (Kasari, Sigman, Yirmiya, & Mundy, 1993; McIntosh, Reichmann-Decker, Winkielman, & Wilbarger, 2006; Moody & McIntosh, 2006; Oberman, Winkielman, & Ramachandran, in press; Stel, van den Heuvel, & Smeets, 2008; Yirmiya et al., 1992). This study examines the ability of people with ASD to extract emotional information from quickly presented faces. This examination should illuminate not only the nature of empathy impairments in autism, but also the role of rapid emotional processing in social functioning of typical individuals.
Several studies have suggested that individuals with ASD are atypical in processing of emotion in general and of facial expression in particular. Much of this research has been inspired by proposals that affective impairments are central to autism, as reflected by the inability to develop reciprocal, affectively based relationships through social interactions (Hobson, 1989). Indeed, there is much research showing impairment of affective exchanges between young autistic children and their mothers (Dawson, Hill, Spencer, Galpert & Watson, 1990; Kasari et al., 1993; Rogers & Pennington, 1991).

One limitation of the literature on emotion processing in autism is the scarcity of data on extraction of emotional information from briefly presented stimuli. This is important because, as mentioned earlier, adaptive social functioning, including rapid and spontaneous mimicry, requires online extraction of expression information from stimuli presented for as little as one-thirtieth of a second (30 ms). Longer presentation durations potentially allow ASD individuals with high verbal intelligence to deploy strategically, in a top–down fashion, their cognitive skills to determine affective information and decide whether to use such information in judgment (Santos, Rondan, Rosset, Da Fonseca, & Deruelle, 2008). In fact, there is evidence that responsiveness to others’ emotions in ASD increases with cognitive functioning (Dissanayake, Sigman, & Kasari, 1996). Consistent with this notion, neurotypical individuals tend to use a template-based strategy to perceive emotional facial expressions, but people with ASD rely more on a rule-based strategy (Rutherford & McIntosh, 2007). Individuals with ASD are also less likely than controls to use facial expression in judging the pleasantness of a face (Celani, Battacchi, & Arcidiacono, 1999). Finally, during extremely short exposures (such as 15 ms or 30 ms), emotional discrimination should depend less on potential group differences in facial gaze patterns, such as latency of saccades to look at or look away from the stimulus. Specifically, the shortest latency of saccades is on the order of 150 ms (Rayner, 1998). Furthermore, ASD-related group differences in fixation patterns tend to emerge around 200 ms (Sasson et al., 2007). In short, these reasons suggest that examining discrimination performance of briefly presented facial expressions is a useful way to characterize potential ASD deficits in early emotion processing.

Present Study

In this study, we investigated the ability of individuals with ASD to extract emotional information from stimuli with exposure durations in the time range of microexpressions (15 ms and 30 ms). We predicted that control participants (typically developing individuals and individuals with a reading disability) would perform relatively well at extracting emotional information from briefly presented faces. We based this prediction on research showing that typical young adults can extract valence from faces presented for as little as 10 ms, even before they are able to identify other nonaffective dimensions of the faces (Murphy & Zajonc, 1993). There is also research suggesting that typical individuals can generate automatic mimicry responses to expressions presented for only 30 ms (Dimberg et al., 2000). However, with limited perceptual input, individuals with ASD should be impaired at extracting emotional dimensions of the face as compared with nonemotional dimensions.

We also examined the specificity of possible ASD impairment in extracting affective information. To do that, we tested whether extraction of affective information from faces (used in empathy processes) is impaired relative to extraction of descriptive information from affectively neutral faces (gender) and nonfaces (neutral animals and objects). It is worth noting that we did not predict any general ASD deficits in face processing. This is because our paradigm required all individuals to maintain focus on the stimulus, there were no subtle changes in faces across presentations, and the task did not require configural (holistic) face-processing strategies or depend on attention to the eyes—factors that underlie face-processing deficits in ASD (Dalton et al., 2005; Hobson, Ousten, & Lee, 1988; Schultz, 2005).

To summarize, in this study we compared the ability of individuals with ASD and controls to extract emotional information from quickly presented faces, an important initial step in empathy processes. The use of very short presentation durations substantially reduced participants’ ability to use higher level cognitive skills to infer emotional information from the stimuli. Furthermore, we explored the performance of individuals with ASD in extracting nonaffective information (gender) from neutral faces and their ability to perceive nonfacial, nonaffective stimuli (animals and objects). This allowed us to separate deficits in automatic extraction of emotion from more general processing impairments.

Method

Participants

Participants were a group of 15 high-functioning adolescents and adults with ASD and two comparison groups: 10 nonautistic individuals with a history of reading disability (RD) and 11 typically developing individuals.

Typically developing participants were recruited from the departmental participant pool and advertisements in the community. Individuals with ASD and reading disorders were recruited from the existing departmental participant pool and included people who had participated in one of several research studies of autism and developmental dyslexia. We selected the latter group as a control because dyslexia, like ASD, is a relatively specific developmental disorder that can be diagnosed despite above-average general IQ. Dyslexia also matches ASD in the relatively high male-to-female ratio. Finally, because participants with ASD and dyslexia were both recruited from area clinics and parent groups, potential biases because of recruitment strategy were reduced. Considering all this, if ASD participants show differences from not only the typical group but also the dyslexia group, then the inference that these differences are specific to ASD rather than to developmental disability in general will be strengthened.

The individuals with ASD were provisionally accepted into the study if they had received a diagnosis of infantile autism or Asperger’s syndrome from a child psychiatrist, developmental pediatrician, or licensed clinical psychologist. Actual participation required that this diagnosis be recently confirmed, with each having met the criteria for ASD within the past 3 years on the basis of scores on the Autism Diagnostic Inventory—Revised (Lord, Rutter, & Le Couteur, 1994) or the Autistic Diagnostic Observation—Generic (Lord et al., 2000).

All individuals in the RD group had a history of early reading problems and either were identified through participation in a
study of developmental dyslexia conducted at the university or were evaluated at the departmental neuropsychology clinic. The presence of a reading disorder was verified by the staff of the neuropsychology clinic using a combination of reading, achievement, and intelligence test scores.

Individuals were excluded from participation if they had a significant hearing or visual impairment or if they had another medical condition that would prohibit full participation in this study (premature birth greater than 4 weeks, brain injury, and seizures). Individuals were also excluded if they had autism associated with another condition such as fragile X syndrome or tuberous sclerosis.

Descriptive information for each group is found in Table 1. The three participant groups were similar in gender, with each group consisting predominantly of men. All effects described later hold when controlling for gender. The three groups were matched on verbal ability, as measured by the Peabody Picture Vocabulary Test (within 1 standard deviation, or 15 points of the standard score). Although it would be preferable to have groups more closely matched on verbal functioning (Peabody Picture Vocabulary Test), the primary task itself is not strongly verbal, and as described later, all primary statistical effects hold when controlling for verbal ability. Typical participants and those with RD were matched on mean age such that we found no statistically significant difference between the two groups ($t < 1$). The ASD group was slightly older than the typical group, $t(21) = 2.23, p < .05$, but not the RD group ($t = 1.48$). This age analysis was completed after bringing one outlier (a 64-year-old individual with ASD) down to the second highest age (age 41). Moreover, as described later, we found no significant correlations between age and accuracy or the second highest age (age 41). Moreover, as described later, all primary statistical effects hold when controlling for verbal ability. Typical participants and those with RD were matched on on mean age such that we found no statistically significant difference between the two groups ($t < 1$). The ASD group was slightly older than the typical group, $t(21) = 2.23, p < .05$, but not the RD group ($t = 1.48$). This age analysis was completed after bringing one outlier (a 64-year-old individual with ASD) down to the second highest age (age 41). Moreover, as described later, we found no significant correlations between age and accuracy or response times across item type and duration, and the critical findings held even when we excluded the oldest participants.

Procedure

For each participant, the experiment contained three blocks, each involving a different judgment type: (a) emotion, (b) gender, and (c) animal or object. Blocks were counterbalanced across participants according to a Latin Square design. In the emotion block, participants saw a series of faces varying in expression and decided whether each face was happy or angry. In the gender block, participants saw a series of neutral faces varying in gender and decided whether each face was male or female. In the animal–object block, participants saw a series of neutral images (e.g., a bird or a clock) and decided whether each image was an object or an animal.

Each block consisted of a practice session followed by an experimental session. During the practice session, 12 stimuli were presented 1 at a time for 3 s each, 6 for each dimension (e.g., happy versus angry). Each stimulus was preceded by a 2-s fixation point and immediately followed by a pattern mask. After the mask, a prompt appeared on the screen asking participants to make a decision about the stimulus (“Was the face happy or angry?”), “Was the face a male or a female?”, or “Was the picture of an animal or an object?” depending on the block). During practice sessions, participants were instructed to say their answers out loud and then press the appropriately labeled key on the computer keyboard (e.g., $H$ for happy and $A$ for angry). The practice sessions allowed the participants to become acquainted with the task and the experimenters to be sure participants could achieve 100% accuracy in the long-duration presentations before continuing to the short-duration presentations in the experimental session. The experimenter monitored and recorded participants’ responses during the practice session and immediately gave feedback on any incorrect responses before moving to the next trial.

The experimental session for each block immediately followed the practice session. Each experimental session consisted of 48 trials, 24 of each stimulus dimension. For each dimension, half the stimuli were presented for 15 ms, and half were presented for 30 ms, and this was freshly randomized for each participant. We selected these durations because in the pretests the 15-ms presentations were just above the threshold of awareness, and the 30-ms duration yielded performance around 75% (thus being most sensitive to any group effects). Furthermore, the variation in duration also allowed us to assess how much the participants’ performance benefited from additional input. Participants were warned that the stimuli would be presented very quickly and were encouraged to be as accurate as possible. In addition, participants made their judgments on the keyboard, so that reaction times could be collected, and were no longer required to say them out loud. Everything else was the same as in the practice session.

Pictues were presented to participants using PCs equipped with high-quality video cards and E-Prime experimental software and CRT monitors with refresh rate set at 85 Hz. We verified presentation duration parameters by means of E-Prime timing records, which revealed that the effective screen refresh rate was 12 ms, and the effective onscreen duration for the 15-ms condition was two refresh rates (24 ms) and for the 30-ms condition, three refresh rates (36 ms).

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Age (M ± SD)</th>
<th>PPVT score (M ± SD)</th>
<th>Gender (male)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism spectrum disorder</td>
<td>26.0 ± 7.5</td>
<td>99.5 ± 23.8</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>Reading disorder</td>
<td>21.6 ± 5.4</td>
<td>105.8 ± 9.8</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Typical</td>
<td>19.7 ± 4.6</td>
<td>112.9 ± 4.8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Note. PPVT = Peabody Picture Vocabulary Test.

1 Happy and angry expressions were used to provide the clearest valence contrast. Happiness is typically used for positive emotion because it is easily recognizable and unambiguous in valence. Anger is typically the most easily recognizable negative expression and the least confused with happiness (Ekman, 1984). In addition, using anger reduces issues related to known ASD impairments in processing of fear expressions (Corden, Chivers, & Skuse, 2008). Note also that to a perceiver, seeing fear is more ambiguous than seeing anger. Seeing another’s anger clearly implies that the person is afraid of the perceiver or of something else in the shared environment—a process that requires more complex processing (Whalen et al., 2001). In short, using anger avoids several issues associated with the presentation of fear faces.
Stimuli

Stimuli used for the emotion and gender blocks consisted of black and white faces of males and females taken from the Pictures of Facial Affect set (Ekman & Friesen, 1976). The pictures were digitized, cropped, and normalized according to Cottrell, Dailey, Padgett, and Adolphs (2001). In the emotion block, the pictures consisted of 12 different individuals, each displaying a happy expression in one picture and an angry expression in a second picture. The 24 pictures were each repeated once to make up the 48 trials, with stimulus order freshly randomized for each participant. The pictures used for the gender block consisted of the same 12 individuals as in the emotion block, each displaying neutral facial expressions. The 12 pictures were repeated four times to make up the 48 trials, with presentation order again randomized for each participant. For the animal–object block, 24 affectively neutral animal (12) and object (12) pictures were chosen from the International Affective Picture System (Lang, Bradley, & Cuthbert, 1997). Each picture was repeated once to make up the 48 trials, randomized for each participant.

Results

Practice Session

In the practice session, the stimuli were presented for 3 s each. Under this presentation condition, the mean performance exceeded 95% for all three response categories (emotion, gender, and animal–object tasks), even before the experimenter gave any feedback. In short, participants had no deficits in emotion perception under conditions that afforded long stimulus input.

Experimental Session

During the experimental session, the stimuli were presented in the time range of microexpressions (15 ms and 30 ms). The relevant means for correct responses are presented in Table 2. We first analyzed the proportion of correct responses in a full design, as a function of judgment type (emotion, gender, and object–animal), diagnosis (typical, RD, and ASD), and exposure (15 ms and 30 ms) in a 3 × 3 × 2 mixed factorial analysis of variance (ANOVA). This analysis revealed a theoretically uninteresting main effect for judgment type, with participants doing overall slightly better on emotion and animal–object tasks than on the gender task, F(2, 66) = 6.16, p < .01. There was also a large main effect of duration, with participants doing much better with 30-ms than with 15-ms presentations, F(2, 66) = 115, p < .001. There was also a two-way interaction between duration and diagnosis, F(2, 66) = 10.1, p < .001. This interaction was qualified by a higher order three-way interaction between diagnosis, duration, and judgment, F(4, 66) = 4.39, p < .01. As one can see in Table 2, this interaction is driven by the fact that on emotion judgment, participants with ASD did not benefit from additional stimulus input conveyed by longer duration (59% vs. 60%), even though they significantly benefited from longer duration on gender (60% vs. 63%) and animal–object judgments (66% vs. 79%). In contrast, post hoc comparisons revealed that the RD and typical groups significantly benefited from longer duration on all three judgments (p < .05). Subsequent analysis focused on performance in individual duration conditions. As one can see in Figure 1, the overall performance in the 15-ms condition was just slightly above chance (60%). This makes the condition relatively uninformative regarding group differences as a function of judgment type because they might be obscured by the floor effect. However, as shown in Figure 2, the overall performance at 30 ms was at the targeted level of 75%. Thus, we focused more detailed analysis on this condition. A 2 × 3 ANOVA revealed a two-way interaction of judgment and diagnosis, F(4, 66) = 4.55, p < .01. This interaction was diagnosed with separate one-way ANOVAS on each judgment type. As predicted, on emotion judgment, the ANOVA was significant, F(2, 35) = 6.83, p < .005, with simple least significant differences tests showing that the ASD group performed worse as compared with the typical group and the RD group (both ps < .005). No other simple effects or pairwise comparisons were significant. We also obtained this selective ASD impairment on emotion discrimination even when 2 of the oldest participants were eliminated from the sample, F(2, 33) = 4.9, p < .05, with simple least significant differences tests showing significant impairments on emotion compared with both typical and RD groups (both ps < .05). Critically, as shown in Figure 2, in the 30-ms condition participants with ASD did not perform significantly worse than participants with RD or typical controls on gender or animal–object judgments.

We also tested reaction times for possible speed–accuracy tradeoffs. Analyses complementary to the ones described above showed no significant interactions with diagnosis. The only significant effect was for duration, with longer presentations associated with faster responses, presumably reflecting an easier task (p < .05). Separate analyses within each presentation condition found no diagnosis effects. This suggests that the recognition results do not reflect speed–accuracy tradeoffs. Finally, controlling for age, verbal ability, and gender did not change any of the results for detection performance or reaction time.

Discussion

In this study, we investigated the ability of individuals with ASD to extract emotional information from stimuli presented in

2 Note that in Table 2 the typical group appears to be performing worse than chance (.44) in making gender judgments at 15 ms. One possible reason for this is that the Ekman pictures used as stimuli include pictures of a few young women with long hair and a few older looking women without visible hair. To the extent that hair was perceived as a cue, it may have resulted in some women being mistaken for men. Furthermore, because these are neutral expressions, the women generally follow the cultural pattern for masculine pictures. These possible explanations are post hoc, however. Regardless, this finding is not critical for our hypotheses because it occurs only when responses are at floor and thus does not bear on the critical analyses in the 30-ms condition.

3 We also analyzed whether the emotion detection impairment differed by expression (happiness or anger). Expression type never interacted with diagnosis, with participants with ASD equally impaired on detecting happiness or anger (all Fs < 1). This finding suggests again that group differences in the rapid presentation paradigms are probably not driven by attention to different face regions because this would presumably influence happiness and anger differently.
the time range of microexpressions. The results suggest that indi-
viduals with ASD are selectively impaired in the extraction of
emotional information. In contrast, they performed similarly to
typically developing and RD controls on extraction of major
descriptive features from neutral faces (i.e., gender) and features
from neutral nonface stimuli (i.e., animals and objects). Thus, this
study supports the notion that individuals with ASD have a spe-
cific impairment in early extraction of emotion that is not neces-
sarily reflective of problems with general face processing, process-
ing of briefly presented stimuli, or general task motivation or
understanding.

In this study, we explored early and rapid emotion processes,
which rely more on automatic components and less on the use
of verbal intelligence and other top–down strategies that can
often compensate for deficits in emotional perception (Ruther-
ford & McIntosh, 2007). As such, the findings suggest that
individuals with ASD may have special difficulty in situations
that require the use of a fast emotion extraction mechanism.
Because of this perception deficit, certain kinds of downstream
empathy processes and related social–emotional functions
could be impaired (see Preston & de Waal, 2002). A deficit in
the ability to rapidly extract affective information may be
particularly disruptive in typical ongoing social interactions, in
which affective displays of one’s interaction partner may be con-
veyed fleetingly. For those who cannot extract emotional
information rapidly, empathic processes may be based more on
broader situational cues, scripts, or the perceiver’s own emo-
tional state instead of the dynamic changes in the emotional
state of the interaction partner. For example, impairments in
automatic extraction of valence from emotional facial expres-
sions may explain some of the observed deficits in spontaneous
and rapid, but not voluntary and slower, mimicry of emotional
facial expression (McIntosh et al., 2006; Oberman et al., in
press). Future research should examine the extent to which
spontaneous mimicry impairments derive from deficits in the
initial stage of emotion perception, the intermediate stage of
emotion elicitation (contagion), or the production of the match-
ing motor response (Moody & McIntosh, 2006). Because in
typical individuals facial mimicry can also facilitate detection
of emotional expressions (Niedenthal, Brauer, Halberstadt, &
Innes-Ker, 2001), it is also important to examine the hypothesis
that mimicry impairment in autism is a contributor to emotion
perception deficits, perhaps via somatosensory feedback mech-
isms (Heberlein & Adolphs, 2007).

Table 2
Mean Proportion of Correct Responses From the Three Participant Groups for the Three
Judgment Tasks by Exposure Duration

<table>
<thead>
<tr>
<th>Diagnosis and emotion</th>
<th>Emotion</th>
<th>Gender</th>
<th>Animal–object</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15 ms</td>
<td>30 ms</td>
<td>Total</td>
</tr>
<tr>
<td>Typical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>.65</td>
<td>.86</td>
<td>.76</td>
</tr>
<tr>
<td>SE</td>
<td>.03</td>
<td>.04</td>
<td>.03</td>
</tr>
<tr>
<td>Reading disability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>.63</td>
<td>.85</td>
<td>.74</td>
</tr>
<tr>
<td>SE</td>
<td>.04</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td>Autism spectrum disorder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>.59</td>
<td>.60</td>
<td>.59</td>
</tr>
<tr>
<td>SE</td>
<td>.04</td>
<td>.07</td>
<td>.06</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>.62</td>
<td>.75</td>
<td>.68</td>
</tr>
<tr>
<td>SE</td>
<td>.02</td>
<td>.04</td>
<td>.03</td>
</tr>
</tbody>
</table>

Figure 1. Proportion of correct discriminations at 15-ms exposure as a
function of stimulus dimension and diagnosis. Error bars are ±1 standard
error of the mean. RD = reading disability.

Figure 2. Proportion of correct discriminations at 30-ms exposure as a
function of stimulus dimension and diagnosis. Error bars are ±1 standard
error of the mean. RD = reading disability.
This study suggests future directions to further specify the exact nature of the observed difference in emotional extraction. First, because our sample consisted of adults, it is important to examine the role of the relative experience that autistic participants have with extracting affective and gender information from faces and potential time points at which their developmental trajectory diverges from typical and control individuals. Children with ASD also show atypical facial reactions to others’ emotional facial expressions, suggesting early differences in responses to others’ emotions (Beall, Moody, McIntosh, Hepburn, & Reed, in press). Examining extraction of emotional information in younger samples is an important next step. Second, this study focused only on responses to facial expressions. Emotional processing involves much more than extraction of emotion from others’ faces. Thus, a complete picture of typical and atypical emotion and empathy processes needs to consider a wide range of interacting processes, including information from nonfacial channels of emotional communication. As this picture develops, these data indicate that consideration of very rapid responses will be necessary to understand how social emotional functioning plays out in the complexities of everyday social interactions.

References


