Examining the bidirectional association between emotion recognition and ASD symptoms using observational and genetic analyses

Zoe E. Reed[1,2], Liam Mahedy[1,2], Abigail Jackson[1,2], George Davey Smith[2], Ian Penton-Voak[1,3], Angela S Attwood[1,2], Marcus R Munafò[1,2,3]

1 School of Psychological Science, University of Bristol, Bristol, UK

2 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

3 National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK

Short running head: Association between emotion recognition and ASD symptoms

Word count: 5,999

Abstract

Background: There is mixed evidence for an association between autism spectrum disorder (ASD) and emotion recognition deficits. We sought to assess the bidirectionality of this association using phenotypic and genetic data in a large community sample.

Methods: Analyses were conducted in three stages. First, we examined the bidirectional association between ASD symptoms at age 16 years and emotion recognition task (ERT) responses at age 24 years (Study 1; N=3,579); and between Diagnostic Analysis of Nonverbal Accuracy (DANVA) emotion recognition responses at age 8 years and ASD symptoms at age 16 years (Study 2; N=8,545). Next, we examined the association between these phenotypes and polygenic risk scores for these. Finally, the genetic correlation between ASD and ERT responses at age 24 was estimated. Analyses were conducted in the Avon Longitudinal Study of Parents and Children.

Results: ASD symptoms at age 16 years were associated with later total correct responses on ERT in Study 1 (ß=-0.17; 95% CI: -0.26 to -0.08). However, there was no clear evidence of an association in Study 2 after adjusting for confounders (ß=-0.004; 95% CI: -0.03 to 0.02). We observed a negative association between ERT polygenic risk score and ASD symptoms (ß=-0.23; 95% CI: -0.33 to -0.13;). However, we found the opposite association between the ASD polygenic risk score and ERT (ß=0.40; 95% CI: 0.10 to 0.70). We observed a moderate genetic correlation between ASD and ERT (r=0.30), but the statistical evidence for this was very weak.

Conclusion: We found an observational association between poorer emotion recognition and increased ASD symptoms. Our genetic analyses revealed an association between ERT polygenic risk scores and ASD symptoms, which may suggest a shared genetic aetiology between these or a potential causal pathway. Our results may inform interventions targeting emotion recognition.

Keywords: autism spectrum disorder, emotion recognition, polygenic risk score, ALSPAC.

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that manifests in childhood and is characterised by persistent difficulties in social communication and interaction and restricted and repetitive behaviours (Campisi, Imran, Nazeer, Skokauskas, & Azeem, 2018). Emotion recognition is thought to be impaired in individuals with ASD (Lozier, Vanmeter, & Marsh, 2014; Uljarevic & Hamilton, 2013), which may contribute to impairments in social function. However, the direction of effect is unclear; ASD symptoms may contribute to deficits in emotion recognition, or these deficits may contribute to, or exacerbate symptoms. Since emotion recognition is a potentially modifiable target (Penton-Voak et al., 2018; Penton-Voak, Bate, Lewis, & Munafò, 2012; Rawdon et al., 2018), a better understanding of its relationship with ASD, including the direction of association, may help inform management and treatment of ASD.

There is inconsistency in the literature on the association between ASD and emotion recognition of facial expressions. Numerous studies have found that those with ASD (or greater symptom severity) are less accurate and slower at identifying emotions than controls (or those with lower symptom severity) (Fridenson-Hayo et al., 2016; Griffiths et al., 2019; Loth et al., 2018; Lozier et al., 2014; Rump, Giovannelli, Minshew, & Strauss, 2009; Uljarevic & Hamilton, 2013; Wallace et al., 2011; Wingenbach, Ashwin, & Brosnan, 2017; Xavier et al., 2015). However, others have failed to show evidence of an association (Evers, Kerkhof, Steyaert, Noens, & Wagemans, 2014; Fink, De Rosnay, Wierda, Koot, & Begeer, 2014; Russo-Ponsaran et al., 2019; Tracy, Robins, Schriber, & Solomon, 2011). Although these studies have sought to establish whether an association between ASD and emotion

recognition deficits exist, they have all been hampered by the use of relatively small sample sizes, cross-sectional study designs, and lack of consideration of the direction of the association. In addition, while these studies focused on individuals with a diagnosis of ASD it is also valuable to examine this relationship in non-clinical populations, as this presents the opportunity to assess how emotion recognition ability varies across a broader continuum of ASD symptoms.

We examined the association, and bidirectionality of the association, between ASD symptoms and emotion recognition using data from a large prospective birth cohort, the Avon Longitudinal Study of Parents and Children (ALSPAC). We conducted the study in two parts. First, we performed observational analyses to investigate the association between a continuous measure of ASD symptoms at age 16 and performance on two distinct emotion recognition tasks at age 24 (Study 1) and age 8 (Study 2). Next, in an effort to strengthen the evidence, we performed genetic score modelling assessing the likely direction of this association further. The aim of this study was to identify the direction between ASD symptoms and emotion recognition.

Methods

Cohort description

We included children from the ALSPAC birth cohort in this study. ALSPAC initially recruited 14,541 pregnant women residing in Avon, UK, with expected dates of delivery between 1st April 1991 and 31st December 1992. Of these initial pregnancies, 13,988 children were alive at age 1. When children were approximately age 7, additional eligible cases who had failed to join the study originally were recruited. As a result, there are data available for more than the 14,541 pregnancies mentioned above after age 7. There is a total sample size, for analyses using any data after age 7, of 14,901. Further details of the cohort and enrolment can be found in the cohort profile papers (Boyd et al., 2013; Fraser et al., 2013; Northstone et al., 2019). Please note that the study website contains details of all the data that are available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data/). Data collection from questionnaires or clinics from 2014 onwards was collected and managed by REDcap electronic data capture tools hosted at the University of Bristol (Harris et al., 2009, 2019). Ethics approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Phenotypic measures

Autism spectrum disorder symptoms

We used parent reported questionnaire responses from the Social and Communication Disorders Checklist (SCDC) (Skuse, Mandy, & Scourfield, 2005) total scores at mean age 16.84 (SD=0.36). The SCDC is a 12-item screening questionnaire for social and communication difficulties. Although it does not directly assess ASD, it is a good indicator of symptoms. Response options were 'Not true', 'Quite/Sometimes true' or 'Very/Often true', ranging from zero to two, with a maximum score of 24. Data was available for 5,410 participants (48% were male) with a mean score of 2.83 (SD=3.78). The distribution of scores is presented in Supplementary FigS1. As there is a zero-skewed distribution we also report the median score which is 1.00 (interquartile range [IQR]=4.00).

Emotion recognition at age 8

We used data from clinic visits (mean age=8.65 [SD=0.32]) for the facial expression subtest of the Diagnostic Analysis of Nonverbal Accuracy (DANVA) (Nowicki & Duke, 1994). A series of 24 images of child faces expressing different emotions were shown and the child was asked to select which emotion they would assign to this face (i.e., happy, sad, angry or fearful). Facial expressions were presented in low or high intensity. Images were shown on a computer screen for two seconds, or if this was not available a manual version was used. Participants indicated their response and the tester clicked this option on the screen or noted this down. The number of errors for each emotion were also recorded. In this study we calculated the number of correct responses for each emotion, by subtracting the number of errors and missing data from the total number of tests. Data was available for 6,714 participants (50% were male) with a mean score of 19.39 (SD=2.78). The distribution of the number of correct responses for the DANVA is presented in Supplementary FigS2.

Emotion recognition at age 24

Data from clinic visits (mean age=24.46 [SD=0.78]) for the Emotion Recognition Task (ERT) (Attwood et al., 2017; Penton-Voak et al., 2012), from the 'E-Prime' session comprising three cognitive tasks, were used. During the ERT, participants were asked to assign emotions to facial images (i.e., happy, sad, angry, disgusted, surprised and fearful). Faces were shown on a screen for 0.2 seconds and then the participant selected one of the six emotions. There were eight levels of intensity for each emotion and each was presented twice, resulting in a total of 96 trials. We used the number of correct responses in analyses. Data was available for 3,562 participants (37% were male) with a mean score of 66.38 (SD=7.89). The distribution of the number of correct responses in the ERT is shown in Supplementary FigS3.

N-back working memory task

A working memory measure was also obtained for participants at the age 24 clinic: the *N*-back task (2-back condition) (Kirchner, 1958). During this task participants were presented with a series of numbers from zero to nine on a computer screen. They were asked to press the '1' key when a number presented was the same as the number presented two trials previously, or a '2' key if this was not the case. Overall performance on this task was determined with a measure of discriminability (*d'*), where a higher score indicates greater accuracy on the task. Further detail on this measure can be found in the Supplementary Material.

Stop signal task

At the same clinic, participants were asked to complete the stop signal task (Logan, Cowan, & Davis, 1984). Participants were presented with the letter's 'X' or 'O' sequentially on a computer screen and they were asked to press the relevant keys when the corresponding letter appeared, unless the stop signal tone was heard at which point, they should stop responding. There was a total of 256 trials, with four trials without the stop signal for every trial with the signal. Mean response times were calculated for the stop signal reaction time (SSRT) and used in our analyses as an indicator of response inhibition. Further detail on this measure can be found in the Supplementary Material.

Potential confounders

We included a number of potential confounders in the observational analyses between ASD symptoms and emotion recognition at age 24 (Study 1) and between emotion recognition at age 8 and later ASD symptoms (Study 2). These are shown on a timeline in Figure 1. We selected these potential confounders based on established risk factors for ASD and those that may theoretically be predictive of emotion recognition ability. First, we examined the unadjusted association between ASD and the outcome or exposure of interest (Model 1). Second, we included child's sex, measures of socioeconomic status which included highest education and highest social class of mother and partner, income and tenure and measures for mothers smoking and binge drinking when children were aged 12 or 4 (depending on the analysis) (Model 2). Third, we further included a measure for the child from the Wechsler Intelligence Scale for Children (WISC) (Model 3). Fourth, we further included experience of head injuries assessed prior to the exposure (Model 4). Fifth, we further included the Strengths and Difficulties Questionnaire (SDQ) (Goodman,

1997) which assesses behaviour (Model 5). Sixth, we further included a measure of depression assessed using the short Mood and Feelings Questionnaire (MFQ) (Angold, Costello, Van Kämmen, & Stouthamer-Loeber, 1996) and a binary measure for any anxiety disorder (Model 6). Finally, we included an earlier SCDC measure of ASD (Model 7). Further detail on these can be found in the Supplementary Material.

FIGURE 1

Missing data

To ensure analyses provided sufficient power to detect associations, we imputed missing data for the exposure and covariates up to the number that data were available for the outcome. Supplementary FigS4 and S5 show the sample attrition relevant to Studies 1 and 2, respectively. Data for the SSRT and working memory cognitive measures were collected at the same time as the ERT measure, therefore where any of these measures were the outcome we imputed to the number where data was available for at least one of these outcomes (N=3,579). Where ASD symptoms was the outcome we imputed up to the number who had data available for at least one ASD measure at ages 10, 13 and 16 (N=8,545). Data were imputed in Stata version 15 using the *ice* and *match* commands. Further detail on the variables used for imputation is presented in the Supplementary Material.

Genetic data

Children's genetic data were obtained from blood samples collected during clinic visits. Samples were genotyped using the Illumina HumanHap 550 quad chip. Genome-wide data were generated by Sample Logistics and Genotyping Facilities at

the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. Quality control measures were used, and individuals excluded based on gender mismatches, minimal or excessive heterozygosity, disproportionate missingness and insufficient sample replication. Data from 9,115 children and 500,527 SNPs passed filters and data was imputed with a phased version of the 1000genomes reference panel from the Impute2 reference data repository. After these procedures, removing participants with cryptic relatedness >5% and those who had withdrawn consent, there were 8,252 children with genotype data available.

Polygenic risk score

We constructed polygenic risk scores using Plink (version 2) (Purcell et al., 2007) for each individual in ALSPAC using summary statistics from GWAS for ASD (Grove et al., 2019), the DANVA proportion index (Coleman et al., 2017) and the number of correct responses in the ERT (Mahedy et al., 2019). The ERT and DANVA GWAS were conducted within ALSPAC. We used genetic data in ALSPAC filtered by an imputation score of 0.8 and a minor allele frequency of 0.01. For the DANVA GWAS imputation quality was filtered using a MACH r-squared >0.3. After filtering and selecting SNPs where genotype data was available, SNPs were clumped for linkage disequilibrium in Plink, using an R^2 of 0.1. We generated weighted polygenic risk scores for each phenotype by summing the number of risk alleles present for each SNP (0, 1 or 2) weighted by the beta of that SNP from the discovery sample, using the --score command in Plink. Further details can be found in the Supplementary Material. Polygenic risk scores were z-standardised; therefore, results can be interpreted as per standard deviation (SD) increase in score.

Statistical analysis

Observational associations

Analyses were conducted in R version 3.5.1. First, the associations between ASD symptoms at age 16 years and ERT, response inhibition and working memory outcomes at age 24 years were examined (Study 1). Next, we tested the association between emotion recognition at age 8 years and ASD symptoms at age 16 years (Study 2). For each of these associations we ran seven different models adjusting for different numbers of potential confounders, as described above. The models for these were similar, however for Study 2, earlier measures of head injury, mother's smoking and binge drinking variables were used. The MFQ was not included as a covariate in Study 2 as it was obtained after the exposure was collected. Models using imputed data are reported in the main text. Non-imputed models are reported in the Supplementary Material.

Sensitivity analyses

We tested whether any associations between ASD symptoms and emotion recognition were due to more general cognitive function, using models including working memory and response inhibition as outcomes. If we found an association with either of these, we ran an additional model for Study 1 to include these as additional covariates, in order to see whether the emotion recognition association is independent of a more general cognitive deficit.

Association between polygenic risk scores and phenotypes

We investigated the association between the polygenic risk scores which explained the most variance in their respective outcomes for ASD, the DANVA proportion index and the number of correct responses in the ERT with their respective outcomes. We also looked at the associations of PRS for ERT and DANVA with the ASD phenotype and of the ASD PRS with the ERT and DANVA phenotypes. We used linear models adjusted for child's sex and the first 10 principal components (PCs).

Genetic correlation

We used summary statistics from GWAS for ASD, the DANVA proportion index and the number of correct responses in the ERT to estimate genetic correlations between these phenotypes using linkage disequilibrium score (LDSC) regression. However, we were unable to calculate genetic correlations with the DANVA proportion index as the SNP heritability was too small.

Results

Observational associations between ASD symptoms and emotion recognition at age 24 using imputed data (Study 1)

Results are shown for ASD symptoms only, for each level of adjustment for Study 1 (Table 1). There was a negative association between ASD symptoms and the total number of correct emotion recognition responses in the fully adjusted model (ß=-0.17; 95% CI: -0.26 to -0.08; *p*=0.0003). These results are similar to the non-adjusted imputed results, although the effect is slightly attenuated. Results from the complete case non-imputed model are consistent (Supplementary Table S4). Supplementary Fig S6 demonstrates comparable distributions of the imputed and original variables.

TABLE 1

Sensitivity analyses for study 1

Results for the models with response inhibition and working memory as outcomes are shown in Supplementary Tables S5 and S6. There was evidence of a negative association between ASD symptoms and working memory in the fully adjusted model (ß=-0.02; 95% CI: -0.03 to -0.005; *p*=0.004). However, there was no evidence of an association between ASD symptoms and response inhibition. We therefore further adjusted for working memory. The association of ASD symptoms at age 16 with the later ERT outcome remained even after adjustment (ß=-0.14; 95% CI: -0.23 to -0.05; *p*=0.002).

Observational associations between emotion recognition at age 8 *and ASD symptoms using imputed data (study 2)*

Results from Study 2 are shown in Table 2. Results are shown for emotion recognition at age 8 only for each level of adjustment. There was a negative association between the total number of correct responses at age 8 and ASD symptoms at age 16. However, there is insufficient evidence for this association in the fully adjusted model (ß=-0.004; 95% CI: -0.03 to 0.02; *p*=0.71). Results from the complete case non-imputed model are consistent (Supplementary Table S7).

TABLE 2

Association between polygenic risk scores and phenotypes

Results from linear regression models between polygenic risk scores for ASD, the proportion index from the emotion recognition task at age 8 (DANVA) and the number of correct responses in the ERT with phenotypes for these are shown in Table 3. As expected, the ERT and DANVA polygenic risk scores are strongly associated with their respective outcomes due to these being obtained in the same sample. However, the ASD polygenic risk score is not associated with the SCDC measure. The ASD polygenic risk score was associated with the ERT outcome, in the opposite direction to what we observe phenotypically, where an increased genetic risk for ASD is associated with a greater number of correct responses in the ERT. We do not observe an association between the ASD polygenic risk score and the DANVA outcome. Finally, a decreased genetic score for the number of correct responses in the ERT was associated with a greater number of ASD symptoms. There was a lack of evidence of this association with the DANVA polygenic risk

score. There was a modest genetic correlation between ERT responses and ASD (r_G=0.30, SE=0.006), but the statistical evidence for this was very weak (*p*=0.21).

TABLE 3

Discussion

We examined the association between ASD symptoms and emotion recognition in a large population-based birth cohort. Higher symptoms at age 16 were associated with poorer emotion recognition at age 24. In addition, better emotion recognition at age 8 was associated with lower symptoms at age 16; however, evidence for this was weaker and the effect was attenuated when adjusting for a measure of behaviour. We also examined the association between response inhibition and working memory with ASD symptoms. We did not find any association with response inhibition; however, we did find an association with working memory. Nevertheless, even after adjusting for working memory, the association between ASD symptoms at age 16 and ERT performance at age 24 remained, suggesting these emotion recognition deficits are at least partially independent from any general cognitive deficit.

Comparison with previous literature

Our findings are in line with previous studies which demonstrate a negative association between ASD (or greater symptom severity) and emotion recognition (Fridenson-Hayo et al., 2016; Griffiths et al., 2019; Loth et al., 2018; Lozier et al., 2014; Rump et al., 2009; Uljarevic & Hamilton, 2013; Wallace et al., 2011; Wingenbach et al., 2017; Xavier et al., 2015). A lack of evidence of this association in other studies (Evers et al., 2014; Fink et al., 2014; Russo-Ponsaran et al., 2019; Tracy et al., 2011) may be due to much smaller sample sizes and lower power in these studies, compared to ours, and/or reliance on cross-sectional designs. Utilising longitudinal data from a large prospective cohort, provides robust evidence of an

association between poorer emotion recognition and ASD symptoms. We increased our sample size and thus power by using multiple imputation on our data and also reduced biases that might be present when using complete case analyses. The tasks used in different studies also vary so it may be that some tasks are better designed to detect these effects than others. Our observational findings suggest that there is an association between earlier ASD symptoms and later emotion recognition deficits, and vice versa (although this was attenuated when including a behavioural measure in the model). This attenuation may be due to including variables which also measure some aspects of ASD, suggesting emotion recognition may be a symptom or cognitive biomarker of ASD, or on the causal pathway from ASD. To the best of our knowledge, this is the first study to use this approach of observational and genetic techniques to examine this association.

We used polygenic risk scores to further explore the direction of effect. We found that a decreased polygenic risk score for ERT was associated with greater ASD symptoms at age 16, which may indicate shared genetic liability to emotion recognition and ASD symptoms, or potentially a causal pathway from emotion recognition deficits to ASD risk. No previous study has examined polygenic risk scores for emotion recognition and ASD symptoms. Therefore, our findings of this association are novel, but replication of this using larger samples for emotion recognition GWAS will be important. It would be useful to investigate this further, potentially with Mendelian randomisation type or genomic structural equation modelling (Grotzinger et al., 2019) approaches, but the current GWAS for emotion recognition is underpowered to do this. To investigate whether our findings could be a result of shared genetic aetiology we looked at genetic correlations between these outcomes. However, we were unable to examine this fully in our study due to poor

genetic signal for the DANVA. We did observe a modest genetic correlation between ASD and number of correct responses in the ERT, but statistical evidence was very weak, which may be due to the small sample size in the GWAS of emotion recognition.

We also found, somewhat unexpectedly, that an increased genetic risk for ASD was associated with a greater number of correct responses in the emotion recognition task at age 24, but not at age 8. Only one other study, to our knowledge, has examined the association between ASD polygenic risk score and emotion recognition outcomes, and they did not find an association with total correct responses after adjusting for multiple testing (Wendt, Carvalho, Gelernter, & Polimanti, 2019). Therefore, further examination of this association would help to identify whether this is robust, and if so, what is the direction of effect.

Limitations

Whilst we observe an association between ASD symptoms and emotion recognition in this study, there are a few considerations to take into account when interpreting these results. Firstly, as with any study there will be a degree of measurement error in the variables we used. Secondly, we used a measure of ASD symptoms at age 16, which may be less generalisable to younger age groups, although symptoms will most likely be apparent by age 16, whereas they may not be at a younger age. Thirdly, the accuracy of the polygenic risk scores can also be influenced by both the discovery sample size, which may particularly be the case with the ERT and DANVA GWAS and the target sample size, which is relatively small for ASD symptoms compared to the discovery sample size. In addition, the ASD polygenic risk score

was not predictive of our SCDC measure, although it is important to highlight that the original GWAS was conducted based on diagnosis and not a continuous measure. Conducting GWAS of emotion recognition in larger samples and perhaps continuous measures of ASD may give us greater power to detect effects in future studies.

Conclusion

Our results indicate that there is an association between emotion recognition deficits and increased ASD symptoms. This is supported by our genetic analyses which revealed an association between polygenic risk scores for emotion recognition and ASD symptoms. Our genetic analyses were underpowered to identify a causal direction of effect, but suggest that there may be a shared genetic aetiology which could explain some of our observational findings, and could in principle reflect a causal pathway from emotion recognition to ASD. Given that emotion recognition is a modifiable target, this is worth exploring further. Future research would benefit from using better powered GWAS; therefore, there is a need to collect these data in larger samples. Our results may be useful in supporting development of interventions that target emotion recognition deficits in ASD.

Acknowledgements:

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Funding: This work was supported in part by the UK Medical Research Council Integrative Epidemiology Unit at the University of Bristol (Grant ref: MC_UU_00011/7). The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and ZER, LM, AJ, GDS, IPV, ASA and MRM will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. ZER is supported by the Elizabeth Blackwell Institute for Health Research, University of Bristol and the Wellcome Trust Institutional Strategic Support Fund (Grant ref:204813/Z/16/Z). LM is supported by the Elizabeth Blackwell Institute for Health Research, University of Bristol and the Wellcome Trust Institutional Strategic Support Fund (Grant ref: 204813/Z/16/Z). IPV and MRM are supported by the National Institute for Health Research Bristol Biomedical Research Centre.

IPV and MRM are co-directors of Jericoe Ltd. a company that licenses software for the assessment and modification of emotion recognition. All other authors have declared that they have no competing or potential conflicts of interest.

Correspondence to:

Zoe Reed

Address: School of Psychological Science, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK

Email: zoe.reed@bristol.ac.uk

Key points

- Emotion recognition deficits are thought to be present in individuals with ASD which may impair social function, but findings are mixed.
- In our study we found that there was a negative association between ASD symptoms and later emotion recognition.
- We also found a negative association between a polygenic risk score for emotion recognition with symptoms of ASD.
- This may indicate that genetic variants for emotion recognition are also shared with ASD, or that emotion recognition causally influences ASD.
- Our findings may have implications for intervention development targeting emotion recognition, if this is in fact causal.

References

Angold, A., Costello, J., Van Kämmen, W., & Stouthamer-Loeber, M. (1996). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: factor composition and structure across development. *International Journal of Methods in Psychiatric Research*.

Attwood, A. S., Easey, K. E., Dalili, M. N., Skinner, A. L., Woods, A., Crick, L., Ilett, E., et al. (2017). State anxiety and emotional face recognition in healthy volunteers. *Royal Society Open Science*, *4*(5), 160855. The Royal Society.

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., et al. (2013). Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *International journal of epidemiology*, *42*(1), 111–27.

Campisi, L., Imran, N., Nazeer, A., Skokauskas, N., & Azeem, M. W. (2018). Autism spectrum disorder. *British Medical Bulletin*, *127*(1), 91–100. Narnia.

Coleman, J. R. I., Lester, K. J., Keers, R., Munafò, M. R., Breen, G., & Eley, T. C. (2017). Genome-wide association study of facial emotion recognition in children and association with polygenic risk for mental health disorders. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *174*(7), 701–711. John Wiley & Sons, Ltd.

Evers, K., Kerkhof, I., Steyaert, J., Noens, I., & Wagemans, J. (2014). No Differences in Emotion Recognition Strategies in Children with Autism Spectrum Disorder:

Evidence from Hybrid Faces. *Autism Research and Treatment*, *2014*, 1–8.

Fink, E., De Rosnay, M., Wierda, M., Koot, H. M., & Begeer, S. (2014). Brief report: Accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(9), 2363–2368. Springer US.

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., et al. (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International journal of epidemiology*, *42*(1), 97–110.

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., et al. (2016). Basic and complex emotion recognition in children with autism: Cross-cultural findings. *Molecular Autism*, *7*(1), 52. BioMed Central.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. Wiley/Blackwell (10.1111).

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2019). Impaired Recognition of Basic Emotions from Facial Expressions in Young People with Autism Spectrum Disorder: Assessing the Importance of Expression Intensity. *Journal of Autism and Developmental Disorders*, *49*(7), 2768–2778. Springer US.

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill,

W. D., Ip, H. F., et al. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525.

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431–444. Nature Publishing Group.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., et al. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. Academic Press.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. Academic Press.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(2), 276–291.

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial

expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. BioMed Central.

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. Cambridge University Press.

Mahedy, L., Suddell, S., Skirrow, C., Wootton, R., Fernandes, G. S., Heron, J., Field, M., et al. (2019). Alcohol use and cognitive functioning in young adults: an observational and Mendelian randomisation study. *medRxiv*, 19003327. Cold Spring Harbor Laboratory Press.

Northstone, K., Lewcock, M., Groom, A., Boyd, A., Macleod, J., Timpson, N., & Wells, N. (2019). The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019 [version 1; peer review: 2 approved]. *Wellcome Open Research*, *4*, 51. F1000 Research Limited.

Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The diagnostic analysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior*, *18*(1), 9–35.

Penton-Voak, I. S., Adams, S., Button, K. S., Fluharty, M., Dalili, M., Browning, M., Holmes, E., et al. (2018). Emotional recognition training modifies neural response to emotional faces but does not improve mood in healthy volunteers

with high levels of depressive symptoms. *bioRxiv*, 1–9. Cambridge University Press.

Penton-Voak, I. S., Bate, H., Lewis, G., & Munafò, M. R. (2012). Effects of emotion perception training on mood in undergraduate students: Randomised controlled trial. *British Journal of Psychiatry*, *201*(1), 71–72.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. Elsevier.

Rawdon, C., Murphy, D., Motyer, G., Munafò, M. R., Penton-Voak, I., & Fitzgerald, A. (2018). An investigation of emotion recognition training to reduce symptoms of social anxiety in adolescence. *Psychiatry Research*, *263*, 257–267. Elsevier.

Rump, K. M., Giovannelli, J. L., Minshew, N. J., & Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child Development*, *80*(5), 1434–1447. NIH Public Access.

Russo-Ponsaran, N. M., Lerner, M. D., McKown, C., Weber, R. J., Karls, A., Kang, E., & Sommer, S. L. (2019). Web-based Assessment of Social–Emotional Skills in School-Aged Youth with Autism Spectrum Disorder. *Autism Research*, *12*(8), 1260–1271.

Skuse, D. H., Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders

Checklist. *British Journal of Psychiatry*, *187*(06), 568–572. Cambridge University Press.

Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*, *41*(1), 102–109. Springer.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526.

Wallace, G. L., Case, L. K., Harms, M. B., Silvers, J. A., Kenworthy, L., & Martin, A. (2011). Diminished sensitivity to sad facial expressions in high functioning autism spectrum disorders is associated with symptomatology and adaptive functioning. *Journal of Autism and Developmental Disorders*, *41*(11), 1475–1486. NIH Public Access.

Wendt, F. R., Carvalho, C. M., Gelernter, J., & Polimanti, R. (2019). The effect of the genetic liability to autism spectrum disorder on emotion recognition in young unaffected probands from a population-based cohort. *medRxiv*, *1*(203).

Wingenbach, T. S. H., Ashwin, C., & Brosnan, M. (2017). Diminished sensitivity and specificity at recognising facial emotional expressions of varying intensity underlie emotion-specific recognition deficits in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *34*, 52–61. Elsevier.

Xavier, J., Vignaud, V., Ruggiero, R., Bodeau, N., Cohen, D., & Chaby, L. (2015). A

multidimensional approach to the study of emotion recognition in autism spectrum disorders. *Frontiers in Psychology*, *6*, 1954. Frontiers.

Figure 1. Timeline of exposure, outcome and potential confounders for studies 1 and 2.

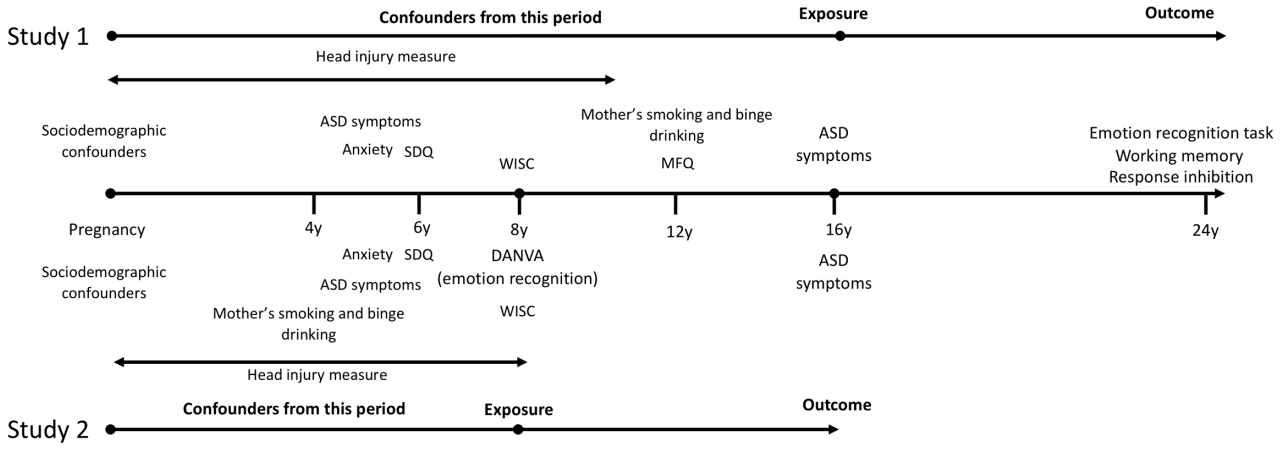

Table 1. Associations between ASD symptoms at age 16 and emotion recognition at age 24 (N=3,579)

Model	Beta	2.5% CI	97.5% CI	P
Model 1[a]	-0.24	-0.32	-0.15	1.15×10^{-07}
Model 2[b]	-0.22	-0.31	-0.14	2.79×10^{-07}
Model 3[c]	-0.20	-0.28	-0.12	2.83×10^{-06}
Model 4[d]	-0.20	-0.28	-0.12	2.78×10^{-06}
Model 5[e]	-0.18	-0.27	-0.09	6.37×10^{-05}
Model 6[f]	-0.19	-0.28	-0.10	2.19×10^{-05}
Model 7[g]	-0.17	-0.26	-0.08	0.0003

Models with number of correct responses in the emotion recognition task as the outcome and ASD symptoms as the exposure were adjusted as follows, in addition to covariates in the previous model. a: No covariates, b: sex + highest education + highest social class + mothers smoking + mothers' binge drinking + income + tenure, c: WISC, d: head injury, e: SDQ, f: MFQ + anxiety, g: earlier ASD symptoms

Table 2. Associations between emotion recognition at age 8 and ASD symptoms at age 16 (N=8,545)

Model	Beta	2.5% CI	97.5% CI	P
Model 1[a]	-0.05	-0.07	-0.03	0.000003
Model 2[b]	-0.05	-0.07	-0.03	0.000007
Model 3[c]	-0.03	-0.05	-0.01	0.002
Model 4[d]	-0.03	-0.05	-0.01	0.002
Model 5[e]	-0.02	-0.04	0.002	0.07
Model 6[f]	-0.02	-0.04	0.001	0.07
Model 7[g]	-0.004	-0.03	0.02	0.71

Models with the ASD symptoms as the outcome and number of correct responses in the DANVA as the exposure were adjusted as follows, in addition to covariates in the previous model. a: No covariates, b: sex + highest education + highest social class + mothers smoking + mothers' binge drinking + income + tenure, c: WISC, d: head injury, e: SDQ, f: anxiety, g: earlier ASD symptoms

Table 3. Associations between ASD, ERT and DANVA polygenic risk scores and phenotypes

Association examined	**N**	**Beta**	**2.5% CI**	**97.5% CI**	**P**	**R^2**
ERT ~ ASD PRS	2555	0.40	0.10	0.70	8.55×10^{-03}	0.002
DANVA ~ ASD PRS	4901	-0.001	-0.08	0.07	0.99	0.000003
ASD ~ ASD PRS	3922	-0.05	-0.17	0.07	0.39	0.0002
ERT ~ ERT PRS	2555	4.40	4.37	4.44	0.00	0.95
ASD ~ ERT PRS	3922	-0.23	-0.33	-0.13	4.12×10^{-06}	0.005
DANVA ~ DANVA PRS	4901	1.76	1.72	1.79	0.00	0.64
ASD ~ DANVA PRS	3922	-0.03	-0.13	0.07	0.60	0.0003