

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

Connor Tom Keating (✉ [cxk655@bham.ac.uk](mailto:cxk655@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/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

For many years, research has suggested that autistic individuals have difficulties recognising the emotions of other people. However, a burgeoning literature argues that these difficulties may be better explained by co-occurring alexithymia rather than autistic characteristics. Importantly, extant studies in this field have focused on the recognition of emotion from *static* images. Here we investigated whether there are differences with respect to emotion recognition from *dynamic* facial stimuli between autistic and non-autistic groups matched on alexithymia.

## Methods

29 control and 31 autistic adults, matched on age, gender, non-verbal reasoning ability and alexithymia, completed a facial emotion recognition task which employed dynamic point light displays of happy, angry and sad facial expressions. Stimuli were manipulated such that expressions were reproduced at 50%, 100% and 150% of their normal speed and spatial extent.

## Results

The ASD group exhibited significantly lower emotion recognition accuracy for angry, but not happy or sad, expressions at the normal (100%) spatial and speed level. Whilst the control group exhibited increasing accuracy across all levels of the speed manipulation, the ASD group only showed improvement from the 100% to 150% level. Non-verbal reasoning and level of autistic traits (and not age, gender or alexithymia) were significant predictors of accuracy for angry videos at the 100% spatial and speed level.

## Limitations

Due to COVID-19 restrictions, only 22 members of the ASD group completed the ADOS-2 assessments and 7 of those who did, scored below threshold for an autism or ASD diagnosis. Therefore, our ASD group may display less frequent or lower intensity autistic behaviours than would typically be seen in an ASD population. The TAS, which has recently been questioned for its construct validity, was used to measure alexithymia.

## Conclusions

Since our participants were matched on alexithymia, and we identified that level of autistic traits (and not alexithymic traits) was a significant predictor of the accuracy of angry expression recognition at the normal level, we conclude that a difficulty with recognising angry expressions *is* relevant to autism and cannot be explained by alexithymia. Future research should elucidate why autistic individuals exhibit differences with angry expressions in particular.

# Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder, characterized by difficulties in social communication, and restricted and repetitive interests (1). 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 (2). However, whilst many studies suggest a disparity in the facial emotion recognition ability of autistic[1] and neurotypical individuals (3–6), there have been inconsistent findings, ranging from no differences between these individuals to large disparities (see 5–7 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 (10–13). Alexithymia is a subclinical condition, characterized by difficulties identifying and expressing emotions (14), which is often comorbid with ASD (in the neurotypical population the incidence of alexithymia is 13% (15) in autistic populations the incidence is estimated at 40-65% (10,16,17)). Cook and colleagues (18) 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 (18). Similarly, Milosavljevic and colleagues (19) 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 (10) 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 (20,21). 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 (22). 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 (23) 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 (23) 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 (23). 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 (24), 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 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 (25) 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 (26). The findings of this study support the alexithymia hypothesis: high alexithymic, but not autistic, traits were associated with less accurate facial expression recognition (26). 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 (27), and likely differ in behavioural phenotype (28,29), 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 (23) 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.

## Methods

**Participants.** 60 individuals, 31 with an ASD diagnosis and 29 neurotypical controls, participated in the study (See Supplementary Materials 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; 26)) and alexithymia (as measured by the 20-item Toronto Alexithymia Scale (TAS-20; 31)). The ASD group had significantly higher Autism Quotient (AQ; 28) scores (see Table 1). 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 (33). 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

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. The level of autistic characteristics of 22 individuals in the ASD group was assessed using the Autism Diagnostic Observation Schedule (version 2 (ADOS-2; 30)). Of the 22 who completed the ADOS-2 assessments, 15 met ADOS criteria for ASD (5 autism, 10 autism spectrum). Although seven individuals

in the ASD group did not meet criteria for ASD according to the ADOS, they had previously received diagnoses from independent clinicians, and thus still participated in the study. Unfortunately, it was not feasible to complete observational assessments on all ASD participants due to restrictions on face-to-face testing during the COVID-19 pandemic.

## **Materials and stimuli.**

**PLF stimuli.** The PLF task was an adapted version of that developed by Sowden and colleagues (23) which was re-programmed in Gorilla.sc (35) to facilitate online testing due to pandemic-related restrictions on face-to-face 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 (32). 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 (36,37), and has strong psychometric properties, including internal consistency ( $\alpha \geq 0.7$ ) and test-retest reliability ( $r \geq 0.8$ ; 38).

**Alexithymia.** Alexithymia was measured via the 20-item Toronto Alexithymia Scale (27). 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 ( $\alpha \geq 0.7$ ) and test-retest reliability ( $r \geq 0.7$ ) (31,39).

**Non-verbal reasoning.** Non-verbal reasoning was assessed via the Matrix Reasoning Item bank (MaRs-IB; 26). 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) (30).

**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 **Figure 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; 19).

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. 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 (40) 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.

## 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  $2 \times 3 \times 3 \times 3$  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 Materials B], a main effect of spatial level [ $F(2,116) = 259.57$ ,  $p < .001$ ,  $\eta_p^2 = .82$ ,  $BF_{10} = 9.05e^{57}$ ; see Supplementary Materials B] which was qualified by an emotion  $\times$  spatial interaction [ $F(4,232) = 88.42$ ,  $p < .001$ ,  $\eta_p^2 = .60$ ,  $BF_{10} = 7.53e^{58}$ ], and an emotion  $\times$  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  $\times$  spatial  $\times$  kinematic  $\times$  group interaction [ $F(8,464) = 2.438$ ,  $p < .05$ ,  $\eta_p^2 = .04$ ,  $BF_{10} = 0.07$ ]. Note that no kinematic  $\times$  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  $\times$  kinematic and emotion  $\times$  spatial interactions. Subsequently we fully unpack the emotion  $\times$  spatial  $\times$  kinematic  $\times$  group interaction.

In line with Sowden et al., (23), we observed an emotion  $\times$  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); **Figure 2**]. This pattern matches the results reported by Sowden et al., (23) 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); **Figure 2**. Bottom panel]. This pattern of results also matches the findings from Sowden et al. (23).[1]

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 [ $t(58) = 2.78, p_{\text{bonf}} < .05$ , mean difference = 1.48,  $BF_{10} = 6.09$ ; **Figure 3**]. There were no significant group differences in emotion recognition accuracy for happy [ $p = .757, BF_{10} = 0.27$ ] or sad [ $p = .085, BF_{10} = 0.93$ ] 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 [ $F(2,116) = 4.59, p < .05, \eta_p^2 = .07, BF_{10} = 3.49$ ] but not happy [ $p = .070, BF_{10} = 0.95$ ] or sad [ $p = .123, BF_{10} = 0.53$ ] 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 [K1-K2:  $t(28) = -4.31, p_{\text{bonf}} < .001$ , mean difference = -1.62,  $BF_{10} = 153.77$ ; K2-K3:  $t(28) = -2.86, p_{\text{bonf}} < .05$ , mean difference = -0.95,  $BF_{10} = 5.52$ ], the ASD group only showed improvement from the K2 to K3 [ $t(30) = -3.46, p_{\text{bonf}} < .01$ , mean difference = -1.16,  $BF_{10} = 21.10$ ] and not K1 to K2 [ $p = .865, BF_{10} = 0.19$ ; **Figure 4**]. Furthermore, the groups did not significantly differ at K1 ( $F(1,58) = .18, p > .05$ ) or K3 ( $F(1,58) = 3.53, p > .05$ ) but at K2, controls out-performed autistic participants ( $F(1,58) = 7.75, p < 0.01, \eta_p^2 = .12$ ). 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 [standardized  $\beta = -.17, t(57) = -1.10, p = .274$ ] nor TAS-20 [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 [standardized  $\beta = .20, t(57) = 1.33, p = .189$ ] nor TAS-20 [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 [standardized  $\beta = -.19, t(57) = -1.24, p = .220$ ] nor TAS-20 [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 [ $t(48) = -.18, p = .861, \text{mean difference} = -.05, BF_{10} = 0.29$ ].

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., (18). 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 change (1, 55) = 9.80,  $p < .01$ ,  $R^2$  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 change (1, 54) = 1.58,  $p = .214$ ,  $R^2$  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 multicollinearity. 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 change (1, 55) = .31,  $p = .581$ ,  $R^2$  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 change (1, 54) = 11.12,  $p < .01$ ,  $R^2$  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 <sup>2</sup>	Adjusted R <sup>2</sup>	SEE	R <sup>2</sup> 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

## Discussion

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. 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., (23), 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). Our 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 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. (18) (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 result is in line with previous research suggesting that angry expressions are better recognized by non-autistic compared to autistic individuals (3,41–44) 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 (45). Importantly, however, some of these previous studies did not measure alexithymia (3,41,43,44) and in those that did, alexithymic and ASD traits were highly confounded (42), 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.

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 (43). 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 (8,42). 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 (46–49) 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 (50–53). 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; 51), some authors (e.g. 36,37) have questioned their utility as “people with alexithymia, by definition, should not be able to report their psychological state” (56). However, endeavours to develop objective measures of alexithymia are in their infancy and early attempts are yet to be replicated (e.g. 50) 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 51,52,54–57). Further studies may try to replicate our results using alternative measures of alexithymia such as the Perth Alexithymia Questionnaire (62) or Bermond Vorst Alexithymia Questionnaire (BVAQ; 42), which have been argued to index an alexithymia construct that is distinct from individuals' current level of psychological distress (61). 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 (13,18,19,26).

## 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.

## Abbreviations

ASD- Autism Spectrum Disorder

PLF- Point-light face

NVR- Non-verbal reasoning

AQ- Autism Quotient

TAS-20- 20-item Toronto Alexithymia Scale

MaRs-IB- The Matrix Reasoning Item Bank

SEM- Standard Error of the Mean

BF- Bayes Factor

## 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**

1. Association AP. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5). 5th ed. BMC Med 17; 2013. 133–137 p.

2. Hobson RP. the Autistic Child'S Appraisal of Expressions of Emotion. *J Child Psychol Psychiatry*. 1986;27(3):321–42.
3. Ashwin C, Chapman E, Colle L, Baron-Cohen S. Impaired recognition of negative basic emotions in autism: a test of the amygdala theory. *Soc Neurosci*. 2006;1(3–4):349–63.
4. Dziobek I, Bahnemann M, Convit A, Heekeren HR. The role of the fusiform-amygdala system in the pathophysiology of Autism. *Arch Gen Psychiatry*. 2010;67(4):397–405.
5. Lindner JL, Rosén LA. Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *J Autism Dev Disord*. 2006;36(6):769–77.
6. Philip RCM, Whalley HC, Stanfield AC, Sprengelmeyer R, Santos IM, Young AW, et al. Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychol Med*. 2010;40(11):1919–29.
7. Harms MB, Martin A, Wallace GL. Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychol Rev*. 2010;20(3):290–322.
8. Keating CT, Cook JL. Facial Expression Production and Recognition in Autism Spectrum Disorders: A Shifting Landscape. *Child Adolesc Psychiatr Clin N Am*. 2020;29(3):557–71.
9. Uljarevic M, Hamilton A. Recognition of emotions in autism: A formal meta-analysis. *J Autism Dev Disord*. 2013;43(7):1517–26.
10. Bird G, Cook R. Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Transl Psychiatry*. 2013;3(7):e285-8. Available from: <http://dx.doi.org/10.1038/tp.2013.61>
11. Kinnaird E, Stewart C, Tchanturia K. Investigating alexithymia in autism: A systematic review and meta-analysis. *Eur Psychiatry*. 2019;55:80–9.
12. Poquérousse J, Pastore L, Dellantonio S, Esposito G. Alexithymia and autism spectrum disorder: A complex relationship. *Front Psychol*. 2018;9(JUL):1–10.
13. Oakley BFM, Brewer R, Bird G, Catmur C. Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. *J Abnorm Psychol*. 2016;125(6):818–23.
14. Nemiah, J. C., Freyberger, H., & Sifneos PE. Alexithymia: A view of the psychosomatic process. *Mod trends Psychosom Med*. 1976;3:430–9.
15. Salminen JK, Saarijärvi S, Äärelä E, Toikka T, Kauhanen J. Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J Psychosom Res*. 1999;46(1):75–82.
16. Berthoz S, Hill EL. The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *Eur Psychiatry*. 2005;20(3):291–8.
17. Hill E, Berthoz S, Frith U. Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *Journal of autism and developmental disorders*. 2004 Apr 1;34(2):229-35.
18. Cook R, Brewer R, Shah P, Bird G. Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions. *Psychol Sci*. 2013;24(5):723–32.

19. Milosavljevic B, Carter Leno V, Simonoff E, Baird G, Pickles A, Jones CRG, et al. Alexithymia in Adolescents with Autism Spectrum Disorder: Its Relationship to Internalising Difficulties, Sensory Modulation and Social Cognition. *J Autism Dev Disord*. 2016;46(4):1354–67.
20. Kilts CD, Egan G, Gideon DA, Ely TD, Hoffman JM. Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage*. 2003;18(1):156–68.
21. Sato W, Kochiyama T, Yoshikawa S, Naito E, Matsumura M. Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cogn Brain Res*. 2004;20(1):81–91.
22. Dobs K, Bühlhoff I, Schultz J. Use and usefulness of dynamic face stimuli for face perception studies- a review of behavioral findings and methodology. *Frontiers in psychology*. 2018 Aug 3;9: 1355
23. Sowden S, Schuster BA, Keating CT, Fraser DS, Cook JL. The role of movement kinematics in facial emotion expression production and recognition. *Emotion*. (in press.).
24. Krumhuber EG, Kappas A, Manstead ASR. Effects of dynamic aspects of facial expressions: A review. *Emot Rev*. 2013;5(1):41–6.
25. Sato W, Uono S, Toichi M. Atypical recognition of dynamic changes in facial expressions in autism spectrum disorders. *Res Autism Spectr Disord*. 2013;7(7):906–12.
26. Ola L, Gullon-Scott F. Facial emotion recognition in autistic adult females correlates with alexithymia, not autism. *Autism*. 2020; 1362361320932727
27. Loomes R, Hull L, Mandy WPL. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *J Am Acad Child Adolesc Psychiatry*. 2017;56(6):466–74.
28. Ketelaars MP, In'T Velt A, Mol A, Swaab H, Van Rijn S. Emotion recognition and alexithymia in high functioning females with autism spectrum disorder. *Res Autism Spectr Disord*. 2016;21:51–60.
29. Rivet TT, Matson JL. Review of gender differences in core symptomatology in autism spectrum disorders. *Res Autism Spectr Disord*. 2011;5(3):957–76.
30. Chierchia G, Fuhrmann D, Knoll LJ, Pi-Sunyer BP, Sakhardande AL, Blakemore SJ. The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *R Soc Open Sci*. 2019;6(10).
31. Bagby RM, Parker JDA, Taylor GJ. The twenty-item Toronto Alexithymia scale-I. Item selection and cross-validation of the factor structure. *J Psychosom Res*. 1994;38(1):23–32.
32. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The Autism-Spectrum Quotient (AQ): Evidence from ... *J Autism Dev Disord*. 2001;31(1):5–17.
33. Mottron L. Matching Strategies in Cognitive Research with Individuals with High-Functioning Autism: Current Practices, Instrument Biases, and Recommendations. *J Autism Dev Disord*. 2004;34(1):19–27.
34. Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S. *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1-4*. West Psychol Serv. 2012;
35. Anwyl-Irvine AL, Massonnié J, Flitton A, Kirkham N, Evershed JK. Gorilla in our midst: An online behavioral experiment builder. *Behav Res Methods*. 2020;52(1):388–407.

36. Ruzich E, Allison C, Smith P, Watson P, Auyeung B, Ring H, et al. Subgrouping siblings of people with autism: Identifying the broader autism phenotype. *Autism Res.* 2016;9(6):658–65.
37. Ruzich, Emily Allison, Carrie Smith P, Watson P et al. 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 ). 2015;6(1):1–12.
38. Stevenson JL, Hart KR. Psychometric properties of the autism-spectrum quotient for assessing low and high levels of autistic traits in college students. *J Autism Dev Disord.* 2017;47(6):1838–53.
39. Taylor GJ, Bagby RM, Parker JDA. The 20-Item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. *J Psychosom Res.* 2003;55(3):277–83.
40. Lee MD, Wagenmakers E-J. *Bayesian Cognitive Modelling: A Practical Course.* Cambridge University Press; 2014.
41. Bal E, Harden E, Lamb D, Van Hecke AV, Denver JW, Porges SW. Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *J Autism Dev Disord.* 2010;40(3):358–70.
42. Brewer R, Biotti F, Catmur C, Press C, Happé F, Cook R, et al. Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders. *Autism Res.* 2016;9(2):262–71.
43. Song Y, Hakoda Y. Selective Impairment of Basic Emotion Recognition in People with Autism: Discrimination Thresholds for Recognition of Facial Expressions of Varying Intensities. *J Autism Dev Disord.* 2018;48(6):1886–94.
44. Leung RC, Pang EW, Brian JA, Taylor MJ. Happy and Angry Faces Elicit Atypical Neural Activation in Children With Autism Spectrum Disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2019;4(12):1021–30.
45. Lozier LM, Vanmeter JW, Marsh AA. Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Dev Psychopathol.* 2014;26(4):933–45.
46. Eddy CM, Cook JL. Emotions in action: The relationship between motor function and social cognition across multiple clinical populations. *Prog Neuro-Psychopharmacology Biol Psychiatry.* 2018;86(March):229–44.
47. Happé F, Cook JL, Bird G. The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annu Rev Psychol.* 2017;68:243–67.
48. Edey R, Yon D, Cook J, Dumontheil I, Press C. Our own action kinematics predict the perceived affective states of others. *J Exp Psychol Hum Percept Perform.* 2017;43(7):1263–8.
49. Cook J. From movement kinematics to social cognition: The case of autism. *Philos Trans R Soc B Biol Sci.* 2016;371(1693):8–10.
50. Amrhein V, Greenland S. Remove, rather than redefine, statistical significance. *Nat Hum Behav.* 2018;2(1):4.

51. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers EJ, Berk R, et al. Redefine statistical significance. *Nat Hum Behav.* 2018;2(1):6–10.
52. Halsey LG, Curran-Everett D, Vowler SL, Drummond GB. The fickle P value generates irreproducible results. *Nat Methods.* 2015;12(3):179–85.
53. Lakens D, Adolphi FG, Albers CJ, Anvari F, Apps MAJ, Argamon SE, et al. Justify your alpha. *Nat Hum Behav.* 2018;2(3):168–71.
54. Huggins CF, Donnan G, Cameron IM, Williams JHG. A systematic review of how emotional self-awareness is defined and measured when comparing autistic and non-autistic groups. *Res Autism Spectr Disord.* 2020;77(December 2019):101612.
55. Leising D, Grande T, Faber R. The Toronto Alexithymia Scale (TAS-20): A measure of general psychological distress. *J Res Pers.* 2009;43(4):707–10.
56. Marchesi C, Ossola P, Tonna M, De Panfilis C. 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. *Compr Psychiatry.* 2014;55(4):972–8.
57. Gaigg SB, Cornell ASF, Bird G. The psychophysiological mechanisms of alexithymia in autism spectrum disorder. *Autism.* 2018;22(2):227–31.
58. Badura AS. Theoretical and empirical exploration of the similarities between emotional numbing in posttraumatic stress disorder and alexithymia. *J Anxiety Disord.* 2003;17(3):349–60.
59. Helmes E, McNeill PD, Holden RR, Jackson C. The construct of alexithymia: Associations with delense mechanisms. *J Clin Psychol.* 2008;64(3):318–31.
60. Rief W, Heuser J, Fichter MM. What does the Toronto Alexithymia Scale TAS-R measure? *J Clin Psychol.* 1996;52(4):423–329.
61. Preece DA, Becerra R, Boyes ME, Northcott C, McGillivray L, Hasking PA. Do self-report measures of alexithymia measure alexithymia or general psychological distress? A factor analytic examination across five samples. *Pers Individ Dif.* 2020;155(August 2019):109721.
62. Preece D, Becerra R, Robinson K, Dandy J, Allan A. The psychometric assessment of alexithymia: Development and validation of the Perth Alexithymia Questionnaire. *Pers Individ Dif.* 2018;132(February):32–44.
63. Vorst HCM, Bermond B. Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire. *Pers Individ Dif.* 2001;30(3):413–34.
64. Kenny L, Hattersley C, Molins B, Buckley C, Povey C, Pellicano E. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism.* 2016;20(4):442–62.