Probing the overarching continuum theory:

Data-driven phenotypic clustering of children with ASD or ADHD

Deserno, M.K.[1,3*],
Bathelt, J.[1*],
Groenman, A.P.[1],
Geurts, H.M.[1,2]

[1] Dutch Autism & ADHD Research Centre (d'Arc), Department of Psychology, University of Amsterdam
[2] Dr. Leo Kannerhuis, Amsterdam, The Netherlands
[3] Max Planck Institute for Human Development, Berlin

*these authors contributed equally to this work

Corresponding Author.
Deserno, M.K., Max Planck Institute for Human Development, Berlin, Germany.
Email: marie.deserno@gmail.com

Conflict of Interest.
None.

Financial Support. This work was supported by an Innovational research incentives scheme VICI from the Netherlands Organization for Scientific Research (NWO, VICI Grant No. 453-16-006) awarded to HMG, a Rubicon fellowship from NWO (No. 019.191SG.005) awarded to MKD, and an Amsterdam Brain & Cognition (ABC) Talent Grant awarded to HMG and JB.

Word Count. 4657
Running Head. ASD and ADHD – an overarching continuum?

Abstract

The clinical validity of the distinction between ADHD and ASD is a longstanding discussion. Recent advances in the realm of data-driven analytic techniques now enable us to formally investigate theories aiming to explain the frequent co-occurrence of these neurodevelopmental conditions. In this study, we probe different theoretical positions by means of a pre-registered integrative approach of novel classification, subgrouping and taxometric techniques in a representative sample (N=434) and replicate the results in an independent sample (N=219) of children (ADHD, ASD, and typically developing) aged 7 to 14 years. First, Random Forest Classification could predict diagnostic groups based on questionnaire data with limited accuracy - suggesting some remaining overlap in behavioural symptoms between them. Second, community detection identified four distinct groups, but none of them showed a symptom profile clearly related to either ADHD or ASD in neither the original sample nor the replication sample. Third, taxometric analyses showed evidence for a categorical distinction between ASD and typically developing children, a dimensional characterization of the difference between ADHD and typically developing children and mixed results for the distinction between the diagnostic groups. We present a novel framework of cutting-edge statistical techniques which represent recent advances in both the models and the data used for research in psychiatric nosology. Our results suggest that that ASD and ADHD cannot be unambiguously characterised as either two separate clinical entities or opposite ends of a spectrum and highlight the need to study ADHD and ASD traits in tandem.

The prevalent co-occurrence of Autism Spectrum Disorder (ASD) and Attention-Deficit Hyperactivity Disorder (ADHD) reflects a pressing problem for diagnosis and treatment in child psychiatry (Melegari et al., 2015; Rommelse, Franke, Geurts, Hartman, & Buitelaar, 2010; Simonoff et al., 2008). The two diagnostic categories share etiological factors, overlapping characteristics (e.g., symptoms of inattention and impulsivity; Ronald, Larsson, Anckarsäter & Lichtenstein, 2014) and are both associated with Generalized Anxiety Disorder (GAD), Obsessive-Compulsive Disorder (OCD) and Major Depression (MD; Mulligan et al., 2009; Rommelse et al., 2010). Common practices of small sample size studies and case control models, however, have stalled progress in the pursuit of a better understanding of the discriminant properties between these two neurodevelopmental conditions. Here, we employ a data-driven clustering approach to investigate whether these neurodevelopmental conditions comprise of subtypes that cross clinical boundaries in a large cohort of atypically and typically developing children, and cross-validate our results subsequently.

A growing body of literature concerned with the diagnostic validity of consensus-driven methods (Insel et al., 2010), such as the International Disease Classification [ICD] (World Health Organisation, 2018) and the Diagnostic and Statistical Manual of Mental Disorders [DSM] (American Psychiatric Association, 2013), alludes to an unresolved issue: the intensive search of discriminatory biomarkers for psychiatric conditions, so far, did not result in traceable pathogenic pathways that enable precision medicine and person-centred support. Rather, as both ASD and ADHD remain *behaviourally* diagnosed conditions, about 30-60% of autistic individuals meet diagnostic criteria for ADHD (Carlsson et al., 2013) and about 21-40% of individuals with ADHD meet criteria for an ASD diagnosis (Antshel, Zhang-James, Wagner, Ledesma, & Faraone, 2016; Grzadzinski, Dick, Lord & Bishop, 2016; Rommelse et al., 2010; Ames & White, 2011). With its recent changes, the DSM-5

acknowledges this frequent co-occurrence with revised criteria to explicitly allow for combined diagnosis of ADHD and ASD. Recent work, however, even hypothesizes that these two conditions should not be conceptualized as distinct disorders but rather as manifestations of one overarching disorder with a similar aetiology (Van der Meer et al., 2012). The hypothesis underlying this theory considers ASD to be a manifestation of the most severe subtype on one end of the overarching continuum, while mild ADHD would be located on the other end of this hypothesized continuum. If this hypothesis holds, its theoretical implication would be that ASD cannot exist without ADHD. One would, therefore, expect the categorical classification of individuals as either ASD or ADHD-cases to be difficult since there is a sliding scale between the symptoms rather than two distinct clinical entities. In a parallel line of reasoning, it is often suggested that symptoms related to attention may stem from inhibitory atypicalities when it comes to ADHD, but social atypicalities in the case of ASD phenotypes (Visser, Rommelse, Greven, Buitelaar, 2016; Ingram, Takahashi & Miles, 2008). Also, several studies have shown disorder-specific effects for the exact same psychiatric drugs, such as different normalisation effects in ADHD and ASD of brain dysfunction through serotonin reuptake inhibitors (Chantiluke et al., 2015), while others report shared dysfunction in ADHD and ASD, such as reduced activation in the right PFC indicating shared inhibitory dysfunction (Xiao et al., 2012). The underlying explanation would then be that similar symptoms could have different underlying causes instead of a common cause (Happé & Ronald, 2008; James, Dubey, Smith, Ropar & Tunney, 2016). In the current study, we investigated if behavioural characteristics are sufficient to classify children into diagnostic categories of ADHD and ASD. In a first step, using taxometric analyses, we tested whether the classification performance may arise because of an underlying continuum. Alternatively, the overlap may arise from subgroups within each diagnostic category that

share behavioural features. In a second step we, therefore, investigated this alternative hypothesis using a data-driven clustering approach to identify potential subgroups.

Methods

Participants

The current analysis was based on data obtained from the Child Mind Institute Biobank database (https://childmind.org, date of access: February 21st, 2019). The initial sample consisted of 475 children (ADHD: 249, ASD: 90, TD: 136) between 7 and 13 years of age. This sample is part of a larger cohort of the Healthy Brain Network Biobank based on a community-referred recruitment model of children with developmental psychopathology (Alexander et al., 2017). One participant in the ASD group and 8 participants in the TD group were removed because of missing questionnaire data or missing diagnostic information. Diagnostic classifications were based on extensive clinicians-administered assessments, including the Autism Diagnostic Observation Schedule (ADOS) for suspected autism (Alexander et al., 2017). For the current analyses we used structured questionnaire data from the self-administered assessment protocol entered through the online patient portal (Alexander et al., 2017). Questionnaire measures may show extreme responses that are not related to the content of the questionnaire that arise due to unintended, extreme, fake, or random responses. We, therefore, calculated the Mahalanobis distance to detect and remove these respondents (Zijlstra, van der Ark, & Sijtsma, 2011). According to this measure, 32 participants were identified as multivariate outliers and were removed from the analysis (Mahalanobis distance > 14.07; 18 ADHD, 21 ASD, 2 TD). The final sample consisted of 434 children (231 ADHD, 77 ASD, 126 TD). There were no significant differences in age between the diagnostic groups, but there was a disproportionate number of boys in the ADHD and ASD groups (see Table 1) consistent with the greater prevalence of these diagnoses in males (Loomes, Hull & Mandy,

2017). A third of the children in the ADHD group had an additional diagnosis. The most common were Oppositional Defiant Disorder (n=72 [29.38%]), Autism Spectrum Disorder (n=39 [15.92%]), Specific Learning Disorder with Impairment in Reading (n=34 [13.88%]), Language Disorder (n=33 [13.47%]), Generalized Anxiety Disorder (n=23 [9.39%]), and other less frequent diagnoses (e.g. Enuresis, Specific Phobias, Separation Anxiety; n<20 [≦5%][1]. Around a fifth of children in the ASD group had an additional diagnosis. The most common diagnoses were ADHD-Combined Type (n=36 [14.69%]), ADHD-Inattentive Type (n=13 [5.31%]), and other less frequent diagnoses (e.g., Oppositional Defiant Disorder, Specific Learning Impairment, Generalized Anxiety Disorder; n<10 [≦5%]).

Table 1 Comparison of demographic information between the diagnostic groups.

	ADHD	ASD	TD	Statistics		
N	249	89	128			
Male [%]	186 [42]	65 [12]	68 [58]	χ^2=86.06*, p<0.001	χ^2=36.48*, p<0.001	χ^2=0.79*, p=0.373
				ADHD vs ASD	ADHD vs TD	ASD vs TD
Age [mean+±SE]	9.39±0.109	9.26±0.188	9.41±0.137	t(131)=0.58, p=0.566	t(273)=-0.12, p=0.908	t(173)=-0.63, p=0.527

* Compared to equal split

Pre-registration

The analysis steps (see also Figure 1) and expected results were preregistered before accessing the data. The preregistration can be accessed online (https://aspredicted.org/ya7wr.pdf).

Analysis code

The code for the analyses is available via the Open Science Framework (https://osf.io/vkwma/?view_only=1e66771d9b8c4f1dab7af35918345432).

[1] Please note that multiple comorbid diagnoses were possible, i.e. the percentages are not additive.

[INSERT FIGURE 1 ABOUT HERE]

Figure 1 Overview of the analysis steps. First, item scores were summarized within questionnaire scales to obtain individual profiles. Then, the profiles were used to predict the diagnostic labels using random forest classification. The proximity matrix generated by the random forest classification was used to detect subtypes. In addition, the questionnaire scales that best distinguished the diagnostic groups were used for taxometric analysis to determine if a categorical or a dimensional account provided a better fit to the data.

Materials

The Strengths and Weaknesses of ADHD-symptoms and Normal-behaviors ratings Scale (SWAN; Swanson et al., 2001) is a questionnaire with 18 items that assesses potential strengths and weaknesses related to ADHD symptoms on a single parent-rated scale. It uses items from the Swanson Nolan And Pelham IV (SNAP IV; Swanson, 2003). The SNAP-IV teacher and parent rating scale is often used to assess ADHD symptoms, but the SWAN rephrases the symptoms into strength-based statements making them follow a normal distribution instead of a skewed distribution (Alexander, Salum, Swanson, & Milham, 2019). For example, "Often does not seem to listen when spoken to directly" from the SNAP IV is reworded to "Listens when spoken to directly". The SWAN items are grouped into the Hyperactivity/Impulsivity (HY) and the Inattention (IN) subscale. A validation study of the SWAN indicated high internal consistent (Cronbach's alpha=0.95) and adequate test-retest reliability (r=0.66; Lakes, Swanson, & Riggs, 2012).

The Social Responsiveness Scale (SRS; Constantino & Todd, 2003) is a 65-item scale that is designed to obtain parents- or teacher-ratings of autistic symptomatology as observed in naturalistic social settings. The SRS-assessed symptoms are combined into five subscales: Social Awareness (AWR), Social Cognition (COG), Social Communication (COM), Social Motivation (MOT), and Restricted Interests and Repetitive Behaviours (RRB). Also, the

assessment scale suggests combining these scales into two symptom comparison subscales: the DSM Social communication & interaction subscale and the DSM RBB. We did not include the latter DSM subscale since it is only based on the RRB subscale and, therefore, redundant in the here employed analyses. Validation studies have shown that the SRS has good psychometric properties (3-month test-retest reliability: 0.88, inter-rated reliability: 0.8, correlation with the Autism Diagnostic Interview Revised (ADI-R) score: 0.7; Constantino, Przybeck, Friesen & Todd, 2000, Constantino & Todd, 2003). Of note to the current investigation is that although the SRS was originally designed to produce continuously distributed scores, recent results indicated a bimodal distribution within affected and unaffected family members of children with ASD (Constantino, Zhang, Frazier, Abbacchi & Law, 2010; Virkud, Todd, Abbacchi, Zhang, & Constantino, 2009).

Replication sample. The independent replication sample consisted of 219 children (73 female, ADHD: 87 [39.73%], ASD: 69 [31.51%], TD: 63 [28.77%]) between 8 and 12 years (mean: 10.11, SE: 0.092). For the purpose of replication, we focus solely on the SWAN and SRS data.

Random Forest Classification

First, we applied random forest classification (RFC) to investigate if the selected questionnaire scales can be used to classify participants into diagnostic groups (ADHD, ASD, TD). For multi-class classification, the diagnostic groups were recoded according to a one-versus-all coding scheme, e.g., ADHD vs ASD and TD. The RFC model was trained in a random subsample of 75% of the participants and 25% of the data were held-out for the final validation. To identify the optimal tree depth (i.e., the more splits the more detailed information is explained), bootstrap cross-validation with 10 random resamples was

213 employed. Synthetic minority oversampling (SMOTE) was used to account for class
214 imbalance in the subsets (Chawla, Bowyer, Hall, & Kegelmeyer, 2002) and the area under the
215 receiver-operating characteristic curve (AUROC) averaged across all classes was used to tune
216 the model. These procedures were implemented in R v3.5.2 using the *randomForest* v4.6 (Liaw
217 & Wiener, 2002) and *caret* v.6.0 (Kuhn, 2018) packages. In order to work with the best
218 performing classification approach, we evaluated and compared the classification
219 performance of alternative machine learning approaches (l1-/l2-regularised support vector
220 classification, ridge regression) and cross-validation strategies (k-fold, stratified shuffle split).
221 The machine learning approach presented in the main analysis showed better or equivalent
222 performance as these alternatives (the detailed results are included in the Supplementary
223 Materials).
224 As an adjunct to the random forest classification, we employed an additional method that
225 uses dimensional factors to discriminate between classes. In contrast to the random forest
226 classification that splits participants according to scores on a scale, this approach can assign
227 weights to individuals scales. For instance, ADHD and ASD may share features of executive
228 function difficulties, but this characteristic may be more important, i.e. has a higher weight,
229 for the discrimination of ADHD versus CMP compared to ASD versus CMP. Partial least
230 squares (PLS) analysis creates linear combinations of input variables, in this case
231 questionnaire scales, that are aligned with outcome variables, here one-hot encoded
232 diagnostic labels. An extension of the PLS approach has been developed for classification
233 problems that also incorporates regularisation for better discrimination (Le Cao, Boitard, &
234 Besse, 2011). The current analysis used the implementation in the mixOmics package v6.7.2
235 in R (Rohart, Gautier, Singh & Le Cao, 2017). The full analysis code is available online
236 (https://osf.io/vkwma/).
237

238

239 **Community detection**

240 As a second step, we then employed a community detection approach to investigate the
241 possibility of subtypes across the ADHD-ASD spectrum based on the clinically sensitive
242 questionnaire scales from the RFC in step one. Community detection is an optimization
243 clustering method to detect communities, or subgroups, of nodes (e.g., people), within
244 networks. In the current analysis, the network is based on the RFC proximity matrix which
245 represents the proximity of each participant to all other participants in the sample according
246 to the RFC solution. The proximity indicates how often two participants were assigned to the
247 same leaf node across decision trees in the random forest that aimed to predict the diagnostic
248 label using splits on the questionnaire ratings. The advantage of applying the community
249 detection to the proximity matrix is that the subgroups are necessarily relevant to the
250 diagnostic categorisation (Feczko et al., 2018), whereas grouping based on e.g. the correlation
251 of questionnaire scales may be influenced by other characteristics such as variance of the
252 scale. Here, the Louvain algorithm was used for community detection (Blondel, Guillaume,
253 Lambiotte, & Lefebvre, 2008) followed by a fine-tuning step using the Keringhan-Lin
254 algorithm (Kernighan & Lin, 1970). Due to randomness in the initial assignment of nodes to
255 communities, the algorithm may produce slightly different results at different instantiations.
256 In order to reach a stable assignment, the algorithm was run 50 times to construct an
257 agreement matrix, which was then used to obtain a consensus community partition
258 (Lancichinetti & Fortunato, 2012). We repeated this procedure for multiple resolutions
259 (varying γ between 0.1 and 5.0; Reichardt & Bornholdt, 2006). We selected the solution that
260 provided the best separation and internal consistency of groups (maximal modularity index)
261 while providing the highest agreement across different resolutions (maximal normalized
262 mutual information between successive values of γ). This solution was indicated at resolution

263 γ=0.2. To estimate the reliability of the clustering at this resolution, we repeated the
264 clustering with a randomly selected subset of 80% of the data and compared the results to
265 clustering with the full dataset across 100 repetitions (Tibshirani & Walter, 2005). The results
266 indicated very high stability of the clustering (mutual information: 0.93, 95%-CI: 0.90-98).
267 Both the random forest classification and community detection analyses were repeated in the
268 independent replication sample.
269
270
271
272 **Taxometric Analysis**
273 Because the community clustering solution suggested a dimensional distribution of groups
274 and scores, we conducted an additional exploratory analysis, which was not part of the pre-
275 registration. In subsequent steps we carried out taxometric analysis to assess if a dimensional
276 or categorical account provided a better fit to the questionnaire data including either the
277 diagnostic information or the clustering information. Prior to taxometric analysis, we assessed
278 the suitability of the data (Ruscio, Ruscio, & Carney, 2011). Tables with the corresponding *a*
279 *priori* parameters can be found in the Supplementary Materials. A solution with the three
280 most important indicator variables as determined by the random forest classification (SWAN
281 HY, SRS RRB, SRS AWR) is presented in the main text below. The solution with three
282 indicators is shown, because three indicators are the recommended minimum for taxometric
283 analysis (Ruscio, Ruscio, & Carney, 2011). Solutions with two and four indicators variables
284 can be found in the Supplementary Materials. As recommended in an authoritative review
285 (Ruscio, Ruscio, & Carney, 2011), we used a combination of fit indices for taxometric analysis
286 that are implemented in the RTaxometrics package v2.3 (Ruscio & Wang, 2017). For two
287 indicators, the Mean Above Minus Below A Cut (MAMBAC) and Maximum Slope

(MAXSLOPE) procedure were employed. For three or more indicators, MAMBAC, the maximum eigenvalue (MAXEIG), and the Latent Mode (L-Mode) procedure were used. The consensus result across the procedures is presented in the main text, i.e. the mean comparison curve fit index (CCFI). The comparative curve fit index (CCFI) can be used to investigate if a latent construct is dimensional (CCFI < 0.4) or categorical (>0.6; Ruscio, Ruscio & Carney, 2011) through comparison to simulated data in parallel analysis. Results for each separate procedure can be found in the Supplementary Materials.

Statistical analysis

Group-wise comparisons were based on Welch-corrected t-tests that account for differences in variance between the groups. Bonferroni correction was used to account for multiple comparisons and corrected p-values are reported in the main text.

Results

Diagnostic groups show different profiles on questionnaires of social communication and ADHD symptoms

The diagnostic groups showed different profiles of scores on the SRS and SWAN questionnaires (analysis of variance (ANOVA) – group: $F(2, 3017)=801.8$, $p<0.001$; group x scale: $F(12, 3016)=10.7$, $p<0.001$). While both diagnostic groups showed higher scores compared to the TD group across all questionnaire scales (see Table 2 & Figure 2), the ASD group scored higher on the SRS compared to the ADHD group. In contrast, there was no significant difference between the ADHD and ASD group for any of the SWAN subscales. Highly similar results were obtained in the replication sample (see Figure S1).

Table 2 Comparison of questionnaire profiles between the ADHD, ASD, and TD groups. Significant differences are shown in bold print.

	ADHD		ASD		TD		ADHD vs TD			ASD vs TD			ADHD vs ASD		
	mean	SE	mean	SE	mean	SE	t	df	d	t	df	d	t	df	d
AWR	0.17	0.056	0.83	0.084	-0.81	0.073	-10.66	266.04	**-1.18**	-14.76	173.93	**-2.12**	-6.55	149.33	**-0.83**
COG	0.10	0.058	0.97	0.093	-0.78	0.057	-10.80	329.26	**-1.14**	-16.04	132.32	**-2.40**	-7.92	140.95	**-1.02**
COM	0.13	0.056	1.00	0.087	-0.85	0.055	-12.45	328.52	**-1.32**	-18.00	136.23	**-2.68**	-8.40	145.44	**-1.08**
MOT	0.04	0.060	0.85	0.117	-0.60	0.063	-7.42	313.71	**-0.79**	-10.92	120.30	**-1.66**	-6.15	118.33	**-0.84**
RRB	0.11	0.059	1.02	0.091	-0.82	0.044	-12.51	354.97	**-1.28**	-18.17	112.51	**-2.78**	-8.40	146.28	**-1.08**
HY	0.44	0.040	0.35	0.087	-1.02	0.086	-15.40	181.81	**-1.81**	-11.21	188.34	**-1.58**	0.98	111.12	0.14
IN	0.39	0.044	0.49	0.081	-1.02	0.082	-15.17	198.00	**-1.76**	-13.11	190.97	**-1.85**	-1.05	124.55	-0.14

Abbreviations: d – Cohen's d, df – degrees of freedom, SE – standard error, AWR – Social Awareness, COG – Social Cognition, COM – Social Communication, MOT – Social Motivation, RRB – Restricted Interests and Repetitive Behaviours, HY – Hyperactivity/Impulsivity, IN – Inattention.

Random Forest Classification can predict diagnostic groups based on questionnaire data with some accuracy

Our results indicated that the optimal classification accuracy was achieved at a tree-depth of two, i.e. two questionnaire scales were sufficient to discriminate the groups. Cross-validation supported that a tree depth of two was optimal for classification. At this depth, the accuracy of the model for the training set was 87% (CI= 83.03-90.44, κ= 0.79, McNemar's p-value=$4.6e^{-8}$) and 72% for the test set[2] (CI=62.76-80.17%, κ=0.56, McNemar's p-value=0.037; f1-score: 0.72, precision: 0.77, recall: 0.79). Sensitivity and specificity of the model indicated that diagnostic groups could be distinguished (see **Table 3**, ADHD: 0.67/0.84; ASD: 0.68/0.78; TD: 0.83/0.96 [sensitivity/specificity]). The most important scales for classification were SWAN HY (variable important as indicated by the percentage of trees that used the variable to split for classification: 100%) and SRS RRB (77.27%), followed by SRS AWR (63.37%), SRS COG (62.65%), SRS COM (57.20%), SRS IN (29.17), and

[2] Classification using sPLS-DA led to similar accuracy: 72% accuracy in the validation dataset, 73.87% when comorbid cases were excluded. The details of the full analysis are presented in the Supplementary Materials.

SRS MOT (0.00%). The accuracy of the classification model was similar when applied to the independent replication sample (overall accuracy: 76%, ADHD: 0.69/0.84, ASD: 0.68/0.94, TD: 0.94/0.85 [sensitivity/specificity]).

When excluding comorbid cases (ASD with a secondary diagnosis of ADHD or vice versa), the random forest classification reached an accuracy of 94.41% (n=340, CI: 91.41-96.61%) in the training set and 71.17% (n=111, CI: 61.81-79.37%; f1-score: 0.787, precision: 0.787, recall: 0.787) in the held-out test set. The specificity and sensitivity were acceptable for all classes (sensitivity/specificity, ADHD: 0.77/0.74; ASD: 0.63/0.82; TD: 0.65/0.98). Without the cases with a dual diagnosis the SRS RRB seemed less important. The most important scales for classification were SRS HY (100%) followed by SRS COM (76.59%), SRS AWR (72.65%), SRS COG (71.68%), SRS RRB (58.53%), SWAN IN (42.47%), and SRS MOT (0.00%).

Table 3 Confusion matrix for the test data. Rows indicate the predicted (Pred.) diagnostic group, columns indicate the actual diagnostic group (Reference [Ref.]).

Pred. / Ref.	ADHD	ASD	TD
ADHD	41	6	2
ASD	17	13	3
TD	3	0	26

Community detection identifies subgroups that cross diagnostic boundaries

The community solution consisted of five groups with four large groups (see Figure 2, C1: n=141 [31.26%], C2: n=86 [19.7%], C4: n=85 [18.85%], C5: n=136 [31.16%]) and one small group[3] (C3: n=3 [0.67%]). The community detection algorithm converged at a stable solution that showed a good separation between the identified groups (Q=0.92). The four

[3] Please note that we do not include this group (C3) in between-group comparisons since there are only three people in that group.

large groups showed different profiles of questionnaire scores (ANOVA: group – $F(2, 3108)=738.01$, $p<0.001$; group x scale: $F(18, 3109)=63.49$, $p<0.001$, see Figure 2 & Table 4). One group (C2: low symptoms) scored around 1 standard deviation (SD) below the other groups across all questionnaire scales and mostly contained children without a diagnosis (TD: 79 [85.11 %], ADHD: 5 [10.64 %], ASD: 2 [4.26%], comparison of proportions to the whole sample: $\chi^2=288.10$, $p<0.001$). A second group (C5: high symptoms) had scores around 1 SD above the mean and consisted of two thirds of children with ADHD and one third children with ASD (TD: 3 [2.92%], ADHD: 92 [67.15%], ASD: 41 [29.93%], $\chi^2=226.05$, $p<0.001$). The other groups had contrasting symptom profiles. One group (C1: SWAN ↑) showed low symptoms on the SRS scales, but high symptoms on the SWAN scales and consisted mostly of children with ADHD (TD: 15 [10.64%], ADHD: 120 [81.11%], ASD: 6 [5.26%]). Another group (C4: SRS ↑) showed elevated symptoms on the SRS scales with lower ratings on the SWAN scales and consisted to equal proportion of children from all diagnostic categories (TD: 29 [34.12%], ADHD: 27 [31.76%], ASD: 29 [34.12%]).

The different groups were associated with differences in demographics and comorbid profiles: children in the cluster with higher SRS scores (C4) were slightly older compared to the rest of the sample, there were more females in the cluster with low symptoms (C2) and more males in the cluster with high symptoms (C5). The other clusters did not deviate in sex ratio or age from the rest of the sample (see Table 5). Furthermore, the cluster with low symptoms (C2) and the cluster with relatively high SWAN scores (C1) contained fewer cases with a dual diagnosis of ASD and ADHD than would be expected given the proportion observed across the whole sample (see Table 5). In contrast, the cluster with high symptoms (C5) and the cluster with high SRS scores (C4) contained more ADHD-ASD comorbid cases than expected (see Table 5).

[INSERT FIGURE 2 ABOUT HERE]

Figure 2 Profiles of diagnostic groups and groups identified through community clustering. **A)** The proximity between participants according to the random forest classification is shown in Force Atlas layout (Jacomy, Venturini, Heymann, & Bastian, 2014) either coloured according to the diagnostic group (top) or according to the groups identified through community detection (bottom). The smaller plots show the proximity matrix ordered according to either diagnostic or community detection labels. The figure illustrates the separation and overlap of the diagnostic groups as seen by the RFC algorithm. **B)** Profiles of the groups according to diagnosis (left) or community detection (right). The lower part of the figure shows the effect size of comparisons between the group. The circular plots in the right figure indicate the relative proportion of diagnoses within the groups identified through community detection. The error bars indicate one standard error around the mean. Abbreviations: AWR – Social Awareness, COG – Social Cognition, COM – Social Communication, MOT – Social Motivation, RRB – Restricted Interests and Repetitive Behaviours, HY – Hyperactivity/Impulsivity, IN – Inattention.

Table 4 Comparison of questionnaire profiles between the community clustering-defined groups. Significant differences are shown in bold print. Abbreviations: d – Cohen's d, df – degrees of freedom, SE – standard error, AWR – Social Awareness, COG – Social Cognition, COM – Social Communication, MOT – Social Motivation, RRB – Restricted Interests and Repetitive Behaviours, HY – Hyperactivity/Impulsivity, IN – Inattention.

	C1 (SWAN↑)		C2 (low symp)		C4 (SRS↑)		C5 (high symp)	
	mean	SE	mean	SE	mean	SE	mean	SE
AWR	-1.11	0.068	-0.15	0.055	-1.11	0.068	0.80	0.069
COG	-1.03	0.053	-0.33	0.055	-1.03	0.053	0.81	0.070
COM	-1.07	0.052	-0.35	0.047	-1.07	0.052	0.90	0.067
MOT	-0.81	0.067	-0.38	0.053	-0.8	0.067	0.68	0.083
RRB	-1.08	0.018	-0.53	0.033	-1.08	0.018	1.09	0.055
HY	-1.43	0.098	0.34	0.039	-1.43	0.098	0.84	0.040
IN	-1.32	0.098	0.22	0.055	-1.32	0.098	0.68	0.055

	C1 vs C2			C1 vs C4			C1 vs C5		
	t	df	d	t	df	d	t	df	d
AWR	-10.92	183.37	**-1.50**	-10.43	150.88	**-1.61**	-19.73	209.51	**-2.66**
COG	-9.19	216.36	**-1.22**	-11.94	126.56	**-1.84**	-20.97	219.31	**-2.74**
COM	-10.24	199.40	**-1.38**	-11.40	124.77	**-1.76**	-23.08	219.76	**-3.02**
MOT	-4.95	182.05	**-0.68**	-8.89	137.81	**-1.37**	-13.87	219.97	**-1.83**
RRB	-14.67	207.41	**-1.84**	-14.35	90.62	**-2.21**	-37.63	163.19	**-4.67**
HY	-16.74	111.86	**-2.46**	-9.18	118.14	**-1.41**	-21.41	113.89	**-3.15**
IN	-13.78	138.47	**-1.97**	-9.51	161.87	**-1.46**	-17.86	138.46	**-2.56**

	C2 vs C4			C2 vs C5			C4 vs C5		
	t	df	d	t	df	d	t	df	d
AWR	-2.56	137.27	-0.37	-10.90	259.47	**-1.32**	-5.64	163.70	**-0.79**
COG	-5.72	134.73	**-0.82**	-12.74	258.47	**-1.54**	-3.90	161.51	**-0.55**
COM	-5.26	119.82	**-0.77**	-15.23	241.72	**-1.84**	-5.39	155.32	**-0.76**
MOT	-5.93	122.75	**-0.86**	-10.72	230.38	**-1.30**	-2.34	170.97	-0.33
RRB	-8.15	105.72	**-1.21**	-25.47	220.97	**-3.09**	-7.78	143.59	**-1.11**
HY	13.23	193.68	**1.80**	-9.07	274.15	**-1.09**	-21.47	196.52	**-2.93**
IN	3.66	162.57	**0.51**	-5.93	274.88	**-0.71**	-8.45	162.18	**-1.19**

Table 5 Comparison of demographic information between the clusters. For the statistical analysis, groups were compared to the frequencies observed across the whole sample regarding sex and comorbidity, and to the rest of the sample regarding age.

	C1 (SWAN↑)	C2 (low sym)	C4 (SRS↑)	C5 (high sym)
N	141	86	85	136
Male [%]	110 [78.72%]	45 [52.33%]	60 [70.59%]	113 [83.09%]
Stat.	χ^2=2.05, p=0.152	χ^2=19.55, p<0.001	χ^2=0.34, p=0.559	χ^2=6.55, p=0.011
Age [mean±SE]	9.34±0.144	9.47±0.161	9.72±0.186	9.22±0.133
Stat.	t(254.21)=-0.44 p=0.659	t(139.96)=0.5 p=0.617	t(120.23)=1.98 p=0.05	t(272.61)=-1.54 p=0.124
n comor. [%]	8 [5.67%]	0 [0.00%]	24 [28.24%]	60 [44.12%]
Stat.	χ^2=19.66, p<0.001	χ^2=22.64, p<0.001	χ^2=2.82, p=0.093	χ^2=44.66, p<0.001

Taxometric Analyses

Given these results, we next tested whether based on such discriminatory measures (SRS and SWAN) taxometric analyses would yield clear evidence in favour of either a dimensional or a categorical account of the differences between the diagnostic groups. The comparative curve fit index (CCFI) can be used to investigate if a latent construct is dimensional (CCFI < 0.4) or categorical (>0.6; Ruscio, Ruscio & Carney, 2011) through comparison to simulated data in parallel analysis. Across different measures of curve fit, we found support for a dimensional distribution when including the typical and *both* atypical groups (mean=0.36), and when including the ADHD and TD groups (mean=0.35). The ASD-TD comparison was consistent with a categorical account (mean=0.76). There was no strong support for either a categorical or a dimensional account for the comparison between ADHD and ASD (mean=0.49, all based on 3 indicators; see Supplementary Materials for similar results obtained with 2 or 4 indicators). To test if the results were influenced by edge cases, we conducted a further taxometric analysis that only included cases that were assigned to one of the major clusters in the consensus community clustering analysis. Taxometric analysis indicated that all groups identified through community clustering were more compatible with a dimensional than a categorical account (see Table 6).

Table 6 Comparative curve fit index (CCFI) using three indicators for comparisons of the community clustering defined groups. CCFI values <0.4 indicate a dimensional distribution, CCFI >0.6 are more compatible with a categorical account.

	l-mode	mambac	maxeig	mean
C1 vs C2	0.37	0.33	0.34	0.35
C1 vs C4	0.37	0.33	0.35	0.35
C1 vs C5	0.33	0.35	0.37	0.35
C2 vs C4	0.33	0.35	0.39	0.36
C2 vs C5	0.36	0.31	0.34	0.34
C4 vs C5	0.37	0.34	0.36	0.36

Discussion

By adopting multiple analytical routes to subtyping, we investigate subgroups within a large cohort of typically and atypically developing children that either (a) represent a taxometric difference between ADHD and ASD or (b) indicate an underlying condition with ADHD and ASD as opposite ends of a dimension. Our results suggest that neither a *categorical* nor a *dimensional* characterization of the indicators used in this study (standard symptom scales) to define ADHD and ASD is more sensible than the other. In other words, the autistic and ADHD related behavioural traits as assessed in the current sample cannot, unambiguously, be characterized as either two separate clinical entities or as two opposite ends of a spectrum. In contrast, we show that the difference between ADHD children and typical children in the current sample is *dimensional*, while the difference between ASD children and typical children can best be characterized as *categorical*. Whereas our results do not support recent literature arguing for an underlying dimension that explains the frequent overlap of the two conditions, they do highlight the importance of studying ASD and ADHD in tandem, as has been suggested by many developmental researchers before (e.g., Rommelse et al., 2011; Van der Meer, 2012; Johnson, Gliga, Jones, & Charman, 2015; Geurts et al., 2004). The classification algorithm we applied was able to classify the diagnostic groups to some extent, but never reaches particularly high accuracy in distinguishing them. Moreover, we find that even with community detection techniques - focused on making detected groups most distinct - results

444 do not show separate groups of ADHD vs. ASD, but suggesting that the behavioural
445 symptom scales are not sufficient to fully distinguish the diagnostic group. This is in line with
446 current clinical practice in which clinicians are trained to not only base their diagnosis on
447 these type of proxy reports of behavioural symptoms but take additional factors into account.
448 Essentially, the current study underlines the potential of studying risk factors in
449 relation to both, ADHD and ASD symptoms. First, our community detection and taxometric
450 results suggests that ADHD and ASD cannot, unambiguously, be characterized as two
451 separate clinical entities or as two opposite ends of a spectrum based on behavioural traits.
452 However, although the definitions of ADHD and ASD are based on behavioural traits, they
453 are associated with a wide range of atypicalities in other areas such as neurobiology and
454 genetics. Taking these into account could lead to more clear-cut results concerning the
455 distinction between ADHD and ASD. Nevertheless, as all identified clusters contain some
456 combination of ADHD and ASD diagnoses, clinically, our results imply that screening for
457 ADHD in ASD is imperative. Theoretically, our results underline taking a dimensional
458 approach that could advance knowledge about genetical, brain, cognitive, and behavioral
459 underpinnings of symptomatology. Dimensional analyses, however, are only useful when it
460 can be demonstrated that the association of predictors with dimensional scores are constant
461 throughout the relevant dimensional severity range (Kessler, 2002). In order to draw strong
462 clinical policy-related conclusions such dimensionality first needs to be justified by
463 demonstrating the absence of non-linear effects outside the clinical range that cause
464 predictors to be significant for dimensional scores (Kessler, 2002). Moreover, we show that a
465 completely dimensional view might be adequate for the relation between typical developing
466 children and ADHD children but does not do just to the complexity of the relationship
467 between ADHD and ASD. Given the dimensional characterization that our results suggest
468 regarding ADHD, the clinician will still need external criteria, such as impairment or

suffering, to determine cut points on such dimensional measures that indicate the existence of impaired functioning.

Second, this study is methodologically a first in the developmental literature. It is the first to adopt a multiverse approach with replication of cutting-edge subtyping *and* taxometric procedures to shed new light on one of the oldest psychometric issues in the field of atypical development research, the question of whether mental disorders should be thought of as discrete categories or as continua. In their review of psychometric modelling approaches, Borsboom and colleagues (2016) note that the psychometric work related to this issue has not been able to put forward a systematic methodological procedure to investigate the kind vs. continua question. The authors suggest that this might be due to the limited range of hypotheses tested by common approaches as these procedures do not test the exhaustive hypotheses space of latent structures, but treat the potential answer as binary: evidence *in favor* of categorical distinction is treated as evidence *against* the hypothesis of a dimensional structure leaving no room for other (hybrid) possibilities, such as some alternative factor mixture models or network models do. We here proposed a combined framework of analytic steps that cover a wider hypothesis space from different methodological angles avoiding the abovementioned issue. However, even with such systematic methodological procedure, we were unable to yield clear results regarding the question whether ADHD and ASD lie along an underlying continuum.

Third, previous autism research has suggested a taxon higher up the proposed gradient scale than DSM classification suggests, i.e. a 'highly severe' ASD subgroup (Frazier et al., 2014). We find that our taxometric results are ambiguous when performed comparing the ASD children with the ADHD children, instead of in all three groups. This ambiguity may be explained by the presence of a such a specific ASD subgroup. Also, the fact that the diagnostic procedure used in this sample was contingent on a clinician's suspicion of ASD (see

494 Alexander et al., 2017) might explain the 70/30 division of ADHD/ASD diagnoses in the
495 high symptom subgroup we identified. Nevertheless, our taxometric analysis underlines the
496 dimensional account of ADHD symptoms in typical children.
497
498 **Limitations**
499 Despite several strong points of this study, including pre-registration, cutting-edge statistical
500 techniques, a large sample size, and replication in an independent sample, several limitations
501 should be considered when interpreting our results. First, it should be stressed that our
502 analyses are based on validated ASD and ADHD symptom scales reflecting a wide range of
503 behaviors and symptoms. Naturally, however, this focus does not cover all potential
504 tributaries to ASD and ADHD phenotypes, such as neuropathological and genetic factors
505 (Rutter, 2013). Also, the SRS scale used in the current study mainly covers the ASD social
506 domains, with only a few indicators of repetitive and restrictive behaviors and no assessment
507 of the sensory sensitivities that often go along with ASD. The literature, however, suggests
508 significant clinical difference between ADHD and ASD samples on this specific domain:
509 reports of repetitive behaviours in ADHD are less frequent than reports of communicative
510 and social difficulties (Nijmeijer et al., 2008). Another large epidemiological study reports that
511 repetitive and restrictive behaviors explain a substantial part of the co-occurrence of ASD and
512 ADHD traits (Polderman, Hoekstra, Posthuma & Larsson, 2014). Future studies should,
513 therefore, include extensive assessments of the whole range of symptoms.
514 Additionally, it should be noted that a taxometric approach to unveiling the latent
515 structure of psychological conditions is not uncontroversial in psychometrics (Lubke & Miller,
516 2015; Borsboom et al., 2016). We here explicitly accommodate all recent advances and
517 recommendations by adopting taxometric procedures based on simulation (Ruscio et al.
518 2017) to deal with exceptions in its core assumptions (i.e. the assumption that categorical

structures produce peaked covariance functions might not be true under certain conditions; Molenaar, Dolan & Verhelst, 2010). Our results are, furthermore, based on (i) a large sample to make sure sampling fluctuation has less impact (Lubke & Neale, 2008) and (ii) symptom scales with varying endorsement probabilities of their items (Lubke & Miller, 2015). Although we combine different symptom scales with different response formats for our taxometric analyses, we chose to only include the combined, continuous subscales of the SRS to make sure the response range taps into the gradual differences of scale (Hay, Bennett, Levy, Sergeant & Swanson, 2007).

Third, current research on ADHD and ASD is highly skewed towards childhood, including this study. There are strong indications that the co-occurrence between ADHD and ASD is dependent on age (for review see Hartman 2016). For example, genetic research (Stergiakouli et al., 2017) indicates that although the biological etiology of the symptoms is dependent on similar biological pathways that the influences of these pathways ADHD co-varies throughout development. Therefore, longitudinal research is warranted.

Conclusion

In conclusion, this study supports those voices in the literature that are doubting the categorical differences between the consensus-based sets of ADHD symptoms and ASD symptoms, however we also cannot, unambiguously state that ADHD and ASD should be characterized as two opposite ends of a spectrum or as two separate clinical entities. In the long run, the statistical developments might result in a non-binary answer to the kind vs. continua question in psychiatry based on a novel way of conceptualizing non-linear transitions between different psychiatric conditions that follow from the complex interplay of their symptoms and the individual environment. For now, it is unambiguous that ADHD and ASD traits need to be studied in tandem.

References

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., ... & Litke, S. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific data, 4*, 170181.

Alexander, L. M., Salum, G. A., Swanson, J. M., & Milham, M. P. (2019). Measuring strengths and weaknesses in dimensional psychiatry. *Journal of Child Psychology and Psychiatry, 61*(1), 40-50.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC.

Ames, C. S., & White, S. J. (2011). Brief report: Are ADHD traits dissociable from the autistic profile? Links between cognition and behaviour. *Journal of Autism and Developmental Disorders, 41*(3), 357-363.

Antshel, K. M., Zhang-James, Y., Wagner, K. E., Ledesma, A., & Faraone, S. V. (2016). An update on the comorbidity of ADHD and ASD: A focus on clinical management. *Expert review of neurotherapeutics, 16*(3), 279-293.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008): "Fast unfolding of communities in large networks". *Journal of Statistical Mechanics* P10008 (2008). https://doi.org/ 10.1088/1742-5468/2008/10/P10008

Borsboom, D., Rhemtulla, M., Cramer, A. O. J., van der Maas, H. L., Scheffer, M., & Dolan, C. V. (2016). Kinds versus continua: a review of psychometric approaches to uncover the structure of psychiatric constructs. *Psychological medicine, 46(*8), 1567-1579.

Carlsson, L.H., Norrelgen, F., Kjellmer, L., Westerlund, J., Gillberg, C., & Fernell, E. (2013). Coexisting disorders and problems in preschool children with autism spectrum disorders. The Scientific World Journal, 2013.

Chantiluke, K., Barrett, N., Giampietro, V., Santosh, P., Brammer, M., Simmons, A., ... & Rubia, K. (2015). Inverse fluoxetine effects on inhibitory brain activation in non-comorbid boys with ADHD and with ASD. *Psychopharmacology, 232*(12), 2071-2082.

Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. (2002): "SMOTE: Synthetic Minority Over-sampling Technique", *Journal of Artificial Intelligence Research 16 (2002),* 321-357. https://doi.org/10.1613/jair.953

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000): "Reciprocal social behavior in children with and without perva- sive developmental disorders". *Journal of Developmental and Behavioral Pediatrics, 21,* 2–11. http://dx.doi.org/10.1097/00004703-200002000-00001

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60,* 524–530. https://doi.org/10.1001/archpsyc.60.5.524

Constantino, J. N., Zhang, Y., Frazier, T., Abbacchi, A. M., Law, P. (2010): "Sibling recurrence and the genetic epidemiology of autism". *American Journal of Psychiatry*, *167*, 1349–1356, https://doi.org/10.1176/appi.ajp.2010.09101470

Feczko, E., Balba, N. M., Miranda-Dominguez, O., Cordova, M., Karalunas, S. L., Irwin, L., Demeter, D. V., Hill, A. P., Longhorst, B. H., Grieser Painter, J., Santen, J. Van, Fombonne, E. J., Nigg, J. T., Fair, D. A. (2018): "Subtyping cognitive profiles in autism spectrum disorders using a functional random forest algorithm". *NeuroImage, 172*, 674-688, https://doi.org/10.1016/j.neuroimage.2017.12.044

Feczko, E., Miranda-Dominguez, O., Graham, A. M., Nigg, J. T., Fair, D. A. (2019): "The heterogeneity problem: Approaches to identify psychiatric subtypes" *Trends in Cognitive Science, 23(7)*, 584-601, https://doi.org/10.1016/j.tics.2019.03.009

592 Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014).
593 Confirmatory factor analytic structure and measurement invariance of quantitative autistic
594 traits measured by the Social Responsiveness Scale-2. *Autism, 18*(1), 31-44.
595 Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., Hartman, C. A., Mulder, E. J., ... &
596 Sergeant, J. A. (2004). Can the Children's Communication Checklist differentiate between
597 children with autism, children with ADHD, and normal controls?. *Journal of Child Psychology*
598 *and Psychiatry, 45*(8), 1437-1453.
599 Grzadzinski, R., Dick, C., Lord, C., & Bishop, S. (2016). Parent-reported and clinician-
600 observed autism spectrum disorder (ASD) symptoms in children with attention
601 deficit/hyperactivity disorder (ADHD): implications for practice under DSM-5. *Molecular*
602 *autism, 7*(1), 7.
603 Happé, F., & Ronald, A. (2008). The 'fractionable autism triad': a review of evidence from
604 behavioural, genetic, cognitive and neural research. *Neuropsychology review, 18*(4), 287-304.
605 Hartman, C. A., Geurts, H. M., Franke, B., Buitelaar, J. K., & Rommelse, N. N. (2016).
606 Changing ASD-ADHD symptom co-occurrence across the lifespan with adolescence as
607 crucial time window: illustrating the need to go beyond childhood. *Neuroscience &*
608 *Biobehavioral Reviews, 71*, 529-541.
609 Hay, D. A., Bennett, K. S., Levy, F., Sergeant, J., & Swanson, J. (2007). A twin study of
610 attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses
611 of ADHD-symptoms and normal-behavior (SWAN) scale. *Biological Psychiatry, 61*(5), 700-
612 705.
613 Ingram, D. G., Takahashi, T. N., & Miles, J. H. (2008). Defining autism subgroups: a
614 taxometric solution. *Journal of autism and developmental disorders, 38*(5), 950-960.
615 Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Kozak, M., Pine, D.S., Quinn, K.,
616 Sanislow, C., Steinberg, J., & Wang, P. (2010). Research Domain Criteria (RDoC): toward

a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*, 748-751.

Jacomy, M., Venturini, T., Heymann, S., Bastian, M. (2014): "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software" *PLoS ONE* 9(6), e98679, https://doi.org/10.1371/journal.pone.0098679

James, R. J., Dubey, I., Smith, D., Ropar, D., & Tunney, R. J. (2016). The latent structure of autistic traits: A taxometric, latent class and latent profile analysis of the adult autism spectrum quotient. *Journal of autism and developmental disorders, 46*(12), 3712-3728.

Johnson, M. H., Gliga, T., Jones, E., & Charman, T. (2015). Annual Research Review: Infant development, autism, and ADHD–early pathways to emerging disorders. *Journal of Child Psychology and Psychiatry, 56*(3), 228-247.

Kernighan, B.W. & Lin, S. (1970): "An efficient heuristic procedure for partitioning graphs". *Bell System Technical Journal* 49, 291-307, https://doi.org/10.1002/j.1538-7305.1970.tb01770.x

Kessler, R. C. (2002). The categorical versus dimensional assessment controversy in the sociology of mental illness. *Journal of Health and Social Behavior*, 171-188.

Kuhn, M. (2018): "caret: Classification and regression training." R package version 6.0-81. https://CRAN.R-project.org/package=caret

Lakes, K. D., Swanson, J. M., Riggs, M. (2012). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal Behavior Raring Scale in preschool sample: continuum measures of hyperactivity and inattention., Journal of Attention Disorder, 16(6), 510-516, https://doi.org/10.1177/1087054711413550

Lancichinetti, A., & Fortunato, S. (2012): "Consensus clustering in complex networks", *Scientific Reports.* https://doi.org/10.1038/srep00336

Le Cao, K.-A., Boitard, S., & Besse, P. (2011). Sparse PLS discriminant analysis: biologically relevant feature selection in graphical displays for multiclass problems. *BMC Bioinformatics*, 12(1), 253, https://doi.org/10.1186/1471-2105-12-25

Liaw, A. & Wiener, M. (2002): "Classification and regression by randomForest", *R News 2*(3), 18-22.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(6), 466-474.

Lubke, G., & Neale, M. (2008). Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research, 43*(4), 592-620.

Lubke, G.H., & Miller, P. J. (2015). Does nature have joints worth carving? A discussion of taxometrics, model-based clustering and latent variable mixture modeling. *Psychological medicine, 45*(4), 705-715.

Melegari, M. G., Nanni, V., Lucidi, F., Russo, P. M., Donfrancesco, R., & Cloninger, C. R. (2015). Temperamental and character profiles of preschool children with ODD, ADHD, and anxiety disorders. *Comprehensive Psychiatry, 58*, 94-101.

Mulligan, A., Anney, R. J., O'Regan, M., Chen, W., Butler, L., Fitzgerald, M., ... & Nijmeijer, J. (2009). Autism symptoms in attention-deficit/hyperactivity disorder: a familial trait which correlates with conduct, oppositional defiant, language and motor disorders. *Journal of Autism and Developmental Disorders, 39*(2), 197-209.

Molenaar, D., Dolan, C. V., & Verhelst, N. D. (2010). Testing and modelling non-normality within the one-factor model. *British Journal of Mathematical and Statistical Psychology, 63*(2), 293-317.

665 Nijmeijer, J. S., Hoekstra, P. J., Minderaa, R. B., Buitelaar, J. K., Altink, M. E., Buschgens,
666 C. J. M., ... & Hartman, C. A. (2009). PDD symptoms in ADHD, an independent familial
667 trait?. *Journal of abnormal child psychology, 37*(3), 443-453.
668 Polderman, T. J. C., Hoekstra, R. A., Posthuma, D., & Larsson, H. (2014). The co-
669 occurrence of autistic and ADHD dimensions in adults: an etiological study in 17 770
670 twins. *Translational psychiatry, 4*(9), e435.
671 Reichhardt, J. & Bornholdt, S. (2006): "Statistical mechanics of community detection"
672 *Physical Review E* 1(74), 016110. https://doi.org/10.1103/PhysRevE.74.016110
673 Rohart, F., Gautier, B., Singh, A., & Le Cao, K.-A. (2017). mixOmics: An R package for
674 'omics feature selection and multiple data integration. *PLoS Computation Biology* 13(11),
675 e1005752.
676 Rommelse, N. N., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J. K. (2010).
677 Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum
678 disorder. *European child & adolescent psychiatry, 19*(3), 281-295.
679 Ronald, A., Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2014). Symptoms of autism and
680 ADHD: a Swedish twin study examining their overlap. *Journal of Abnormal Psychology,*
681 *123*(2), 440.
682 Ruscio, J., Ruscio, A.M., Carney, L.M. (2011): "Performing taxometric analysis to distinguish
683 categorical and dimensional variables" *J Exp Psychopathol*, 2 (2011),170-196,
684 https://doi.org/10.5127/jep.010910
685 Ruscio, J., Wang, S. (2017): "RTaxometrics: Taxometric Analysis", R package version 2.3,
686 https://CRAN.R-project.org/package=RTaxometrics
687 Rutter, M. (2013). Changing concepts and findings on autism. *Journal of autism and developmental*
688 *disorders, 43*(8), 1749-1757.

689 Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008).
690 Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity,
691 and associated factors in a population-derived sample. *Journal of the American Academy of Child*
692 *& Adolescent Psychiatry, 47*(8), 921-929.
693 Stergiakouli, E., Smith, G. D., Martin, J., Skuse, D. H., Viechtbauer, W., Ring, S. M., ... & St
694 Pourcain, B. (2017). Shared genetic influences between dimensional ASD and ADHD
695 symptoms during child and adolescent development. *Molecular autism, 8*(1), 18.
696 Swanson, J. M. (2003). The SNAP-IV teacher and parent rating scale. In Therapist's guide to
697 learning and attention disorders (pp. 487-500). Academic Press.
698 Tibshirani & Walther (2005): "Cluster Validation by Prediction Strength", *Journal of*
699 *Computational and Graphical Statistics*, 14(3), 511-528,
700 https://doi.org/10.1198/106186005X59243
701 Van der Meer, J. M., Oerlemans, A. M., van Steijn, D. J., Lappenschaar, M. G., de
702 Sonneville, L. M., Buitelaar, J. K., & Rommelse, N. N. (2012). Are autism spectrum
703 disorder and attention-deficit/hyperactivity disorder different manifestations of one
704 overarching disorder? Cognitive and symptom evidence from a clinical and population-
705 based sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(11), 1160-
706 1172.
707 Virkud, Y. V., Todd, R. D., Abbacchi, A. M., Zhang, Y., Constantino, J. N. (2009): "Familial
708 aggregation of quantitative autistic traits in multiplex versus simplex autism". *American*
709 *Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 150(3)*, 328–334. https://doi.org/
710 10.1002/ajmg.b.30810
711 Visser, J. C., Rommelse, N. N., Greven, C. U., & Buitelaar, J. K. (2016). Autism spectrum
712 disorder and attention-deficit/hyperactivity disorder in early childhood: a review of unique

and shared characteristics and developmental antecedents. *Neuroscience & Biobehavioral Reviews, 65*, 229-263.

World Health Organization. (2018). International statistical classification of diseases and related health problems (11th Revision). Retrieved from https://icd.who.int/browse11/l-m/en

Zijlstra, W. P., Ark, A. van der, & Sijtsma, K. (2011): "Outliers in questionnaire data: Can they be detected and should they be removed?" *Journal of Education and Behavioral Statistics 35(2),* 186-212, https://doi.org/10.3102/1076998610366263