**Additional Supporting Information for:** Developmental Trajectories of Autistic Social Traits in the General Population. Pender, Fearon, St Pourcain, Heron & Mandy. 2021.

**Table S1: Fit indices for LGM**

|  |  |  |
| --- | --- | --- |
| Fit indices | Linear model (intercept and slope) | Quadratic model (intercept, slope and quadratic) |
| Log-likelihood (LL) | -71137.16 | -70988.49 |
| BIC | 142356.91 | 142096.28 |
| SABIC | 142328.31 | 142054.96 |
| Chi-square | 320.61 (d.f = 5, *p* <.001) | 23.27 (d.f = 1, *p* <.001) |
| RMSEA | .081 | .048 |
| CFI | .967 | .998 |
| SRMR | .039 | .009 |

*Bayesian Information Criteria (BIC) and Sample-size Adjusted BIC (SABIC)*

The BIC and SABIC use the log-likelihood value, and penalise based on the number of parameters. SABIC additionally penalises based on the sample-size. BIC and SABIC are the traditional fit statistic used for comparing mixture models – they will typically decrease and then increase following the incremental addition of classes. The model with the lowest BIC would typically be deemed optimal. Both the BIC and SABIC have been shown to perform well in simulation studies (Nylund et al., 2007; Peugh & Fan, 2012)

*Chi-square*

The χ2 statistic is an absolute fit index (based on the characteristics of the model, rather than based on comparison with an alternative specification). It assumes multivariate normality, and is sensitive to sample size (Gerbing & Anderson, 1992). With increasing sample size, the χ2 value increases, which can lead to plausible models being rejected (Schermelleh-Engel & Moosbrugger, 2003).

*Root Mean Square Error of Approximation (RMSEA; Steiger, 1990)*

The RMSEA is an index of the difference between the observed covariance matrix and the hypothesised covariance matrix denoting the model (Chen, 2007). It takes model complexity into account by reflecting degrees of freedom. Scores below .06 are regarded as indicative of a well-fitting model (Hu & Bentler, 1999).

*Comparative Fit Index (CFI; Bentler, 1990)*

The CFI is an incremental fit index (a corrected version of a non-centrality index). Scores above .95 are deemed acceptable, and .97 means it is superior to the independence model (Schermelleh-Engel & Moosbrugger, 2003). The CFI is is relatively independent from sample size and yields better performance with smaller sample sizes (Hu & Bentler, 1999)

*Standardised Root Mean Square Residual (SRMR; Byrne, 1998)*

The SRMR is an index of the average of the standardised residuals between the observed and hypothesised covariance matrices, and is relatively independent from sample size (Chen, 2007). Scores below .08 are indicative of a good fit, and .05 of excellent fit (Hu & Bentler, 1999).

**Table S2: Fit indices for LCGA**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | #param | LL | BIC | SABIC | Entropy | 2LL diff | LMR (*p*) | BLRT (*p*) |
| Two classes | 11 | -70433.58 | 140968.11 | 140933.15 | .94 | 10730.36 | <.001 | <.001 |
| Three classes | 15 | -68597.89 | 137333.44 | 137285.77 | .91 | 3671.38 | .0061 | <.001 |
| Four classes | 19 | -67935.60 | 135909.20 | 135985.19 | .88 | 1324.58 | .12 | <.001 |
| Five classes | 23 | -67354.37 | 134919.83 | 134846.74 | .87 | 1162.45 | .32 | <.001 |
| Sixclasses | 27 | -66850.81 | 133949.41 | 133863.61 | .87 | 1007.12 | .16 | <.001 |

LCGA assumes no within-class variance for the growth factors, meaning that the slopes, intercepts and quadratic terms are assumed identical across individuals within each class. The variance of the quadratic term in our LCGA specifications was constrained to zero, to improve model identification and fit

*Entropy*

Entropy is a measure of classification accuracy and indicates the level of bias one would expect. Commonly .08 is used as a threshold for an acceptable level of fit.

*Lo-Mendell-Rubin Ratio Test (LMR; Lo, Mendell, & Rubin, 2001) and Bootstrap Likelihood Ratio Test (BLRT; McLachlan & Peel, 2000)*

The LMR and BLRT compare the ratio of log-likelihoods for a model with *k* classes to a model with (*k* – 1) classes, and report a significance test to show whether the addition of the *k*th class made a significant improvement. Both perform well as a basis for model selection (Grimm et al., 2016; Nylund et al., 2007).

**Table S3: Fit indices for GMM**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Constraint | #param | LL | BIC | SABIC | Entropy | 2LL diff | LMR (*p*) | BLRT (*p*) |
| Twoclasses | Fully invariant | 11 | -69046.17 | 138193.29 | 138158.34 | .95 | 4012.95 | <.001 | <.001 |
| Three classes | Fully invariant | 15 | -68246.97 | 136631.60 | 136583.93 | .90 | 1598.41 | .18 | <.001 |
| Fourclasses | Fully invariant | 19 | -67551.13 | 135276.63 | 135216.25 | .89 | 1391.68 | .58 | <.001 |
| Fiveclasses | Fully invariant | 23 | -67020.577 | 134252.231 | 134179.141 | 0.882 | 1061.103 | .0022 | <.001 |
| Sixclasses | Fully invariant | 27 | Did not converge |
| Twoclasses | Psi-invariant | 14 | -68992.28 | 138113.04 | 138068.55 | .95 | 4064.21 | <.001 | <.001 |
| Three classes | Psi-invariant | 18 | -68179.35 | 136523.89 | 136466.69 | .89 | 1625.86 | <.001 | <.001 |
| **Four classes** | **Psi-invariant** | **22** | **-67491.69** | **135185.28** | **135115.37** | **.88** | **1375.31** | **.0049** | **<.001** |
| Fiveclasses | Psi-invariant | 26 | -67002.64 | 134243.89 | 134161.27 | .88 | 978.10 | .67 | <.001 |
| Sixclasses | Psi-invariant | 30 | -66655.16 | 133585.64 | 133490.30 | .87 | 694.96 | .21 | <.001 |
| TwoClasses | Psi variant | 15 | -61908.32 | 123954.30 | 123906.64 | 0.90 | 18288.65 | <.001 | <.001 |
| Threeclasses | Psi variant | 23 | Did not converge |
| Twoclasses | Fully variant | 21 | -61431.023 | 123054.769 | 122988.034 | 0.846 | 19186.718 | <.001 | <.001 |
| Threeclasses | Fully variant | 32 | Did not converge |

The variance of the quadratic term in our GMM specifications was constrained to zero, to improve model identification and fit

In GMM, the assumption of zero within-class variance is relaxed so that variances and covariances are obtained around the intercept, slope and quadratic growth factors. By default in MPlus, variances are constrained to be equal across classes, but this can be relaxed to allow these to vary across classes. Occasion-specific residuals (error) can also be constrained or relaxed, both longitudinally (within a class) and cross-sectionally (between classes). Hence prior to running the models, four possible specifications were designed: one fully-invariant (the most conservative model in which all variances and residuals were held equal), one fully-variant (the most freed model in which variances and residuals are allowed to vary) and two partially-variant (one conservative model allowing residuals to vary longitudinally while holding growth factors and residuals equal across classes, and one relaxed model fixing residuals longitudinally but allowing growth factors and residuals to vary between classes). We compared fit indices for the four GMM models for two through to six classes.

As an exploratory step, alternative approaches to modelling the data were specified at this stage, including Inflated Poisson models and two-part models (Muthén, 2001; Olsen & Schafer, 2001), but these did not yield improved fit or interpretability. Therefore we concluded that despite some ambiguity regarding the optimal number of classes, the original GMM model best represented the data.

**Table S4: Demographic and household characteristics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Persistent low** | **Persistent high** | **Increasing** | **Decreasing** |
|   | **n** | **%** | **n** | **%** | **OR** | ***p*** | **n** | **%** | **OR** | ***p*** | **n** | **%** | **OR** | ***p*** |
| Marital status:Never married |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1076 | 14.2 | 45 | 22.1 | 1.77 (1.26 – 2.48) | .001 | 108 | 16.6 | 1.22 (.98 – 1.51) | .080 | 115 | 18.4 | 1.38 (1.11 – 1.71) | .003 |
| Widowed / separated / divorced | 383 | 5.1 | 14 | 6.9 | 1.54 (.88 – 2.70) | .127 | 35 | 5.4 | 1.11 (.77 – 1.58) | .581 | 36 | 5.8 | 1.21 (.85 – 1.73) | .288 |
| Married | 6124 | 80.8 | 145 | 71.1 | .. |  | 506 | 78.0 | .. |  | 475 | 75.9 | .. |  |
| Maternal education:CSE/vocational |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1749 | 23.4 | 63 | 31.5 | 1.83 (1.28 – 2.61) | .001 | 171 | 26.0 | 1.24 (1.01 – 1.52) | .039 | 174 | 28.4 | 1.36 (1.10 – 1.67) | .004 |
| O Level | 2645 | 35.3 | 76 | 38.0 | 1.46 (1.04 – 2.05) | .030 | 242 | 36.8 | 1.16 (.96 – 1.40) | .115 | 212 | 34.6 | 1.09 (.90 – 1.33) | .371 |
| A Level or more | 3095 | 41.3 | 61 | 30.5 | .. |  | 244 | 37.1 | .. |  | 227 | 37.0 | .. |  |
| Child is from BME background | 290 | 3.9 | 12 | 6.1 | 1.58 (.87 – 2.87) | .131 | 30 | 4.7 | 1.19 (.81 – 1.75) | .373 | 27 | 4.5 | 1.14 (.76 – 1.71) | .517 |
| Mother’s age at delivery:Teenage |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 177 | 2.3 | 13 | 6.1 | 2.53(1.399- 4.565) | .002 | 19 | 2.8 | 1.245 (.766 – 2.022) | .376 | 24 | 3.7 | 1.594(1.027 – 2.475) | .038 |
| 20s | 4198 | 54.1 | 122 | 57.0 | .. |  | 362 | 53.6 | .. |  | 357 | 55.7 | .. |  |
| 30s | 3285 | 42.3 | 77 | 36.0 | .807 (.604 – 1.077) | .145 | 285 | 42.2 | 1.006(.856 – 1.183) | .941 | 251 | 39.2 | .898(.760 – 1.063) | .211 |
| 40s | 103 | 1.3 | <5 | - | - | .- | 9 | 1.3 | 1.013(.508 – 2.02) | .970 | 9 | 1.4 | 1.027(.515 – 2.048) | .939 |

**Table S5: IQ by group (age 8)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Persistent Low** | **Persistent High** | **Increasing** | **Decreasing** |
| IQ Component | **Mean** | **95% CI** | **Mean** | **95% CI** | **Mean** | **95% CI** | **Mean** | **95% CI** |
| Total IQ | 105.53 | 105.10 – 105.95 | 94.11 | 91.24 – 96.97 | 101.66 | 100.29 – 103.03 | 99.72 | 98.21 – 101.23 |
| Verbal IQ | 108.32 | 107.90 – 108.75 | 98.33 | 95.42 – 101.24 | 105.09 | 103.70 – 106.49 | 102.79 | 101.26 – 104.33 |
| Performance IQ | 100.87 | 100.43 – 101.30 | 90.07 | 87.11 – 93.04 | 97.17 | 95.74 – 98.59 | 96.02 | 94.46 – 97.58 |

**Table S6: Between-group comparisons: internalising and externalising difficulties, age 9-16 years**

|  |  |  |
| --- | --- | --- |
| **Baseline and final measurement (between groups)** | **Mean difference (95% CI)** | ***p*** |
| High versus Increasing:Mood age 9 | 3.53 (2.59 – 4.46) | <.001 |
| Mood age 16 | -.022 (-1.31 – 1.26) | .973 |
| Conduct age 9 | 1.84 (1.48 – 2.20) | <.001 |
| Conduct age 16 | .49 (.051 – 1.036) | .075 |
| Internalising age 9 | 1.29 (.89 – 1.69) | <.001 |
| Internalising age 16 | .40 (-.187 - .985) | .182 |
| Peer difficulties age 9 | 2.31 (1.87 – 2.75) | <.001 |
| Peer difficulties age 16 | 1.56 (.965 – 2.161) | <.001 |
| Hyperactivity age 9 | 3.17 (2.76 – 3.58) | <.001 |
| Hyperactivity age 16 | 1.36 (.83 – 1.89) | <.001 |
| Decreasing versus Increasing:Mood age 9 | -1.17 (-1.65 – -.69) | <.001 |
| Mood age 16 | 2.80 (2.18 – 3.43) | <.001 |
| Conduct age 9 | -.63 (-.83 – -.43) | <.001 |
| Conduct age 16 | 1.41 (1.18 – 1.65) | <.001 |
| Internalising age 9 | -.51 (-.75 – -.26) | <.001 |
| Internalising age 16 | 1.06 (.75 – 1.38) | <.001 |
| Peer difficulties age 9 | -.54 (-.75 – -.32) | <.001 |
| Peer difficulties age 16 | .33 (.07 – .58) | <.001 |
| Hyperactivity age 9 | -1.41 (-1.69 – -1.13) | <.001 |
| Hyperactivity age 16 | .94 (.62 – 1.26) | <.001 |
| High versus DecreasingMood age 9 | 2.36 (1.39 – 3.32) | <.001 |
| Mood age 16 | 2.78 (1.48 – 4.08) | <.001 |
| Conduct age 9 | 1.21 (.84 – 1.58) | <.001 |
| Conduct age 16 | 1.91 (1.36 – 2.45) | <.001 |
| Internalising age 9 | .78 (.38 – 1.19) | <.001 |
| Internalising age 16 | 1.46 (.86 – 2.07) | <.001 |
| Peer difficulties age 9 | 1.77 (1.32 – 2.22) | <.001 |
| Peer difficulties age 16 | 1.89 (1.29 – 2.50) | <.001 |
| Hyperactivity age 9 | 1.76 (1.35 – 2.18) | <.001 |
| Hyperactivity age 16 | 2.30 (1.74 – 2.86) | <.001 |

**Table S7: Within-group comparisons: internalising and externalising difficulties, age 9-16 years**

|  |  |  |
| --- | --- | --- |
| **Predictor and outcome variables** | **Mean difference (95% CI)** | ***p*** |
| Persistent Low:Mood  | -.42 (-.52 – -.32) | <.001 |
| Conduct  | -.22 (-.26 – -.18) | <.001 |
| Internalising  | -.10 (-.16 – -.05) | <.001 |
| Peer difficulties  | .04 (-.01 – .08) | .117 |
| Hyperactivity  | -.29 (-.35 – -.24) | <.001 |
| Persistent High:Mood  | -1.70 (-3.31 – -.08) | .040 |
| Conduct  | -.49 (-.97 – .001) | .049 |
| Internalising  | .28 (-.38 – .93) | .400 |
| Peer difficulties  | .10 (-.72 – .52) | .759 |
| Hyperactivity  | -.75 (-1.28 – -.23) | .005 |
| Increasing:Mood  | 1.69 (1.11 – 2.27) | <.001 |
| Conduct | 1.00 (.80 – 1.20) | <.001 |
| Internalising | .95 (.70 – 1.20) | <.001 |
| Peer difficulties | .25 (.06 – .45) | .012 |
| Hyperactivity | .94 (.72 – 1.15) | <.001 |
| Decreasing:Mood | -1.78 (-2.35 – -1.22) | <.001 |
| Conduct | -1.00 (-1.21 – -.79) | <.001 |
| Internalising | -.52 (-.80 – -.24) | <.001 |
| Peer difficulties | -.42 (-.65 – -.19) | <.001 |
| Hyperactivity | -1.34 (-1.61 – -1.06) | <.001 |