Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development
Heaton, P.; Reichenbacher, L.; Sauter, D.A.; Allen, R.; Scott, S.; Hill, E.

Published in:
Psychological Medicine

DOI:
10.1017/S0033291712000621

Link to publication

Citation for published version (APA):
Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development

P. Heaton*, L. Reichenbacher, D. Sauter, R. Allen, S. Scott and E. Hill

Goldsmiths College, University of London, London, UK

Background. The results from recent studies suggest that alexithymia, a disorder characterized by impairments in understanding personal experiences of emotion, is frequently co-morbid with autism spectrum disorder (ASD). However, the extent that alexithymia is associated with primary deficits in recognizing external emotional cues, characteristic in ASD, has yet to be determined.

Method. Twenty high-functioning adults with ASD and 20 age- and intelligence-matched typical controls categorized vocal and verbal expressions of emotion and completed an alexithymia assessment.

Results. Emotion recognition scores in the ASD group were significantly poorer than in the control group and performance was influenced by the severity of alexithymia and the psycho-acoustic complexity of the presented stimuli. For controls, the effect of complexity was significantly smaller than for the ASD group, although the association between total emotion recognition scores and alexithymia was still strong.

Conclusions. Higher levels of alexithymia in the ASD group accounted for some, but not all, of the group difference in emotion recognition ability. However, alexithymia was insufficient to explain the different sensitivities of the two groups to the effects of psycho-acoustic complexity on performance. The results showing strong associations between emotion recognition and alexithymia scores in controls suggest a potential explanation for variability in emotion recognition in non-clinical populations.

Received 12 October 2011; Revised 29 February 2012; Accepted 8 March 2012; First published online 4 April 2012

Key words: Alexithymia, autism, emotional vocalizations.

Introduction

Autism spectrum disorder (ASD) covers a range of neurodevelopmental disorders, of biological origin, that are diagnosed on the basis of deficits in social and communication skills and the presence of restricted interests and repetitive behaviours (APA, 2000). Kanner’s first description of autism (Kanner, 1943) made reference to affective difficulties and early group studies of autistic children confirmed poor emotion recognition relative to age- and intelligence-matched typically developing controls (Hobson, 1986; Hobson et al. 1988). However, although the behavioural presentation of ASD clearly encompasses difficulties in understanding emotional cues on faces and in voices, the results from several more recent experimental studies suggest that the severity of emotion processing deficits may be considerably more variable than was originally thought. Indeed, several studies have failed to reveal deficits in emotion processing in ASD relative to matched controls (Grossman et al. 2000; Castelli, 2005; Williams & Happé, 2010). In one such study Jones et al. (2011) presented large samples of adolescents with and without ASD with emotional faces (Ekman & Friesen, 1976) and vocal expressions of emotion (Sauter, 2006). Although the authors observed autism-specific difficulties with identification of surprise, the results showed few group differences, and the error analysis suggested similar confusions across emotion categories in both participant groups.

Factors that influence success and failure on emotion recognition tasks in ASD include stimulus complexity and ecological validity. Law Smith et al. (2010) presented high-functioning adolescents with ASD with the Emotion Recognition Task developed by Montagne et al. (2007). In this task participants are required to make judgements about video clips of actors with facial expressions of anger, sadness, disgust, fear, happiness and surprise that are incrementally morphed with neutral faces. Stimulus morphing

* Address for correspondence: Professor P. Heaton, Goldsmiths College, London, UK.
(Email: p.heaton@gold.ac.uk)
results in faces that vary in emotional intensity, thereby reflecting the highly variable nature of facial emotions experienced in interpersonal encounters. The results from the study showed that, when the stimuli were presented in unmorphed form, deficits in the ASD group were only observed for the emotion category of disgust. For all other emotion categories, deficits emerged when higher levels of neutral were incorporated and the emotions became less intense.

Although differences in results from studies of emotion recognition in ASD are clearly influenced by the type of stimuli used, heterogeneity in ASD samples may also play a role in determining the severity of the emotion processing deficits observed. There is an increasing awareness that ASD is heterogeneous, in terms of both symptom severity and cognitive impairment, and that co-morbidity between ASD and other disabilities is common (Williams et al. 2008). Although estimates of co-occurring intellectual impairment in autism have been as high as 75% (Schalock et al. 2007), more recent work carried out in the UK observed significantly lower rates of intellectual impairment in this group (Williams et al. 2008; Charman et al. 2011). Intellectually able individuals with ASD are better able to correctly categorize emotions than those with significant intellectual impairment (e.g. Loveland et al. 1997) and experiments using simple stimuli to test emotion recognition may observe ceiling effects for these individuals. This may then explain why studies using simple stimuli to compare high-functioning individuals and typically developing controls sometimes fail to reveal significant group differences (e.g. Law Smith et al. 2010).

Another factor that may influence the development of emotion recognition and contribute to the observed variability in emotion recognition skills in ASD is sensory dysfunction, which can be marked at early stages of development (Kern et al. 2006). Karmiloff-Smith (2009) proposed that sensory or attentional abnormalities, present in early infancy, can result in a diverse range of later occurring behavioural abnormalities. In the study by Kern et al. (2006) abnormalities in both auditory and visual modalities were noted and these may impact on the autistic infants’ ability to learn about facial and vocal expressions of emotion. Although the results from the study suggested that sensory difficulties abate with age, the negative effects of early impoverished interpersonal interactions may not be easily compensated at later stages of development.

Recently, there has been a surge of interest in the co-occurrence of ASD and alexithymia. Alexithymia is characterized by difficulties in identifying and describing feelings, difficulties in distinguishing feelings from bodily sensations of emotional arousal, impaired symbolization and a tendency to focus on external events rather than on personal experiences (Nehiam et al. 1976). Although prevalence rates of alexithymia are estimated at around 10% in the typical population (Linden et al. 1995; Salminen et al. 1999), Hill et al. (2004) and Berthoz & Hill (2005) observed slight or severe alexithymia in 85% of a sample of able individuals with ASD. Deficits in mentalizing or in theory of mind (ToM) have been demonstrated in numerous studies of ASD (for a review see Baron-Cohen, 2001) and difficulties in emotional understanding have been linked with these in past studies (Heery et al. 2003). However, alexithymia has been associated with reduced links between physiological arousal and the subjective experience of emotion (McLean, 1949) and a decoupling of these mechanisms might better explain alexithymia in ASD than a deficit in mentalizing ability. Silani et al. (2008) investigated this hypothesis in a functional magnetic resonance imaging (fMRI) study of individuals with ASD and matched controls and concluded that alexithymia in ASD results from an emotion-specific deficit associated with lack of awareness of changes in bodily states, rather than with difficulties in mentalizing.

To date, no published work has addressed the question of whether the types of emotion processing deficits characterized by alexithymia are associated with primary difficulties in understanding expressions of emotions in faces and voices. Although much of the research into emotion perception in ASD has focused on facial displays of basic emotions (e.g. Ekman & Friesen, 1976), these can be expressed both in non-verbal vocal gesture (Scott et al., 1997) and in emotional speech (Murray & Arnott, 1993; Scherer et al. 2001). In the current study we attempted to extend research into alexithymia in ASD and in typical development by looking at associations between scores on the 20-item Toronto Alexithymia Scale (TAS-20; Bagby et al. 1994a, b) and data from a vocal emotion recognition task developed by Sauter et al. (2010). As emotion recognition deficits in intellectually able individuals with ASD seem to emerge more strongly when stimuli are more complex, stimulus complexity will also differ across the two experimental conditions. The null hypotheses for the study are that autism and control groups will not differ in their ability to recognize emotions across simple and complex stimuli, and that emotion recognition scores will be independent of levels of alexithymia.

**Method**

**Participants**

Twenty participants with ASD were recruited through the UK-based National Autistic Society website
(www.autism.org.uk/) and also through research advertisements delivered to specialist social groups and community centres. The group included 15 males and five females with ages ranging between 20 and 67 years. All had been formally diagnosed by a clinical psychologist or psychiatrist in accordance with DSM criteria. Twenty typical controls (15 male and five female) were recruited from higher education colleges and community centres and were matched to the ASD participants for gender, chronological age and intelligence, assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). To control for autistic traits in the control group, all participants completed the 50-item self-report Autism Spectrum Quotient (AQ; Baron-Cohen et al. 2001). The AQ has been shown to have good discriminative validity when compared to diagnostic interviews in clinical settings (Woodbury-Smith et al. 2005). The participants also completed the TAS-20 (Bagby et al. 1994a,b), which assesses emotion processing within three cognitive-affective areas: difficulty identifying feelings, difficulty describing feelings, and externally oriented thinking. The TAS-20 is a reliable and well-validated measure of emotion processing with good psychometric properties (Bagby et al. 1994b), good internal consistency (Cronbach’s $\alpha=0.81$) and high test–retest reliability (Berthoz & Hill, 2005). Severe alexithymia is indicated by a score $>60$, with a borderline range between 52 and 60. Participants’ psychometric data are shown in Table 1.

Procedure and stimuli

Participants were tested individually in a soundproof cubicle. They were asked to complete the AQ, TAS-20 and the WASI before beginning the experimental task. The stimuli used in the experimental task were developed and tested by Sauter (2006) and Sauter et al. (2010). The first set of stimuli (.wav files) were 60 vocal recordings of four actors (two males and two females) expressing one of the six basic emotions (happy, sad, anger, surprise, fear and disgust) described by Ekman & Friesen (1976) non-verbally. Examples of non-verbal stimuli from this set include laughing vocalizations for happiness and crying vocalizations for sadness. The second set of 60 stimuli (.wav files) were vocal recordings of four actors (two males and two females) expressing the same six emotions verbally. The verbal content of these stimuli was three-digit numbers (e.g. five hundred and twenty-three). The average duration of the stimuli was 1.84 s ($\text{S.D.}=0.46$ s). The stimuli were presented in a forced-choice task, in which response alternatives (emotion words: angry, happy, fear, surprise, disgust, sad) were presented on a touch-screen computer.

Participants were required to listen to each of the auditory stimuli and select the emotion word that described the vocal stimulus. There were four practice trials, two for each of the two conditions, followed by two blocks of 60 test trials. To avoid practice or order effects, the two blocks of trials were counterbalanced across participants. Participant responses were coded for accuracy.

Ethics approval for the study was granted by the Goldsmiths College Ethics Committee, following the guidelines laid down by the Economic and Social Research Council (ESRC) and the British Psychological Society.

Results

Inspection of the data showed that the distribution of scores did not depart significantly from normality (skewness and kurtosis both differed from zero by $<2$ S.D.) and Levene’s test for homogeneity of variance was non-significant across diagnostic groups for both simple ($p=0.881$) and complex ($p=0.364$) conditions. The experimental data were then analysed in SPSS version 17 (SPSS Inc., USA) using a $2 \times 2$ mixed ANOVA with group (ASD/controls) as the between-group variable and stimulus type (emotion sounds/emotion speech) as the within-subjects variable.

This revealed a highly significant main effect of group ($F_{1,38}=59.45$, $p<0.001$), with controls achieving higher emotion recognition scores. In addition to the main effect of group, the results from

<table>
<thead>
<tr>
<th>Table 1. Participants’ psychometric data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>ASD group</strong></td>
</tr>
<tr>
<td><strong>Control group</strong></td>
</tr>
<tr>
<td>Age (years), mean (S.D.)</td>
</tr>
<tr>
<td>WASI score, mean (S.D.)</td>
</tr>
<tr>
<td>VIQ</td>
</tr>
<tr>
<td>PIQ</td>
</tr>
<tr>
<td>FIQ</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>TAS total*, mean (S.D.)</td>
</tr>
<tr>
<td>Low</td>
</tr>
<tr>
<td>Borderline</td>
</tr>
<tr>
<td>High</td>
</tr>
</tbody>
</table>

ASD, Autism spectrum disorder; WASI, Wechsler Abbreviated Scale of Intelligence; VIQ, Verbal IQ; PIQ, Performance IQ; FIQ, Full-Scale IQ; AQ, Autism-Spectrum Quotient; TAS, Toronto Alexithymia Scale; S.D., standard deviation.

* Group difference significant at $p=0.001$.
the ANOVA revealed a main effect of stimulus type \((F_{1,38} = 116.69, p < 0.001)\), with better discrimination of simple sounds than more complex emotion speech. The group by stimulus interaction was also significant \((F_{1,38} = 39.51, p < 0.001)\) (Fig. 1).

Four post-hoc t tests with Bonferroni adjustments (modified \(\alpha\) value of \(p = 0.0125\)) showed that the controls made significantly more correct discriminations in both sound \(t[38] = -5.20, p < 0.001\) and speech conditions \(t[38] = -8.35, p < 0.001\) than participants with ASD. Similarly adjusted paired-samples t tests showed that controls obtained significantly higher correct discrimination scores in the sound compared to the speech condition \(t[19] = 3.79, p = 0.001\) and this was also observed for the ASD group \(t[19] = 10.64, p < 0.001\).

Correlations between total emotion recognition scores and TAS scores were significant for the ASD group \((r = -0.66, p = 0.002)\) and the control group \((r = -0.63, p = 0.003)\). Correlations between total emotion recognition scores and full-scale IQ scores were significant for the control group \((r = 0.49, p = 0.028)\) but not for the ASD group \((r = 0.28, p = 0.24)\).

Although the ANOVA showed that both groups obtained lower accuracy scores in the speech than in the sound condition, the extent of this seemed to be greater in the ASD group. Thus, although scores for controls fell by 5% in the speech condition, they rose to 19% for the ASD group. To enable the identification of factors associated with this decrease in correct discrimination scores across conditions, a discrepancy variable was calculated by subtracting the correct speech scores from the correct sound scores. Levene's test for homogeneity of variance was carried out on the discrepancy data and was not significant \((p = 0.082)\). The distribution of the scores did not depart from normality.

An independent t test was carried out on the data and showed that the mean discrepancy score for the ASD group \((\text{mean} = 11.35, \text{s.d.} = 4.77)\) was significantly greater \(t[38] = 6.28, p < 0.001\) than the mean discrepancy score for the control group \((\text{mean} = 3.0, \text{s.d.} = 3.53)\). The correlation between the total discrepancy score and TAS scores narrowly failed to reach significance for the ASD group \((r = 0.43, p = 0.053)\) but was highly significant for the control group \((r = 0.57, p = 0.009)\). The correlations between the total discrepancy scores and full-scale IQ scores were significant for the control group \((r = -0.46, p = 0.043)\) but not for the ASD group \((r = -0.106)\).

The question of whether the effect of complexity was limited to specific emotions or influenced the range of emotion tested in the study was explored. Discrepancy scores (simple–complex) were calculated for the six different emotions. An ANOVA carried out on these data failed to reveal a significant diagnosis by emotion type interaction \((F_{5,190} = 1.56)\), suggesting that the two groups display similar response profiles across the set of emotions tested.

**Discussion**

In contrast to recent studies suggesting that emotion recognition is unimpaired in ASD, the results of the current study show a significant main effect of group with poorer emotion recognition in the ASD group. In the non-verbal emotion sounds condition the ASD group correctly identified 79% of the stimuli compared with 90% for controls. Very similar identification rates have been observed when autistic and control participants have been asked to identify emotions in Ekman faces (controls 92.5%, ASD 76%; Pelphrey et al. 2004). This suggests that emotion processing deficits in ASD are similarly marked across modalities. Also consistent with previous findings (Law Smith et al. 2010), we observed that the nature of the emotion stimuli was an important factor in predicting emotion recognition performance in ASD. Previous research with typical participants has revealed poorer recognition performance for emotions in speech than in non-verbal emotion sounds (e.g. Hawk et al. 2009), and in the current study both groups obtained significantly lower scores on the emotional speech condition than on the non-verbal emotional sound condition. However, the impact of increasing complexity was far greater for participants with ASD than for controls. Hence, although the control group’s correct identification scores fell by 5% in the emotional speech condition, they rose to 19% for the ASD group.

There are several possible reasons why the participants with ASD showed such a large decrease in performance on the speech compared with the non-verbal sounds. It may be that the speech task involves two concurrent types of decoding (emotion and speech) whereas the non-verbal vocalizations only require the listener to decode the emotion cues. When listening to the affective speech stimuli, the listener must disregard the linguistic information and focus on the
affective cues. Non-verbal vocalizations are also far richer in affective cues than speech vocalizations and it is not surprising that this impacts on recognition performance. Several studies have shown unimpaired categorization of emotion cues in musical stimuli in ASD (e.g. Heaton et al. 2008) and this may be subserved by the same mechanisms underlying the relatively preserved (approximately 80% correct) understanding of affective cues in non-verbal emotional sounds.

The analysis of the alexithymia data extended existing work on ASD by showing a strong association between severity of alexithymic symptoms and recognition of external emotion cues. Hill et al. (2004) and Berthoz & Hill (2005) showed greatly increased levels of alexithymia in the ASD populations they studied, and 70% of the current ASD sample, compared with 5% of controls, showed borderline and/or high levels of alexithymia. However, it is important to note that levels of alexithymia within the ASD group showed considerable variability (see Table 1). Several recent experimental studies (Silani et al. 2008; Bird et al. 2010, 2011) have identified co-morbid alexithymia in subgroups of individuals with ASD, and Bird et al. (2010) observed the same pattern of brain activation in the left anterior insula in response to emotion-evoking stimuli in typical controls and ASD participants without alexithymia. Our findings, showing that high levels of alexithymia are also associated with increased difficulties in emotion recognition, highlight the importance of including alexithymia measures in experimental studies of emotion recognition in ASD. Indeed, the current lack of consensus about the prevalence and severity of these and other difficulties associated with ASD (see Bird et al. 2011, for work on eye-gaze in ASD and co-morbid alexithymia) may reflect a past failure to control for co-occurring alexithymia in experimental studies. Even though estimates of alexithymia in the general population are low (Linden et al. 1995; Salminen et al. 1999), the inclusion of alexithymia measures in studies of typically developing emotion recognition may also be warranted. Although the large majority of control participants (95%) obtained low alexithymia scores, the association between these and the scores on the emotion recognition task showed the same strong association that had been observed for the ASD group whose alexithymia scores were significantly higher. This suggests that alexithymia, even at subclinical levels, explains variability in emotion recognition within typical and also clinical populations.

The discrepancy score measured the decrease in recognition performance with linguistic stimuli. The magnitude of the discrepancy score was relatively small for controls, but it was strongly associated with levels of alexithymia. Correlations between total emotion scores, discrepancy scores and measures of intelligence were also significant for the control group. Thus, typical individuals with higher TAS scores and lower IQ showed the largest decrement in performance when the emotion cues were more difficult to access (emotion speech condition). Although the discrepancy scores were significantly greater for the ASD group than for the control group, they did not correlate with IQ measures or severity of alexithymia. Thus, although alexithymia scores strongly correlated with total emotion recognition scores for the ASD group, alexithymic severity seemed to be independent of auditory processing abnormalities revealed by discrepancy scores for the ASD group. The results from several studies have revealed abnormalities in cortical voice processing in adults with ASD (e.g. Gervais et al. 2004), and Samson et al. (2011) have outlined a neural complexity hypothesis to account for enhanced and impoverished perceptual discrimination across different classes of auditory stimuli. Emotion processing difficulties in adults with ASD are widely held to be developmental in origin (e.g. Hobson, 1993). Marcelli et al. (2000) have shown that mothers mimic their infants’ expressions of emotions, and this enables the typically developing infant to learn about the emotions of self and of others. For the autistic individual, early inattention to such communicative acts may result in both an inability to understand other peoples’ expressions of emotion and alexithymia (Allen & Heaton, 2010). However, it is important to note that the developmental sequelae of early inattention to social stimuli encompasses a broad range of abnormalities that may extend beyond emotion processing deficits. For example, Kuhl et al. (2005) have shown that autistic children who exhibit decreased attention to infant-directed speech fail to develop specialized neural mechanisms for processing speech stimuli. Although the ASD adults in the current study performed at significantly lower levels than the controls on the vocal gestures condition, the effect was greatly increased when demands on voice processing mechanisms were also increased. It is therefore plausible to suggest that, for intellectually able adults with ASD, their current degree of perceptual difficulty, in addition to their developmental history, will determine the severity of their emotional speech processing deficits.

Acknowledgements

We are grateful to the individuals who participated in our study and thank our editor and reviewers for their constructive comments.
References


