Differences Between Autistic and Non-autistic Adults in the Recognition of Anger From Dynamic Expressions Remain After Controlling for Alexithymia.

Connor Tom Keating (✉ ck655@bham.ac.uk)  
University of Birmingham  https://orcid.org/0000-0001-5906-1789

Sophie L Sowden  
University of Birmingham

Dagmar S Fraser  
University of Birmingham

Jennifer L Cook  
University of Birmingham

Research

Keywords: Autism Spectrum Disorder, Facial Expression, Emotion Recognition, Movement Kinematics, Alexithymia

DOI: https://doi.org/10.21203/rs.3.rs-88768/v3

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

A burgeoning literature suggests that alexithymia, and not autism, is responsible for the difficulties with static emotion recognition that are documented in the autistic population. Here we investigate whether alexithymia can also account for difficulties with dynamic facial expressions. Autistic and control adults (N=60) matched on age, gender, non-verbal reasoning ability and alexithymia, completed an emotion recognition task, which employed dynamic point light displays of emotional facial expressions that varied in speed and spatial exaggeration. The ASD group exhibited significantly lower recognition accuracy for angry, but not happy or sad, expressions with normal speed and spatial exaggeration. The level of autistic, and not alexithymic, traits was a significant predictor of accuracy for angry expressions with normal speed and spatial exaggeration.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder, characterized by difficulties in social communication, and restricted and repetitive interests (American Psychiatric Association, 2013). Since the ability to infer emotion from facial expressions is important for social interaction, emotion recognition has long been suspected as a difficulty in ASD (Hobson, 1986). However, whilst many studies suggest a disparity in the facial emotion recognition ability of autistic and neurotypical individuals (Ashwin, Chapman, Colle, & Baron-Cohen, 2006; Dziobek, Bahnemann, Convit, & Heekeren, 2010; Lindner & Rosén, 2006; Philip et al., 2010), there have been inconsistent findings, ranging from no differences between these individuals to large disparities (see Harms et al., 2010, Keating & Cook, 2020, Uljarevic & Hamilton, 2013 for reviews). Consequently, the question of whether autistic individuals exhibit atypical facial emotion recognition has been debated for over 30 years.

The most recent contributions to this debate claim that it is not autism per se that is linked to emotion recognition atypicalities but rather alexithymia (Bird & Cook, 2013; Kinnaird, Stewart, & Tchanturia, 2019; Oakley, Brewer, Bird, & Catmur, 2016; Poquérusse, Pastore, Dellantonio, & Esposito, 2018). Alexithymia is a subclinical condition, characterized by difficulties identifying and expressing emotions (Nemiah, Freyberger & Sifneos, 1976), which is often comorbid with ASD (in the neurotypical population the incidence of alexithymia is 13% (Salminen, Saarijärvi, Äärelä, Toikka, & Kauhanen, 1999) in autistic populations the incidence is estimated at 40-65% (Berthoz & Hill, 2005; Bird & Cook, 2013; Hill, Berthoz, & Frith, 2004)). Cook and colleagues (2013) demonstrated that continuous measures of alexithymic, but not autistic, traits are predictive of poorer facial emotion recognition from static face images. Furthermore, when groups are matched in terms of alexithymia, autistic and neurotypical adults perform comparably with respect to the recognition of emotion (Cook et al., 2013). Similarly, Milosavljevic et al., (2016) demonstrated lower emotion recognition scores - again from static face images - for autistic adolescents high in alexithymia relative to those low in alexithymia. Consequently, Bird & Cook (2013) propose ‘the alexithymia hypothesis’: autistic individuals’ difficulties in emotion-processing, including facial emotion recognition, are caused by co-occurring alexithymia not ASD.
To date, the majority of studies that have tested “the alexithymia hypothesis” have focused on the recognition of emotion from static face images and have thus overlooked the inherently dynamic nature of facial expressions (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004). Importantly, dynamic faces carry both spatial information about the configuration of facial features relative to each other and information about the kinematics (e.g. speed) of movement of facial features (Dobs, Bülthoff, & Schultz, 2018). Recent developments in the face processing literature emphasize the importance of both kinematic and spatial cues in neurotypical facial emotion recognition.

Most notably, Sowden and colleagues (in press.) manipulated point-light face (PLF) stimuli (a series of white dots on a black background that convey biological motion and eliminate contrast, texture, colour and luminance cues) such that expressions of happiness, anger and sadness were reproduced at 50%, 100% and 150% of their normal speed, and at 50%, 100% and 150% of their normal range of spatial movement (e.g. at the 150% level a smile would be 50% bigger / more exaggerated than normal). Sowden and colleagues (in press.) found that the emotion recognition accuracy of neurotypical participants was modulated as a function of both spatial and kinematic manipulation. Specifically, when expressions were reduced in their speed and spatial extent (i.e. at the 50% level), participants were less accurate in their labelling of angry and happy expressions and more accurate for sad expressions. Conversely, when expressions were played with exaggerated spatial movement and greater speed (i.e. at the 150% level), participants displayed higher accuracy for angry and happy expressions and lower accuracy for sad expressions (Sowden et al., in press.). Thus, accuracy for labelling high arousal emotions (happy and angry) is improved when the stimulus is faster and more spatially exaggerated, whereas labelling of low arousal emotions (sad) is impaired. Recent literature therefore highlights that, for neurotypical individuals, both spatial and kinematic facial cues contribute to emotion recognition accuracy.

Since dynamic information is particularly important in real life processing of facial expressions (Krumhuber, Kappas, & Manstead, 2013), if the alexithymia hypothesis is to explain functional, everyday challenges, it must be the case that the co-occurrence of alexithymia can account for recognition difficulties with respect to both spatial and kinematic aspects of facial emotional expressions. That is, when autistic and non-autistic groups are matched in terms of alexithymia there should be no differences between the groups in the processing of either spatial or kinematic cues with respect to emotion recognition. Facial expression recognition from static stimuli relies on the processing of spatial cues but overlooks the contribution of kinematic cues. To the best of our knowledge, there are no studies that have investigated autistic versus neurotypical recognition of emotion from dynamic face stimuli whilst controlling for the influence of alexithymia. There are, however, some studies that have compared autistic and neurotypical processing of dynamic facial expressions without controlling for alexithymia. For example, Sato and colleagues (Sato, Uono, & Toichi, 2013) demonstrated that for neurotypical adults reducing the speed of movement of facial morph stimuli[2] reduced naturalness ratings, however, for autistic adults the effect of speed on naturalness ratings was significantly weaker. Sato and colleagues’ results thus demonstrate differences, between autistic and non-autistic adults, in the effects of manipulating facial kinematics. To the best of our knowledge, only one study has examined the contribution of autistic and alexithymic traits to dynamic emotion recognition (Ola & Gullon-Scott, 2020).
The findings of this study support the alexithymia hypothesis: high alexithymic, but not autistic, traits were associated with less accurate facial expression recognition (Ola & Gullon-Scott, 2020). However, this study, conducted by Ola and Gullon-Scott, has two important limitations. First, only female participants were recruited. Since autistic males comprise three quarters of the ASD population (Loomes, Hull, & Mandy, 2017), and likely differ in behavioural phenotype (Ketelaars, In'T Velt, Mol, Swaab, & Van Rijn, 2016; Rivet & Matson, 2011), one must be cautious about extrapolating the findings to autistic males. Second, Ola and Gullon-Scott did not recruit a non-autistic control group. Consequently, the authors were not able to explore whether autistic versus non-autistic group differences in dynamic emotion recognition remain after controlling for alexithymia. That is, although Ola and Gullon-Scott were able to show that some difficulties with emotion recognition from dynamic stimuli were associated with alexithymia, one cannot conclude from this study that there are no differences with respect to emotion recognition from dynamic stimuli that are specifically associated with ASD.

The current study employed the paradigm developed by Sowden and colleagues (Sowden et al., in press.) to test the hypothesis that when autistic and non-autistic groups are matched in terms of alexithymia there will be no differences, between the groups, in the processing of either spatial or kinematic cues with respect to emotion recognition. More specifically, male and female autistic adults and neurotypical controls rated the emotion expressed by PLF stimuli that had been manipulated such that expressions of happiness, anger and sadness were reproduced at 50%, 100% and 150% of their normal speed and spatial extent. The groups were matched in terms of their scores on a self-report measure of alexithymia. We predicted that emotion recognition accuracy would be affected by both kinematic and spatial manipulation and that these effects would not interact with group, but rather that Bayesian statistics would provide support for the null hypothesis that the groups perform comparably. We further predicted that the effects of spatial and kinematic manipulation on emotion recognition accuracy would covary with scores on the self-report alexithymia measure.

Footnote:

[1] 'Disability-first' terminology is used throughout in line with the majority preference expressed in a survey of the autistic community (Kenny et al., 2016)

[2] Facial morph stimuli were constructed by successively presenting 26 images from a neutral (0%) to full emotional (100%) expression with an increase of 4% in emotion from one image to the next. By presenting the images in this way, it gave the illusion of a dynamic emotional expression. The speed of playback was then manipulated to allow the researchers to test their hypotheses.

Method

Participants
The chosen sample size is based on an *a priori* power analysis conducted using GLIMMPSE (Kreidler et al., 2013), which focused on replicating the primary results in the control group (the emotion*spatial and emotion*kinematic interactions). Using data from Sowden et al., (in press), 8 participants are required in the control group in order to have 95% power to detect an effect size of $0.70 (\eta^2)$ at alpha level 0.01 for the emotion*spatial interaction. Moreover, 11 participants are required in the control group in order to have 95% power to detect an effect size of $0.53 (\eta^2)$ for the emotion*kinematic interaction at alpha level 0.01. However, Button et al., (2013) argue that effect size estimates are commonly inflated (“the winners curse”), and that there is “a common misconception that a replication study will have sufficient power to replicate an initial finding if the sample size is similar to that in the original study”. Accordingly, we planned to recruit a larger number of participants (N=30 per group; almost triple the largest sample size generated in our power calculations), in order to obtain adequate power. We pre-registered this sample size via the Open Science Framework (https://osf.io/kpefz).

60 individuals, 31 with an ASD diagnosis and 29 neurotypical controls, participated in the study (See Supplementary Information A for race/ethnicity information). Participants were matched for age, gender, non-verbal reasoning (NVR; as measured by the Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019)) and alexithymia (as measured by the 20-item Toronto Alexithymia Scale (TAS-20; Bagby, Parker, & Taylor, 1994). The ASD group had significantly higher Autism Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) scores (see Table 1). The level of autistic characteristics of those in the ASD group was assessed using the Autism Diagnostic Observation Schedule (version 2 (ADOS-2; Lord et al., 2012). The mean total ADOS-2 score in the ASD group was 10.32 (see Supplementary Information B for information on the quantity of participants that met criteria for diagnosis). The MaRs-IB was used to match participants on the basis that the PLF task relies on non-verbal reasoning ability and, with respect to participant matching, task specific measures of intelligence/ability have been argued to be more appropriate than general measures (Mottron, 2004). A total of four participants (three in the ASD group and one in the control group) had AQ or TAS-20 scores over two standard deviations from their group mean. Since the general pattern of results was unaffected by their removal, these participants are included in the final analysis.

Table 1. Means, standard deviations and group differences of participant characteristics. In the central columns, means are followed by standard deviations in parentheses.

|                         | Control group (n=29) | ASD group (n=31) | Significance |
|-------------------------|----------------------|------------------|--------------|
| Gender                  | 11 Female, 17 Male, 1 Other | 14 Female, 16 Male, 1 Other | p=.850 |
| Age                     | 28.81 (9.54)         | 30.14 (9.08)     | p=.581 |
| NVR                     | 62.91 (15.17)        | 57.05 (17.90)    | p=.178 |
| TAS-20                  | 55.66 (13.57)        | 59.74 (13.14)    | p=.241 |
| AQ                      | 19.86 (7.44)         | 32.52 (10.21)    | p<.001 |
| ADOS-2                  | N/A                  | 10.32 (4.76)     | N/A          |
Twenty-two of the 31 ASD participants were recruited via an existing autism research database kept by the Birmingham Psychology Autism Research Team (B-PART). The control and remaining nine ASD participants were recruited via social media (Facebook and Twitter) and Prolific – an online recruitment platform. All participants in the ASD group had previously received a clinical diagnosis of ASD from a qualified clinician.

**Materials and stimuli**

**PLF stimuli**

The PLF task was an adapted version of that developed by Sowden and colleagues (Sowden et al., in press.) which was re-programmed in Gorilla.sc (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020) to facilitate online testing. The same instructions, stimulus videos, and rating scales were used as in the original study. The stimulus videos comprised dynamic PLF stimuli, created from videos of four actors (two male, two female) verbalising sentences (“My name is John and I’m a scientist”) whilst posing three target emotions (angry, happy and sad). PLFs were adapted (see Sowden et al., for further detail) to achieve three spatial movement levels, ranging from decreased to increased spatial movement (S1: 50% spatial movement; S2: 100% spatial movement; S3: 150% spatial movement), and three kinematic levels, ranging from reduced to increased speed (K1: 50% original stimulus speed; K2: 100% original stimulus speed; K3 – 150% of the original stimulus speed). Consequently, there were 9 manipulations per emotion (e.g. (1) S1, K1, (2) S2, K1, (3) S3, K1, (4) S1, K2, (5) S2, K2, (6) S3, K2, (7) S1, K3, (8), S2, K3, (9) S3, K3).

**Autistic traits**

The autistic traits of all ASD and control participants were assessed via the 50-item Autism Quotient (Baron-Cohen et al., 2001). This self-report questionnaire is scored on a range from 0 to 50, with higher scores representing higher levels of autistic characteristics. The AQ assesses five different domains relevant for ASD traits (attention switching, attention to detail, communication, social skill and imagination). The AQ has been widely used in both the general and the autistic population (Ruzich, Emily Allison, Carrie Smith & Watson, 2015; Ruzich et al., 2016), and has strong psychometric properties, including internal consistency (α ≥ 0.7) and test-retest reliability (r ≥ 0.8; Stevenson & Hart, 2017)

**Alexithymia**

Alexithymia was measured via the 20-item Toronto Alexithymia Scale (Bagby et al., 1994). The TAS-20 comprises 20 items rated on a five-point Likert scale (ranging from 1, strongly disagree, to 5, strongly agree). Total scores on the TAS-20 can range from 20 to 100, with higher scores indicating higher levels of alexithymia. The TAS-20 is the most popular self-report tool for alexithymia and boasts good internal consistency (α ≥ 0.7) and test-retest reliability (r ≥ 0.7) (Bagby et al., 1994; Taylor, Bagby, & Parker, 2003).

**Non-verbal reasoning**
Non-verbal reasoning was assessed via the Matrix Reasoning Item bank (MaRs-IB; Chierchia et al., 2019). Each item in the MaRs-IB consists of a 3 x 3 matrix. Eight of the nine available cells in the matrix are filled with abstract shapes, and one cell in the bottom right-hand corner is left empty. Participants are required to complete the matrix by selecting the missing shape from four possible options. In order to correctly identify the missing shape, participants have to deduce relationships between the shapes in the matrix (which vary in shape, colour, size and position). When participants select an answer, they move on to the next item. If participants do not provide a response within 30 seconds, they continue to the next item without a response. The MaRs-IB assessment lasts 8 minutes regardless of how many trials are completed. There is a total of 80 different items in the MaRs-IB, however participants are not required (or expected) to complete all 80 items within the 8 minutes. If a participant completed all 80 items within 8 minutes, the items were presented again but the responses to these were not analysed (following the procedure established by Chierchia and Fuhrmann et al., 2019). The MaRs-IB has been shown to have acceptable internal consistency (Kuder-Richardson 20 \( \geq \) 0.7) and test-retest reliability (\( r \geq 0.7; \) Chierchia et al., 2019).

Procedure

Following a pre-registered design (see https://osf.io/kpefz), participants first completed the questionnaires (demographics followed by AQ, followed by TAS-20) and then moved on to the PLF task. Each trial in this task (see Fig. 1) began with the presentation of a stimulus, which comprised a silent PLF video of an actor expressing one of 3 emotions whilst saying a sentence at one of the 3 spatial and 3 kinematic levels. After watching the video, participants were asked to rate how angry, happy and sad the person was feeling. Participants made their ratings on a visual analogue scale, with one end representing 'Not at all angry/happy/sad' and the opposite end representing 'Very angry/happy/sad'. Individuals were asked to make ratings for all three target emotions (angry, happy and sad) on scales, which were presented on screen in a random order, after each PLF video. Each trial took approximately 25 seconds to complete. Participants completed 3 practice trials (at the S2 and K2 level) and then 108 randomly ordered experimental trials (12 per condition) across three blocks. Participants were invited to take a break between blocks.

Following PLF task completion participants completed the Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019).

Participants completed all tasks online using Google Chrome or Mozilla Firefox on a computer or laptop. The frame rate (in frames per second; FPS) of their devices was measured to ensure that the quality/fluidity of the stimulus videos was not degraded. All participants’ frame rates were 60 FPS or higher with one exception at 50 FPS. When we ran all analyses with and without the 50 FPS participant, treating them as a potential outlier, the pattern of results was unaffected. Therefore, this participant was included in all analyses.

Statistical Analysis
The three emotion rating responses for each trial were transformed into scores from 0 to 10 (with 0 representing a response of 'Not at all' and 10 representing 'Very') to 3 decimal places. Emotion recognition accuracy scores were calculated as the correct emotion rating minus the mean of the two incorrect emotion ratings[3]. For instance, for a trial in which an angry PLF was presented, the mean ratings of the two incorrect emotions (happy and sad) were subtracted from the correct emotion (angry).

To test our first hypothesis, we submitted these accuracy scores to a 2 x 3 x 3 x 3 Analysis of Variance (ANOVA) with the between-subjects factor group (ASD, control) and the within-subjects factors emotion (happy, angry, sad), stimulus spatial level (S1, S2, S3), and stimulus kinematic level (K1, K2, K3). To test our second hypothesis, we applied a sqrt transformation to all ordinal factors of interest (age, NVR, AQ, TAS-20), computed z-scores for the transformed data, and submitted the transformed z-scored data, along with the nominal predictor gender, to multiple regression analyses. The effect of the spatial manipulation (defined as the difference in accuracy between S3 and S1), the effect of the kinematic manipulation (defined as the difference in accuracy between K3 and K1), mean recognition accuracy and finally accuracy for angry videos at the normal level (S2, K2) were used as the DVs for each of these analyses. For all analyses, we used a p = .05 significance threshold to determine whether to accept or reject the null hypothesis. The frequentist approach was supplemented with the calculation of Bayes Factors, which quantify the relative evidence for one theory or model over another. For all Bayesian analyses, we followed the classification scheme used in JASP (Lee & Wagenmakers, 2014) to classify the strength of evidence given by Bayes factors, with BF$_{10}$ between one and three considered as weak evidence, between three and ten as moderate evidence and greater than ten as strong evidence for the alternative hypothesis respectively.

Footnote:

[3] Many of the studies that have investigated the emotion recognition ability of autistic individuals have used forced-choice paradigms in which there is a binary (correct; 1, or incorrect; 0) accuracy score for each trial. In order to facilitate comparison of our results to those studies, we also completed a binary accuracy analysis, which yielded similar results (see Supplementary Information C). In this analysis, for each trial, participants scored 1 when they gave the highest rating to the correct emotion, and 0 when they rated either of the incorrect emotions higher than the correct emotion.

Results

Our primary hypothesis was that emotion recognition accuracy would be affected by both kinematic and spatial manipulation and that these effects would not interact with group. To test this hypothesis we conducted a mixed 2 x 3 x 3 x 3 ANOVA with the between-subjects factor group (ASD, control) and the within-subjects factors emotion (happy, angry, sad), stimulus spatial level (S1, S2, S3), and stimulus kinematic level (K1, K2, K3). This analysis revealed a significant main effect of emotion [F(2,116) = 17.79,
p < .001, $\eta_p^2 = .24$, BF$_{10} = 1.03e^{15}$; see Supplementary Information D], a main effect of spatial level [F(2,116) = 259.57, p < .001, $\eta_p^2 = .82$, BF$_{10} = 9.05e^{57}$; see Supplementary Information D] which was qualified by an emotion x spatial interaction [F(4,232) = 88.42, p < .001, $\eta_p^2 = .60$, BF$_{10} = 7.53e^{58}$], and an emotion x kinematic interaction [F(4,232) = 53.90, p < .001, $\eta_p^2 = .48$, BF$_{10} = 1.90e^{20}$]. Furthermore, this analysis revealed a significant four-way emotion x spatial x kinematic x group interaction [F(8,464) = 2.438, p < .05, $\eta_p^2 = .04$, BF$_{10} = 0.07$]. Note that no kinematic x group interaction was found [p = .538, BF$_{10} = 0.02$], suggesting that autistic and control participants exhibit similar patterns of accuracy across the kinematic levels. Below, in order to shed light on the effects of the spatial and kinematic manipulations, we first unpack the emotion x kinematic and emotion x spatial interactions. Subsequently we fully unpack the emotion x spatial x kinematic x group interaction.

In line with Sowden et al., (in press.), we observed an emotion x spatial interaction [F(4,232) = 88.42, p < .001, $\eta_p^2 = .60$, BF$_{10} = 7.53e^{58}$]. Post-hoc repeated measures ANOVAs revealed that whilst the effect of the spatial manipulation was present for all three emotions (all F > 7.00, all p < .01), the direction of the effect varied between high and low arousal emotions: recognition scores for angry and happy videos were highest for 150% spatial extent (S3) [angry mean (Standard Error of the Mean; SEM) = 5.21(.21); happy mean(SEM) = 5.70(.24)], followed by 100% spatial extent (S2) [angry mean(SEM) = 3.15(.22); happy mean(SEM) = 4.75(.23)], and finally 50% spatial extent (S1) [angry mean SEM) = 0.53(.22); happy mean(SEM) = 2.10(.25)]. In contrast, for sad videos, recognition scores were highest for S1 [sad mean(SEM) = 3.50(.22)], lowest for S3 [sad mean(SEM) = 2.78(.22)] and intermediate for S2 [sad mean(SEM) = 3.15(.20); Fig. 2]. This pattern matches the results reported by Sowden et al., (in press.) for neurotypical participants.

In addition, our analysis identified an emotion x kinematic interaction [F(4,232) = 53.90, p < .001, $\eta_p^2 = .48$, BF$_{10} = 1.90e^{20}$]. Whilst there was a main effect of the kinematic manipulation for all three emotions (all F > 20, all p < .001), the direction of the effect differed between high and low arousal emotions. For angry and happy videos, emotion recognition improved with increasing speed [angry: K1 mean(SEM) = 2.28(.19); K2 mean(SEM) = 2.87(.19); K3 mean(SEM) = 3.73(.23); happy: K1 mean(SEM) = 3.50(.23); K2 mean(SEM) = 4.50(.22); K3 mean(SEM) = 4.55(.21)]. For sad videos, emotion recognition improved as speed decreased [K3 mean(SEM) = 2.03(.19); K2 mean(SEM) = 3.21(.22); K1 mean(SEM) = 4.18(.23); Fig. 2 Bottom panel]. This pattern of results also matches the findings from Sowden et al., (in press.) for neurotypical participants.

In order to unpack the significant four-way interaction, we conducted post-hoc 2 x 3 x 3 (group, emotion, kinematic) ANOVAs for each spatial level. This analysis revealed a significant emotion x kinematic x group interaction at the S2 [F(4,232) = 4.53, p < 0.01, $\eta_p^2 = .07$, BF$_{10} = 5.92$] but not S1 [p = .265, BF$_{10} = 0.09$] or S3 [p = .208, BF$_{10} = 0.09$] level. To unpack this emotion x kinematic x group interaction at the S2 level, we conducted separate post-hoc ANOVAs for each kinematic level at the 100% (S2) spatial level. This analysis revealed a significant emotion x group interaction at the K2 [F(2,116) = 6.48, p < .01, $\eta_p^2 = .10$, BF$_{10} = 17.09$] but not K1 [p = .244, BF$_{10} = 0.32$] or K3 [p = .082, BF$_{10} = 0.82$] level. Bonferroni-corrected
post-hoc independent sample t-tests revealed that control, relative to ASD, participants had higher accuracy for angry videos at the 100% spatial (S2) and speed (K2) level \( [t(58) = 2.78, p_{\text{bonf}} < .05, \text{mean difference} = 1.48, \text{BF}_{10} = 6.09; \text{Fig. 3}] \). There were no significant group differences in emotion recognition accuracy for happy \( [p = .757, \text{BF}_{10} = 0.27] \) or sad \( [p = .085, \text{BF}_{10} = 0.93] \) videos at the S2K2 level. Thus, the groups significantly differed in accuracy for angry PLFs that were not spatially or kinematically manipulated.

To further unpack the emotion x kinematic x group interaction at the S2 level, we conducted separate post-hoc ANOVAs for each emotion at the S2 level. This analysis identified a significant kinematic x group interaction for angry \( [F(2,116) = 4.59, p < .05, \eta^2_p = .07, \text{BF}_{10} = 3.49] \) but not happy \( [p = .070, \text{BF}_{10} = 0.95] \) or sad \( [p = .123, \text{BF}_{10} = 0.53] \) PLFs. Therefore, for angry videos at the normal spatial level, the effect of the kinematic manipulation varied as a function of group. Bonferroni-corrected paired sample t-tests demonstrated that whilst the control group exhibited increasing accuracy across all kinematic levels \( [\text{K1-K2: } t(28) = -4.31, p_{\text{bonf}} < .001, \text{mean difference} = -1.62, \text{BF}_{10} = 153.77; \text{K2-K3: } t(28) = -2.86, p_{\text{bonf}} < .05, \text{mean difference} = -0.95, \text{BF}_{10} = 5.52] \), the ASD group only showed improvement from the K2 to K3 \( [t(30) = -3.46, p_{\text{bonf}} < .01, \text{mean difference} = -1.16, \text{BF}_{10} = 21.10] \) and not K1 to K2 \( [p = .865, \text{BF}_{10} = 0.19; \text{Fig. 4}] \).

Furthermore, the groups did not significantly differ at K1 \( (F(1,58) = 1.18, p > .05) \) or K3 \( (F(1,58) = 3.53, p > .05) \) but at K2, controls out-performed autistic participants \( (F(1,58) = 7.75, p < 0.01, \eta^2_p = .12) \). These results suggest that, whilst controls improved in their accuracy for angry PLF stimuli across each level of increasing kinematic manipulation, for autistic participants, only the most extreme (K3) level of the kinematic manipulation resulted in an accuracy boost.

**Multiple Regression Analyses**

Our second hypothesis was that variation in emotion recognition accuracy would covary, not with ASD symptomatology but with scores on the self-report alexithymia scale (TAS-20). To test whether autistic or alexithymic traits were predictive of the effect of the spatial and kinematic manipulations, we conducted two multiple regression analyses. For the first analysis, we used the effect of spatial manipulation (defined as the difference in accuracy between S3 and S1) as the dependent variable (DV) and AQ and TAS-20 as predictor variables. This analysis resulted in a non-significant model overall \( [F(2,57) = .87, p = .425] \), neither AQ \( [\text{standardized } \beta = -.17, t(57) = -1.10, p = .274] \) nor TAS-20 \( [\text{standardized } \beta = .19, t(57) = 1.20, p = .236] \) were significant predictors of the effect of the spatial manipulation. In the second analysis, we used the effect of the kinematic manipulation (defined as the difference in accuracy between K3 and K1) as the DV and AQ and TAS-20 as predictors. Again, this analysis resulted in a non-significant model \( [F(2,57) = 1.63, p = .206] \), neither AQ \( [\text{standardized } \beta = .20, t(57) = 1.33, p = .189] \) nor TAS-20 \( [\text{standardized } \beta = .05, t(57) = .32, p = .752] \) were significant predictors of the effect of the kinematic manipulation. We then conducted a third multiple regression with mean emotion recognition accuracy (across all trials) as the DV. Once again, neither AQ \( [\text{standardized } \beta = -.19, t(57) = -1.24, p = .220] \) nor TAS-20 \( [\text{standardized } \beta = .12, t(57) = .81, p = .424] \) were significant predictors of mean recognition accuracy and the overall model did not explain a significant amount of variance in the data \( [F(2,57) = .78, p = .461] \). To explore the
possibility that only extreme scores on the TAS-20 predict performance, we compared mean accuracy for alexithymic (i.e. TAS-20 $\geq 61$) and non-alexithymic (i.e. TAS-20 $\leq 51$) participants (according to the cut-off scores outlined by Bagby, Taylor and Parker; 31), excluding ‘possibly alexithymic’ individuals. An independent samples t-test confirmed that there was no significant difference in mean accuracy between these groups ($t(48) = -18, p = .861$, mean difference = -.05, $BF_{10} = 0.29$).

Finally, building on our previous observation that the ASD and control groups differed in accuracy for angry videos at the normal (100%) spatial and speed level we conducted a multiple regression analysis to identify the extent to which autistic and alexithymic traits were predictive of accuracy for angry videos at the S2 and K2 levels. This analysis revealed that autistic [standardized $\beta = -.44$, $t(57) = -3.05$, $p < .01$], but not alexithymic [standardized $\beta = .22$, $t(57) = 1.54$, $p = .130$], traits were predictive of accuracy for angry videos at the normal spatial and speed level [overall model statistics: $F(2, 57) = 4.67$, $p < .05$, $R^2 = .141$]. Bayesian analyses revealed that AQ [$BF_{\text{inclusion}} = 4.230$] was over 16 times more likely to be included in a model to predict accuracy for angry videos at the normal spatial and speed level than alexithymic traits [$BF_{\text{inclusion}} = 0.263$].

In order to ensure that AQ is not just a significant predictor of accuracy for angry expressions at the normal spatial and speed level due to variation across other co-variables (e.g. age, gender, and non-verbal reasoning), we completed an additional three-step forced entry hierarchical regression analysis following the procedures of Cook et al., (2013). In the first step, the demographic variables (gender, age and NVR) were entered into the model, which overall accounted for 16% of the variance in accuracy at the S2K2 level [$F(3,56) = 3.56$, $p < .05$, $R^2 = .160$]. Importantly, of the three demographic variables, only NVR was a significant predictor of accuracy for angry videos at the normal spatial and speed level [standardized $\beta = .35$, $t(56) = 2.79$, $p < .01$] (and not gender [standardized $\beta = .15$, $t(56) = 1.20$, $p = .233$] or age [standardized $\beta = -.01$, $t(56) = -.06$, $p = .950$]). In the second step, AQ was added [standardized $\beta = -.36$, $t(55) = -3.13$, $p < .01$], producing a statistically significant $R^2$ change [$F \text{ change (1, 55)} = 9.80$, $p < .01$, $R^2 \text{ change} = .127$]. Finally, when TAS-20 was entered into the model, the analysis revealed it was not a significant predictor of accuracy for angry videos at the normal level [standardized $\beta = .17$, $t(54) = 1.26$, $p = .214$] and resulted in a non-significant $R^2$ change [$F \text{ change (1, 54)} = 1.58$, $p = .214$, $R^2 \text{ change} = .020$; see Table 2.]. Hence, this analysis demonstrated that AQ (and not TAS-20) was a significant predictor of accuracy for angry videos at the normal level (S2, K2) even after age, gender and NVR have been accounted for.

These analyses suggest that alexithymia accounts for very little variance in accuracy for angry videos at the normal (S2K2) level once autistic traits have been accounted for. However, since our autism and alexithymia measures were correlated [$R = .53$, $p < .001$], when alexithymia is entered into a multiple regression after autistic traits, it may not be a significant predictor due to multi-collinearity. Consequently, we ran one further hierarchical regression, with the demographic variables entered in Step 1, alexithymia in Step 2 and autistic traits in Step 3. Alexithymia failed to significantly improve the model [$F \text{ change (1, 55)} = .31$, $p = .581$, $R^2 \text{ change} = .005$], explaining only 0.5% more variance than that explained by the demographic variables alone. Despite being highly correlated with alexithymia, autistic traits were again a
significant predictor of accuracy for angry videos at the normal level [standardized \( \beta = -.45 \), \( t(54) = -3.33 \), \( p < .01 \)] when added to the model in Step 3. Adding autistic traits at this step produced a statistically significant \( R^2 \) change \([F \text{ change } (1, 54) = 11.12, p < .01, R^2 \text{ change } = .143]\), explaining an additional 14.3% of the variance in accuracy.

Table 2. Results of the forced entry hierarchical regression for accuracy for angry videos at the normal spatial and speed level. 1. predictors: age, gender, non-verbal reasoning; 2. predictors: age, gender, non-verbal reasoning, AQ; 3. predictors: age, gender, non-verbal reasoning, AQ, TAS-20

| Model | R    | R^2  | Adjusted R^2 | SEE  | R^2 change | F change | Sig. F change |
|-------|------|------|--------------|------|------------|----------|---------------|
| 1     | .400 | .160 | .115         | 1.82 | .160       | 3.556    | .020          |
| 2     | .536 | .287 | .235         | 1.69 | .127       | 9.798    | .003          |
| 3     | .554 | .307 | .243         | 1.68 | .020       | 1.581    | .214          |

Footnote:

[4] See Supplementary Information E for a juxtaposition of the current data against data published by Sowden et al. (in press).

Discussion

The current study tested whether alexithymia accounts for recognition difficulties with respect to both spatial and kinematic aspects of facial emotional expression recognition in ASD. We hypothesized that the effects of spatial and kinematic manipulation on emotion recognition accuracy would covary with scores on a self-report alexithymia measure and would not be explained by group (ASD versus control) membership. In replication of Sowden et al., (2013), our results indicated that emotion recognition accuracy was affected by both spatial and kinematic manipulation. However, in conflict with our hypotheses, emotion recognition accuracy did not covary with alexithymia scores. Rather, we observed a significant emotion x spatial x kinematic x group interaction. Further unpacking this interaction revealed that autistic, relative to control, adults showed reduced recognition of angry expressions at the normal (100%) spatial (S2) and speed (K2) level. Furthermore, whilst control participants improved in accuracy across all kinematic levels, autistic participants only benefitted from the speed increase from the normal (100%) to increased (150%) speed level. In addition, multiple regression analyses revealed that autistic traits and NVR, but not age, gender or alexithymia, were significant predictors of recognition accuracy for angry videos at the normal spatial and speed level (where autistic traits were a negative predictor and NVR was a positive predictor). O results demonstrate that when autistic and control individuals are matched in terms of alexithymia there are group differences in recognition accuracy, though these are restricted to angry (not happy or sad) expressions.

Establishing what can and cannot be explained by the alexithymia hypothesis is of major importance not only to academics working in the field but also to clinicians for whom it is important to understand which
aspects of behaviour and cognition are indicative of autism, and which are more representative of alexithymia. If the alexithymia hypothesis is to explain functional, everyday, challenges with expression recognition it must be the case that alexithymia can account for recognition difficulties with respect to both spatial and kinematic aspects of facial emotional expressions. Here we suggest that this is not the case. Self-reported alexithymia was not predictive of the effect of spatial or kinematic manipulation on emotion recognition emotion recognition accuracy in general, or emotion recognition accuracy specifically relating to angry videos at the normal spatial and speed level. Furthermore, mean accuracy was comparable for alexithymic (i.e. TAS-20 $\geq 61$) and non-alexithymic individuals (i.e. TAS-20 $\leq 51$). Importantly, the mean alexithymia scores in our study (Control mean= 55.66; ASD mean= 59.74) are comparable to that of Cook et al. (2013; Control mean= 46.9; ASD mean= 55.6) and 46.67% of our participants could be identified as having alexithymia (TAS-20 $\geq 61$) according to the TAS-20. Thus, it is not the case that the lack of an effect of alexithymia is due to abnormally low TAS-20 scores in our sample.

Though we do not refute the idea that difficulties with emotion recognition from static images of faces may be better explained by the presence of alexithymia rather than autism, we do not find evidence to support the claim that this argument extends to dynamic stimuli. Our data raise the possibility that alexithymic individuals experience difficulties that are specific to static stimuli and may be able to rely on dynamic information to compensate for difficulties with recognising emotions from static snapshots of faces. An alternative explanation, however, is that dynamic stimuli simply contain more information (i.e. spatial information about the configuration of facial features and kinematic information about changes over time). Consequently, it may be that tasks which use only static stimuli are more challenging and thus more sensitive to individual differences than those that employ dynamic stimuli. Nevertheless, accuracy scores on our task ranged between -6.05 and 9.92 (out of 10) and our task was able to index differences in anger recognition between autistic and control individuals. Thus, our task is clearly sensitive to individual differences in emotion recognition. Future studies, which specifically aim to test whether emotion processing difficulties in alexithymia are specific to static stimuli should titrate the sensitivity of static and dynamic tasks and compare performance on both tasks.

Of particular note is our finding that differences between autistic and control individuals are restricted to the recognition of angry expressions. This finding is in line with previous research suggesting that angry expressions are better recognized by non-autistic compared to autistic individuals (Ashwin et al., 2006; Bal et al., 2010; Brewer et al., 2016; Leung, Pang, Brian, & Taylor, 2019; Song & Hakoda, 2018a) and is supported by meta-analytic evidence demonstrating greater differences between ASD and control groups in the recognition of angry compared to happy and sad expressions (Lozier, Vanmeter, & Marsh, 2014). Importantly, however, some of these previous studies did not measure alexithymia (Ashwin et al., 2006; Bal et al., 2010; Leung, Pang, Brian, & Taylor, 2019; Song & Hakoda, 2018) and in those that did, alexithymic and ASD traits were highly confounded (Brewer et al., 2016), making it impossible to determine whether differences in anger recognition were attributable to alexithymia or ASD. The present study resolves this ambiguity and suggests that difficulties with recognising angry expressions at the ‘normal’ spatial and speed level are related to autism, not alexithymia.
An important observation is that in the current paradigm both groups performed equally well for slowed angry expressions, but whilst the controls benefitted from the K1 to K2 speed increase (i.e. 50% to 100% speed), the autistic participants only benefitted from the K2 to K3 speed increase (i.e. accuracy only increased when the stimulus was played at 150% of normal speed). These findings raise the possibility that autistic individuals may have a higher ‘kinematic threshold’ for perceiving angry expressions (i.e. an angry expression has to be moving quite quickly before it actually appears angry or angrier to ASD participants). This idea builds upon the findings of a previous study that used static photographic stimuli at varying expressive intensities (constructed by repeatedly morphing a full expression with a neutral expression to result in 9 intensity levels for each emotion) to estimate identification thresholds (the intensity at which an expression is identified correctly on two consecutive trials) for autistic and control participants. The authors found that autistic individuals had significantly higher identification thresholds than controls, meaning that a higher intensity was necessary before an expression appeared angry to ASD participants. Importantly, this study also found no significant group differences in identification thresholds for happiness or sadness (Song & Hakoda, 2018). These findings suggest that autistic individuals have a different identification threshold for static angry expressions. For dynamic facial expressions, it may be that autistic and control individuals have a different ‘kinematic identification threshold’ such that the expression must move more quickly (than would be required for control individuals) before it is identified as angry. Further research is necessary to investigate whether the group difference in recognising angry expressions at the S2K2 level is underpinned by a difference in kinematic identification thresholds.

Another (non-mutually exclusive) explanation for why the autistic individuals may have particular difficulty recognizing angry expressions relates to movement production. Previous studies have documented differences between autistic and control participants in the production of facial expressions of emotion (Brewer et al., 2016; Keating & Cook, 2020). In our study, we used PLF videos that were made by filming four neurotypical participants posing different emotional states. Given that autistic and neurotypical individuals may produce different facial expressions and that one’s own movement patterns influence the perception and interpretation of the movements of others (Cook, 2016; Eddy & Cook, 2018; Edey, Yon, Cook, Dumontheil, & Press, 2017; Happé, Cook, & Bird, 2017) our autistic participants might have struggled to read emotion in our PLF videos because the expressions were dissimilar to expressions that they would have adopted themselves. To date studies that have documented differences between autistic and control participants in the production of facial expressions of emotion have used neurotypical observer ratings as a measure of the quality of facial expression (i.e. from the perspective of a neurotypical rater autistic individuals produce expressions which appear “atypical”). Consequently, research has not yet identified what specifically is different about autistic and non-autistic facial expressions. Importantly, differences might be found in the final arrangement of facial features (i.e. spatial differences) or the speed/acceleration/jerk with which individuals reach these expressions (i.e. kinematic differences). Further research is necessary to a) characterize the expressive differences between autistic and neurotypical individuals, b) ascertain whether there are greater expressive
differences between the groups for angry than happy and sad expressions and, c) confirm whether such differences in movement profile contribute to emotion recognition difficulties.

**Limitations**

In the present study, we aimed to produce statistically rigorous and replicable results. The standard alpha level (p < .05) has recently been called into question for its utility and appropriateness in psychological research (Amrhein & Greenland, 2018; Benjamin et al., 2018; Halsey, Curran-Everett, Vowler, & Drummond, 2015; Lakens et al., 2018). Hence, we are reassured to see that our main findings remain significant, after Bonferroni-correction and, when we set a more conservative alpha threshold of 0.025. Importantly, substantial effect sizes and Bayes factors support our low p values, thus providing us with further confidence in our results. Therefore, we believe our findings make sound contributions to the literatures regarding alexithymia, ASD and dynamic facial expression recognition, however, there are several limitations to consider.

One potential limitation is that due to COVID-19-related restrictions on face-to-face testing, only 22 of our ASD group completed ADOS-2 assessments. As a result, we have limited information about whether the remaining 9 participants would surpass the threshold for an autism or autism spectrum diagnosis on the ADOS-2. In addition, of the 22 participants that did complete the observational assessment, just 15 met criteria for a diagnosis. Hence, it is possible that our ASD group display less frequent or lower intensity autistic behaviours than would typically be seen in an ASD population. In spite of this we identified a significant group difference. Note that this limitation may have resulted in false negatives or an underestimation of the true effect size. However, it is highly unlikely that it could have resulted in false positives or inflated effects sizes.

Another potential limitation of this study is that we used the self-report TAS-20 to measure alexithymia. Whilst 89% of studies comparing the emotional self-awareness of autistic and non-autistic participants use self-report measures (and 62% use the TAS-20; Huggins, Donnan, Cameron, & Williams, 2020), some authors (e.g. Leising, Grande, & Faber, 2009; Marchesi, Ossola, Tonna, & De Panfilis, 2014) have questioned their utility as “people with alexithymia, by definition, should not be able to report their psychological state” (Marchesi et al., 2014). However, endeavours to develop objective measures of alexithymia are in their infancy and early attempts are yet to be replicated (e.g. Gaigg, Cornell, & Bird, 2018) and thus self-report measures are necessary. Whilst the TAS-20 has long been the gold-standard tool for assessing alexithymia, there are some concerns that it might actually be a measure of psychopathology symptoms or current levels of psychological distress (see Badura, 2003; Helmes, McNeill, Holden, & Jackson, 2008; Leising et al., 2009; Marchesi et al., 2014; Preece et al., 2020; Rief, Heuser, & Fichter, 1996). Further studies may try to replicate our results using alternative measures of alexithymia such as the Perth Alexithymia Questionnaire (Preece, Becerra, Robinson, Dandy, & Allan, 2018) or Bermond Vorst Alexithymia Questionnaire (BVAQ; Vorst & Bermond, 2001), which have been argued to index an alexithymia construct that is distinct from individuals’ current level of psychological distress (Preece et al., 2020). However, since our aim was to investigate whether the alexithymia
hypothesis applies, not only to emotion recognition from static face stimuli, but also to recognition from dynamic stimuli, it was crucial for the design of the current study that we employ the same measure of alexithymia (i.e. the TAS-20) as has previously been employed in the emotion recognition literature (Cook et al., 2013; Milosavljevic et al., 2016; Oakley et al., 2016; Ola & Gullon-Scott, 2020).

Conclusions

The current study tested whether alexithymia can account for recognition difficulties with respect to both spatial and kinematic aspects of facial emotional expression recognition in ASD. In conflict with our hypotheses, emotion recognition accuracy did not covary with alexithymia scores. Rather, we observed that autistic, relative to control, adults showed reduced recognition of angry expressions at the normal (100%) spatial and speed level. Interestingly whilst for controls, recognition accuracy improved across all levels of the kinematic manipulation for angry videos, the autistic participants only benefitted from the 100% to 150% speed increase. Our results fail to provide evidence to support the idea that ‘the alexithymia hypothesis’ extends to emotion recognition from dynamic face stimuli. Instead our results draw attention to anger specific differences in emotion recognition between autistic and non-autistic individuals. Future research should aim to elucidate why autistic individuals exhibit differences that are specific to angry expressions.

Declarations

Ethics approval and consent to participate

The study has received approval from the STEM Ethics Committee at the University of Birmingham (ERN_16-0281AP9A) and from the Birmingham Participatory Autism Research Team (B-PART) Consultancy group. All participants gave consent to participate.

Consent for publication

Not applicable

Availability of data and materials

The dataset generated and analyzed during the current study are available in the Open Science Framework (OSF) repository, https://osf.io/6jw3f/wiki/home/

Competing interests

The authors declare that they have no competing interests

Funding
This project was supported by the Medical Research Council (MRC, United Kingdom) MR/R015813/1 and the European Union’s Horizon 2020 Research and Innovation Programme under ERC-2017-STG Grant Agreement No 757583.

**Author’s contributions**

CTK is the first author on this paper. CTK sought ethical approval, designed the experiment, recruited the participants and completed the analysis and write-up. SS helped program the study such that testing could occur online and provided regular supervision throughout. DSF helped with processing the data to facilitate analysis. JLC is the last author on this paper. JLC provided continued supervision throughout the process. JLC helped with designing the experiment, conducting the analysis, and writing up the experiment. All authors read and approved the final manuscript.

**Acknowledgments**

We thank the Birmingham Participatory Autism Research Team (B-PART) consultancy group for their advice and input on the study. We are really grateful for their fantastic contribution.

**References**

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)* (5th ed.). BMC Med 17.

Amrhein, V., & Greenland, S. (2018). Remove, rather than redefine, statistical significance. *Nature Human Behaviour, 2*(1), 4. https://doi.org/10.1038/s41562-017-0224-0

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: a test of the amygdala theory. *Social Neuroscience, 1*(3–4), 349–363. https://doi.org/10.1080/17470910601040772

Badura, A. S. (2003). Theoretical and empirical exploration of the similarities between emotional numbing in posttraumatic stress disorder and alexithymia. *Journal of Anxiety Disorders, 17*(3), 349–360. https://doi.org/10.1016/S0887-6185(02)00201-3

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism
and Developmental Disorders, 40(3), 358–370. https://doi.org/10.1007/s10803-009-0884-3

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from ... Journal of Autism and Developmental Disorders, 31(1), 5–17.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... Johnson, V. E. (2018). Redefine statistical significance. Nature Human Behaviour, 2(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. European Psychiatry, 20(3), 291–298. https://doi.org/10.1016/j.eurpsy.2004.06.013

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. Translational Psychiatry, 3(7), e285-8. https://doi.org/10.1038/tp.2013.61

Brewer, R., Biotti, F., Catmur, C., Press, C., Happé, F., Cook, R., & Bird, G. (2016). Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders. Autism Research, 9(2), 262–271. https://doi.org/10.1002/aur.1508

Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S. J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. Royal Society Open Science, 6(10). https://doi.org/10.1098/rsos.190232

Cook, J. (2016). From movement kinematics to social cognition: The case of autism. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1693), 8–10. https://doi.org/10.1098/rstb.2015.0372

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions. Psychological Science, 24(5), 723–732. https://doi.org/10.1177/0956797612463582

Dobs, K., Bülthoff, I., & Schultz, J. (2018). Use and usefulness of dynamic face stimuli for face perception studies-a review of behavioral findings and methodology. Frontiers in Psychology, 9(AUG), 1–7. https://doi.org/10.3389/fpsyg.2018.01355

Dziobek, I., Bahnemann, M., Convit, A., & Heekeren, H. R. (2010). The role of the fusiform-amygdala system in the pathophysiology of Autism. Archives of General Psychiatry, 67(4), 397–405. https://doi.org/10.1001/archgenpsychiatry.2010.31

Eddy, C. M., & Cook, J. L. (2018). Emotions in action: The relationship between motor function and social cognition across multiple clinical populations. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 86(March), 229–244. https://doi.org/10.1016/j.pnpbp.2018.05.021
Edey, R., Yon, D., Cook, J., Dumontheil, I., & Press, C. (2017). Observation our own action kinematics predict the perceived affective states of others. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1263–1268. https://doi.org/10.1037/xhp0000423

Gaigg, S. B., Cornell, A. S. F., & Bird, G. (2018). The psychophysiological mechanisms of alexithymia in autism spectrum disorder. *Autism, 22*(2), 227–231. https://doi.org/10.1177/13623613166667062

Halsey, L. G., Curran-Everett, D., Vowler, S. L., & Drummond, G. B. (2015). The fickle P value generates irreproducible results. *Nature Methods, 12*(3), 179–185. https://doi.org/10.1038/nmeth.3288

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology, 68*, 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review, 20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Helmes, E., McNeill, P. D., Holden, R. R., & Jackson, C. (2008). The construct of alexithymia: Associations with defense mechanisms. *Journal of Clinical Psychology, 64*(3), 318–331. https://doi.org/10.1002/jclp.20461

Hill, E., Berthoz, S., & Frith, U. (2004). Brief Report: Cognitive Processing of Own Emotions in Individuals with Autism...: EBSCOhost. *Journal of Autism & Developmental Disorders, 34*(2), 229.

Hobson, R. P. (1986). The Autistic Child's Appraisal of Expressions of Emotion. *Journal of Child Psychology and Psychiatry, 27*(3), 321–342. https://doi.org/10.1111/j.1469-7610.1986.tb01836.x

Huggins, C. F., Donnan, G., Cameron, I. M., & Williams, J. H. G. (2020). A systematic review of how emotional self-awareness is defined and measured when comparing autistic and non-autistic groups. *Research in Autism Spectrum Disorders, 77*(December 2019), 101612. https://doi.org/10.1016/j.rasd.2020.101612

Keating, C. T., & Cook, J. L. (2020). Facial Expression Production and Recognition in Autism Spectrum Disorders: A Shifting Landscape. *Child and Adolescent Psychiatric Clinics of North America, 29*(3), 557–571. https://doi.org/10.1016/j.chc.2020.02.006

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Ketelaars, M. P., In'T Velt, A., Mol, A., Swaab, H., & Van Rijn, S. (2016). Emotion recognition and alexithymia in high functioning females with autism spectrum disorder. *Research in Autism Spectrum Disorders, 21*, 51–60. https://doi.org/10.1016/j.rasd.2015.09.006
Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18*(1), 156–168. https://doi.org/10.1006/nimg.2002.1323

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry, 55*, 80–89. https://doi.org/10.1016/j.eurpsy.2018.09.004

Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., ... Glueck, D. H. (2013). GLIMMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software, 54*(10), 1–26. https://doi.org/10.18637/jss.v054.i10

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*(1), 41–46. https://doi.org/10.1177/1754073912451349

Lakens, D., Adolli, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., ... Zwaan, R. A. (2018). Justify your alpha. *Nature Human Behaviour, 2*(3), 168–171. https://doi.org/10.1038/s41562-018-0311-x

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modelling: A Practical Course*. https://doi.org/10.1017/CBO9781139087759

Leising, D., Grande, T., & Faber, R. (2009). The Toronto Alexithymia Scale (TAS-20): A measure of general psychological distress. *Journal of Research in Personality, 43*(4), 707–710. https://doi.org/10.1016/j.jrp.2009.03.009

Leung, R. C., Pang, E. W., Brian, J. A., & Taylor, M. J. (2019). Happy and Angry Faces Elicit Atypical Neural Activation in Children With Autism Spectrum Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4*(12), 1021–1030. https://doi.org/10.1016/j.bpsc.2019.03.013

Leung, R. C., Pang, E. W., Brian, J. A., & Taylor, M. J. (2019b). Happy and Angry Faces Elicit Atypical Neural Activation in Children With Autism Spectrum Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4*(12), 1021–1030. https://doi.org/10.1016/j.bpsc.2019.03.013

Lindner, J. L., & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger’s syndrome. *Journal of Autism and Developmental Disorders, 36*(6), 769–777. https://doi.org/10.1007/s10803-006-0105-2

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnositic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1-4*. *Western Psychological Services*. 
Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology, 26*(4), 933–945. https://doi.org/10.1017/S0954579414000479

Marchesi, C., Ossola, P., Tonna, M., & De Panfilis, C. (2014). The TAS-20 more likely measures negative affects rather than alexithymia itself in patients with major depression, panic disorder, eating disorders and substance use disorders. *Comprehensive Psychiatry, 55*(4), 972–978. https://doi.org/10.1016/j.comppsych.2013.12.008

Milosavljevic, B., Carter Leno, V., Simonoff, E., Baird, G., Pickles, A., Jones, C. R. G., ... Happé, F. (2016). Alexithymia in Adolescents with Autism Spectrum Disorder: Its Relationship to Internalising Difficulties, Sensory Modulation and Social Cognition. *Journal of Autism and Developmental Disorders, 46*(4), 1354–1367. https://doi.org/10.1007/s10803-015-2670-8

Mottron, L. (2004). Matching Strategies in Cognitive Research with Individuals with High-Functioning Autism: Current Practices, Instrument Biases, and Recommendations. *Journal of Autism and Developmental Disorders, 34*(1), 19–27. https://doi.org/10.1023/B:JADD.0000018070.88380.83

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. *Modern Trends in Psychosomatic Medicine, 3*, 430–439.

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. *Journal of Abnormal Psychology, 125*(6), 818–823. https://doi.org/10.1037/abn0000182

Ola, L., & Gullon-Scott, F. (2020). Facial emotion recognition in autistic adult females correlates with alexithymia, not autism. *Autism*. https://doi.org/10.1177/1362361320932727

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., ... Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine, 40*(11), 1919–1929. https://doi.org/10.1017/S0033291709992364

Poquérusse, J., Pastore, L., Dellantonio, S., & Esposito, G. (2018). Alexithymia and autism spectrum disorder: A complex relationship. *Frontiers in Psychology, 9*(JUL), 1–10. https://doi.org/10.3389/fpsyg.2018.01196

Preece, D. A., Becerra, R., Boyes, M. E., Northcott, C., McGillivray, L., & Hasking, P. A. (2020). Do self-report measures of alexithymia measure alexithymia or general psychological distress? A factor analytic examination across five samples. *Personality and Individual Differences, 155*(August 2019), 109721. https://doi.org/10.1016/j.paid.2019.109721

Preece, D., Becerra, R., Robinson, K., Dandy, J., & Allan, A. (2018). The psychometric assessment of alexithymia: Development and validation of the Perth Alexithymia Questionnaire. *Personality and
Rief, W., Heuser, J., & Fichter, M. M. (1996). What does the Toronto Alexithymia Scale TAS-R measure? *Journal of Clinical Psychology, 52*(4), 423–329. Retrieved from 10.1002/(SICI)1097-4679(199607)52:4%3C423::AID-JCLP6%3E3.0.CO;2-Q

Rivet, T. T., & Matson, J. L. (2011). Review of gender differences in core symptomatology in autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*(3), 957–976. https://doi.org/10.1016/j.rasd.2010.12.003

Ruzich, Emily Allison, Carrie Smith, P., & Watson, P. et al. (2015). *Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical and fe ... Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ).* 6(1), 1–12. https://doi.org/10.1186/2040-2392-6-2

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2016). Subgrouping siblings of people with autism: Identifying the broader autism phenotype. *Autism Research, 9*(6), 658–665. https://doi.org/10.1002/aur.1544

Salminen, J. K., Saarijärvi, S., Äärelä, E., Toikka, T., & Kauhanen, J. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *Journal of Psychosomatic Research, 46*(1), 75–82. https://doi.org/10.1016/S0022-3999(98)00053-1

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research, 20*(1), 81–91. https://doi.org/10.1016/j.cogbrainres.2004.01.008

Sato, W., Uono, S., & Toichi, M. (2013). Atypical recognition of dynamic changes in facial expressions in autism spectrum disorders. *Research in Autism Spectrum Disorders, 7*(7), 906–912. https://doi.org/10.1016/j.rasd.2013.04.008

Song, Y., & Hakoda, Y. (2018a). Selective Impairment of Basic Emotion Recognition in People with Autism: Discrimination Thresholds for Recognition of Facial Expressions of Varying Intensities. *Journal of Autism and Developmental Disorders, 48*(6), 1886–1894. https://doi.org/10.1007/s10803-017-3428-2

Song, Y., & Hakoda, Y. (2018b). Selective Impairment of Basic Emotion Recognition in People with Autism: Discrimination Thresholds for Recognition of Facial Expressions of Varying Intensities. *Journal of Autism and Developmental Disorders, 48*(6), 1886–1894. https://doi.org/10.1007/s10803-017-3428-2

Sowden, S., Schuster, B. A., Keating, C. T., Fraser, D. S., & Cook, J. L. (n.d.). The role of movement kinematics in facial emotion expression production and recognition. *Emotion.*

Stevenson, J. L., & Hart, K. R. (2017). Psychometric properties of the autism-spectrum quotient for assessing low and high levels of autistic traits in college students. *Journal of Autism and Developmental
Disorders, 47(6), 1838–1853. https://doi.org/10.1007/s10803-017-3109-1

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (2003). The 20-Item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. *Journal of Psychosomatic Research, 55*(3), 277–283. https://doi.org/10.1016/S0022-3999(02)00601-3

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

Vorst, H. C. M., & Bermond, B. (2001). Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire. *Personality and Individual Differences, 30*(3), 413–434. https://doi.org/10.1016/S0191-8869(00)00033-7

**Figures**

![Figure 1](image)

**Figure 1**

Example of one trial in the PLF main task. The fixation cross display is presented for 500ms at the start of each trial. The average length of a stimulus video was 2 seconds. Rating scales remained on screen until participants had rated the stimulus and pressed the space bar.
Figure 1

Example of one trial in the PLF main task. The fixation cross display is presented for 500ms at the start of each trial. The average length of a stimulus video was 2 seconds. Rating scales remained on screen until participants had rated the stimulus and pressed the space bar.
Figure 1

Example of one trial in the PLF main task. The fixation cross display is presented for 500ms at the start of each trial. The average length of a stimulus video was 2 seconds.Rating scales remained on screen until participants had rated the stimulus and pressed the space bar.
Figure 2

Mean accuracy scores, for all participants, for each emotion across the spatial and kinematic levels. The black line represents the mean, the shaded region represents the standard deviation, the coloured box represents 1 standard error around the mean and the dots are individual datapoints.
Figure 2

Mean accuracy scores, for all participants, for each emotion across the spatial and kinematic levels. The black line represents the mean, the shaded region represents the standard deviation, the coloured box represents 1 standard error around the mean and the dots are individual datapoints.
Figure 2

Mean accuracy scores, for all participants, for each emotion across the spatial and kinematic levels. The black line represents the mean, the shaded region represents the standard deviation, the coloured box represents 1 standard error around the mean and the dots are individual datapoints.
Figure 3

Accuracy at the S2, K2 level, as a function of emotion. Control in lilac, ASD in green. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.
Figure 3

Accuracy at the S2, K2 level, as a function of emotion. Control in lilac, ASD in green. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.

*p<.05
Figure 3

Accuracy at the S2, K2 level, as a function of emotion. Control in lilac, ASD in green. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.
Figure 4

Mean accuracy scores for angry videos at the S2 level for control and ASD participants across the kinematic levels. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.
Mean accuracy scores for angry videos at the S2 level for control and ASD participants across the kinematic levels. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.
Figure 4

Mean accuracy scores for angry videos at the S2 level for control and ASD participants across the kinematic levels. The black line represents the mean, the coloured box represents the standard error of the mean, the shaded region represents the standard deviation, and the dots are individual datapoints.
This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialsPLFASDFinal.docx
- SupplementaryMaterialsPLFASDFinal.docx
- SupplementaryMaterialsPLFASDFinal.docx