Title: Differential functional neural circuitry behind autism subtypes with marked imbalance between social-communicative and restricted repetitive behavior symptom domains

Authors: Natasha Bertelsen[1,2], Isotta Landi[1], Richard A. I. Bethlehem[3], Jakob Seidlitz[4,5], Elena Maria Busuoli[1,2], Veronica Mandelli[1,2], Eleonora Satta[1], Stavros Trakoshis[1,6], Bonnie Auyeung[7,8], Prantik Kundu[9], Eva Loth[10,11], Guillaume Dumas[12], Sarah Baumeister[13], Christian F. Beckmann[14], Sven Bölte[15,16,17], Thomas Bourgeron[12], Tony Charman[18], Sarah Durston[19], Christine Ecker[20], Rosemary Holt[8], Mark H. Johnson[21], Emily J. H. Jones[22], Luke Mason[22], Andreas Meyer-Lindenberg[23], Carolin Moessnang[23], Marianne Oldehinkel[14,24], Antonio Persico[25], Julian Tillmann[18,26], Steven C. R. Williams[27], Will Spooren[28], Declan G. M. Murphy[10,11], Jan K. Buitelaar[14], the EU-AIMS LEAP group, Simon Baron-Cohen[8], Meng-Chuan Lai[8,29,30,31,32], & Michael V. Lombardo[1,8]*

Affiliations:

1 Laboratory for Autism and Neurodevelopmental Disorders, Center for Neuroscience and Cognitive Systems @UniTn, Istituto Italiano di Tecnologia, Corso Bettini, 31, 38068 Rovereto (TN), Italy
2 Center for Mind/Brain Sciences, University of Trento, 38068 Rovereto (TN), Italy
3 Department of Psychiatry, University of Cambridge, United Kingdom
4 Department of Child and Adolescent Psychiatry and Behavioral Science, Children's Hospital of Philadelphia, Philadelphia PA USA
5 Department of Psychiatry, University of Pennsylvania, Philadelphia PA USA
6 Department of Psychology, University of Cyprus, Cyprus
7 Department of Psychology, School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, United Kingdom
8 Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
9 Brain Mapping Unit, Department of Psychiatry, University of Cambridge, United Kingdom
10 Sackler Institute for Translational Neurodevelopment, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK
11 Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK
12 Human Genetics and Cognitive Functions, Institut Pasteur, UMR3571 CNRS, Université de Paris, Paris (75015), France
13 Child and Adolescent Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, J5, 68159 Mannheim, Germany
14 Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands
15 Center of Neurodevelopmental Disorders (KIND), Centre for Psychiatry Research; Department of Women's and Children's Health, Karolinska Institutet & Stockholm Health Care Services, Region Stockholm, Stockholm, Gävlegatan 22, BUP-FOU, 11330 Stockholm, Sweden

16 Child and Adolescent Psychiatry, Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden
17 Curtin Autism Research Group, School of Occupational Therapy, Social Work and Speech Pathology, Curtin University, Perth, Western Australia
18 Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK
19 Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands
20 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt am Main, Goethe University, Deutschordenstrasse 50, 60528 Frankfurt, Germany
21 Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, Cambridge, UK
22 Centre for Brain and Cognitive Development, Birkbeck, University of London, Henry Wellcome Building, Malet Street, London WC1E 7HX, UK
23 Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, 68159 Mannheim, Germany
24 Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Victoria, Australia
25 Child and Adolescent Neuropsychiatry Unit, Gaetano Martino University Hospital, University of Messina, Via Consolare Valeria 1, I-98125 Messina, Italy
26 Department of Applied Psychology: Health, Development, Enhancement, and Intervention, University of Vienna, Vienna, Austria
27 Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK
28 Roche Pharma Research and Early Development, Neuroscience, Ophthalmology and Rare Diseases, Roche Innovation Center Basel, Grenzacherstrasse 124, B.001 N.667, CH-4070 Basel, Switzerland
29 The Margaret and Wallace McCain Centre for Child, Youth & Family Mental Health, Azrieli Adult Neurodevelopmental Centre, and Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada
30 Department of Psychiatry, The Hospital for Sick Children, Toronto, Canada
31 Department of Psychiatry, Faculty of Medicine, University of Toronto, Toronto, Canada
32 Department of Psychiatry, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan

Corresponding Author: Michael V. Lombardo (michael.lombardo@iit.it)

Abstract

Social-communication (SC) and restricted repetitive behaviors (RRB) are autism diagnostic symptom domains. SC and RRB severity can markedly differ within and between individuals and is underpinned by different neural circuitry and genetic mechanisms. Modeling SC-RRB balance could help identify how neural circuitry and genetic mechanisms map onto such phenotypic heterogeneity. Here we developed a phenotypic stratification model that makes highly accurate (96-98%) out-of-sample SC=RRB, SC>RRB, and RRB>SC subtype predictions. Applying this model to resting state fMRI data from the EU-AIMS LEAP dataset (n=509), we find replicable somatomotor-perisylvian hypoconnectivity in the SC>RRB subtype versus a typically-developing (TD) comparison group. In contrast, replicable motor-anterior salience hyperconnectivity is apparent in the SC=RRB subtype versus TD. Autism-associated genes affecting astrocytes, excitatory, and inhibitory neurons are highly expressed specifically within SC>RRB hypoconnected networks, but not SC=RRB hyperconnected networks. SC-RRB balance subtypes may indicate different paths individuals take from genome, neural circuitry, to the clinical phenotype.

Autism spectrum disorder (ASD) is a clinical consensus label used to characterize individuals with a collection of early onset developmental difficulties in the domains of social-communicative (SC) and restricted repetitive behaviors (RRB)[1,2]. The single diagnostic label of autism helps many individuals in a variety of ways by being incorporated into a sense of identity, explaining challenging aspects of life, and/or enabling access to services. However, the diagnosis also encapsulates a vast amount of multi-scale heterogeneity. In the face of such heterogeneity, future translational research must develop a deeper understanding of how biological mechanisms affect individuals and must develop more personalized approaches towards interventions to help facilitate positive outcomes[3]. An example of this multi-scale heterogeneity can be seen at the level of the phenotypic symptom dyad of SC and RRB domains. Prior work has suggested that SC and RRB domains are fractionable at behavioral[4] and neural levels[5–7] and are underpinned by different genetic mechanisms[8–11]. The multi-scale fractionation of these domains provides a strong starting point for understanding how multi-scale heterogeneity manifests in autism from genome to phenome. Examining the relative balance between severity of difficulties in SC and RRB within an individual (SC-RRB balance) could help highlight unique biological underpinnings affecting different types of autistic individuals.

In this work, we test the hypothesis that subtyping individuals by the degree of SC-RRB balance can help identify differential biological mechanisms. Past research utilizing 'gold standard' diagnostic tools such as the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) (e.g., [12–14]) have suggested the presence of 3 SC-RRB balance subtypes: 1) medium to high levels of both SC and RRB severity (SC=RRB); 2) medium to high SC severity and comparatively lower RRB severity (SC>RRB); and 3) medium to high RRB severity and comparatively lower SC severity (RRB>SC). These subtypes might be underpinned by a common pathway if they showed similar neural circuit and genomic mechanisms that differ from a typically-developing (TD) comparison group. However, based on the idea that SC and RRB domains are fractionable across multiple levels, it could be that these subtypes diverge onto multiple pathways from genome to phenome[15] (Figure 1). This idea has not yet been tested with respect to macroscale neural circuitry and its link to functional genomic mechanisms. Here we evaluate how SC-RRB balance subtypes link up to differential macroscale connectome phenotypes, measured with resting state fMRI (rsfMRI) functional connectivity. Functional connectivity networks are known to be linked to underlying transcriptomic mechanisms, particularly with regards to the spatial patterning of gene expression across the brain (e.g., [16–18]). Given that subtypes could exhibit different functional connectome phenotypes, we leverage the link between macroscale rsfMRI functional networks and transcriptomic mechanisms to better understand whether autism-relevant functional genomic mechanisms differentially affect such phenotypes.

To test these ideas, we first utilize ADI-R data from thousands of individuals (n=2,628) within the National Database for Autism Research (NDAR) (https://nda.nih.gov). We develop a supervised subtyping approach that can estimate relative balance of SC versus RRB within an individual and make accurate out-of-sample subtype predictions. We then applied this supervised stratification approach to the deeply phenotyped EU-AIMS LEAP dataset[19–21] to examine how SC-RRB balance subtypes may replicably differ from typically-developing control (TD) groups in macroscale rsfMRI connectome phenotypes. Finally, in order to link functional connectome phenotypes to autism-associated genes, we utilize the Allen Institute Human Brain Atlas[22,23] to

identify genes whose spatial expression pattern is highly similar to macroscale functional networks that differ between autism subtypes and controls. These functional network-relevant gene lists are then investigated for enrichment in a variety of autism-associated gene lists derived from evidence at genetic or transcriptomic levels. This will allow for tests of the hypothesis that subtype disruption of imaging-derived phenotypes preferentially occurs to macroscale networks with high levels of gene expression of autism-associated genes[24]. This approach will also allow us to test whether autism-associated genes affect networks similarly or differently across the SC-RRB balance subtypes.

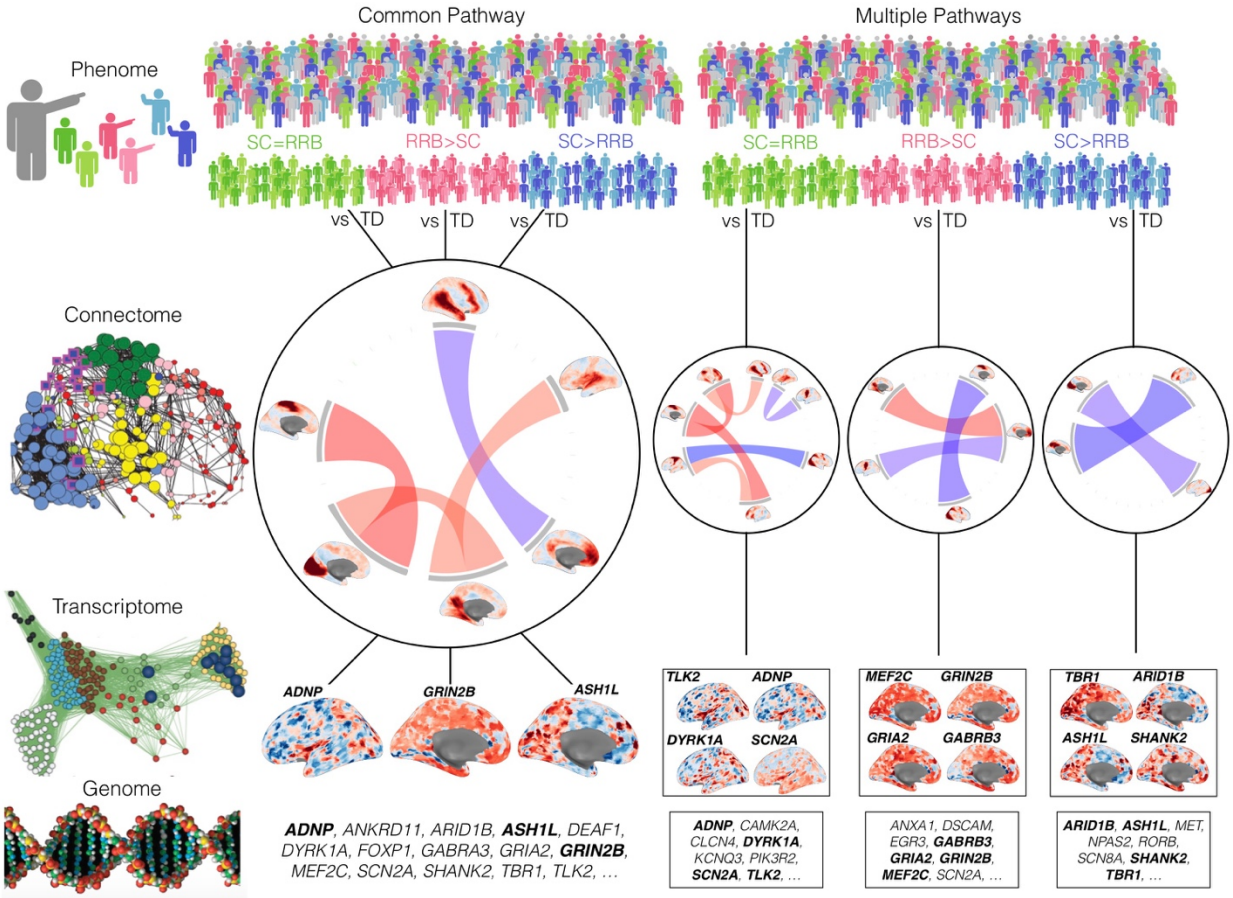


Figure 1: Approach towards testing common pathway versus multiple pathways explanations behind SC-RRB balance in autism. *In this figure we depict two alternatives for how SC-RRB balance subtypes (phenome level; SC=RRB, green; RRB>SC, pink; SC>RRB, blue) could be explained at the level of macroscale functional connectome phenotypes measured with rsfMRI (connectome level) and autism-associated functional genomic mechanisms (e.g., transcriptome and genome levels). Columns in this figure depict the common pathway (middle) and multiple pathways (right) models. The common pathway model predicts that when each subtype is compared to a typically-developing (TD) comparison group, they converge and share a common difference from TD in affected macroscale rsfMRI functional connectome phenotype. Underpinning this shared connectome phenotype are a myriad of differing functional genomic mechanisms. At the level of the transcriptome, we identify genes linked to macroscale functional networks by identifying genes whose spatial expression pattern across the brain is similar to the*

spatial topology of the macroscale functional network. This procedure generates a list of genes relevant for such macroscale networks and these lists are then tested for enrichment in autism-association functional genomic mechanisms. The gene list at the genome level represents an example of possible autism-associated genes that may (bold) or may not (non-bold) be linked to macroscale functional networks. In contrast to the common pathway model, the multiple pathways model would highlight that differential connectome phenotypes when compared to TD are unique to each subtype, and that each of these subtype-specific connectome phenotypes is underpinned by a differing set of autism-associated functional genomic mechanisms.

Results

Highly accurate supervised prediction of SC-RRB balance subtypes

In our first set of analyses, we sought to develop a supervised model to predict ADI-R SC-RRB balance subtypes from the NDAR datasets. Relatively equal Discovery (n=889) and Replication (n=890) datasets were partitioned from the total n=1,781 individuals in NDAR and this split into Discovery and Replication was balanced as a function of the originating datasets and sex. Using z-normalized difference scores, we cut individuals into SC=RRB, SC>RRB, and RRB>SC subtypes (Figure 2). Importantly, the subtype labels were defined separately in Discovery and Replication sets from the norms (mean and SD) estimated on each set. This ensures that the definition of the labels in each set is done independently. Irrespective of the z-threshold used for labeling the subtypes (e.g., z = 0.5 up to z =1 in steps of 0.1), we find that a multiclass SVM classifier trained on the NDAR Discovery set and tested on the NDAR Replication set is highly accurate in the range around 96-98.6% (p = 9.99e-5). Contrasting this z-score approach to subtyping with unsupervised clustering methods, we found that such SC-RRB balance subtypes are not easily identifiable in a consistent fashion across Discovery and Replication cohorts with such blind methods (see Supplementary Figures 2-3). Examination of sex across these subtypes did not yield any significant between subtype differences (Discovery: χ^2 = 1.91; p = 0.38; Replication: χ^2 = 3.50; p = 0.17), with a 3:5 to 5:1 sex ratio of males to females. Subtypes did differ in age at the time of ADI-R interview, with the SC>RRB group being younger than the other subtypes (Discovery: $F(2,886)$ = 10.61, p = 2.77e-5; Replication: $F(2,887)$ = 10.80, p = 2.31e-5). See Table 1 for descriptive statistics.

Subtype differences in adaptive behavior

Applying our subtyping method to the deeply phenotyped EU-AIMS LEAP dataset, we had the opportunity to investigate whether there are other phenotypic differences between the subtypes. No replicable differences were apparent between SC=RRB and SC>RRB across the ADOS-2, SRS-2, RBS-R, and SSP at any z-threshold. However, at a z-threshold of 1 we found replicable reductions in SC>RRB compared to SC=RRB on the Vineland Adaptive Behavior Scales-Second Edition (VABS-II; henceforth VABS) Daily Living Skills subscale and the overall Adaptive Behavior Composite (ABC) score (see Table 2 and Figure 2D-E). While these effects appear only at z=1, it is apparent that as the z-threshold for defining subtypes increases, the effect size for a subtype difference also increases (Supplementary Figure 4). This could imply that continuous variation on the z-normalized SC-RRB difference score would be associated with

Vineland scores. To test this, we ran a dimensional model but could not find statistically significant effects on both Discovery and Replication sets (Supplementary Figure 4).

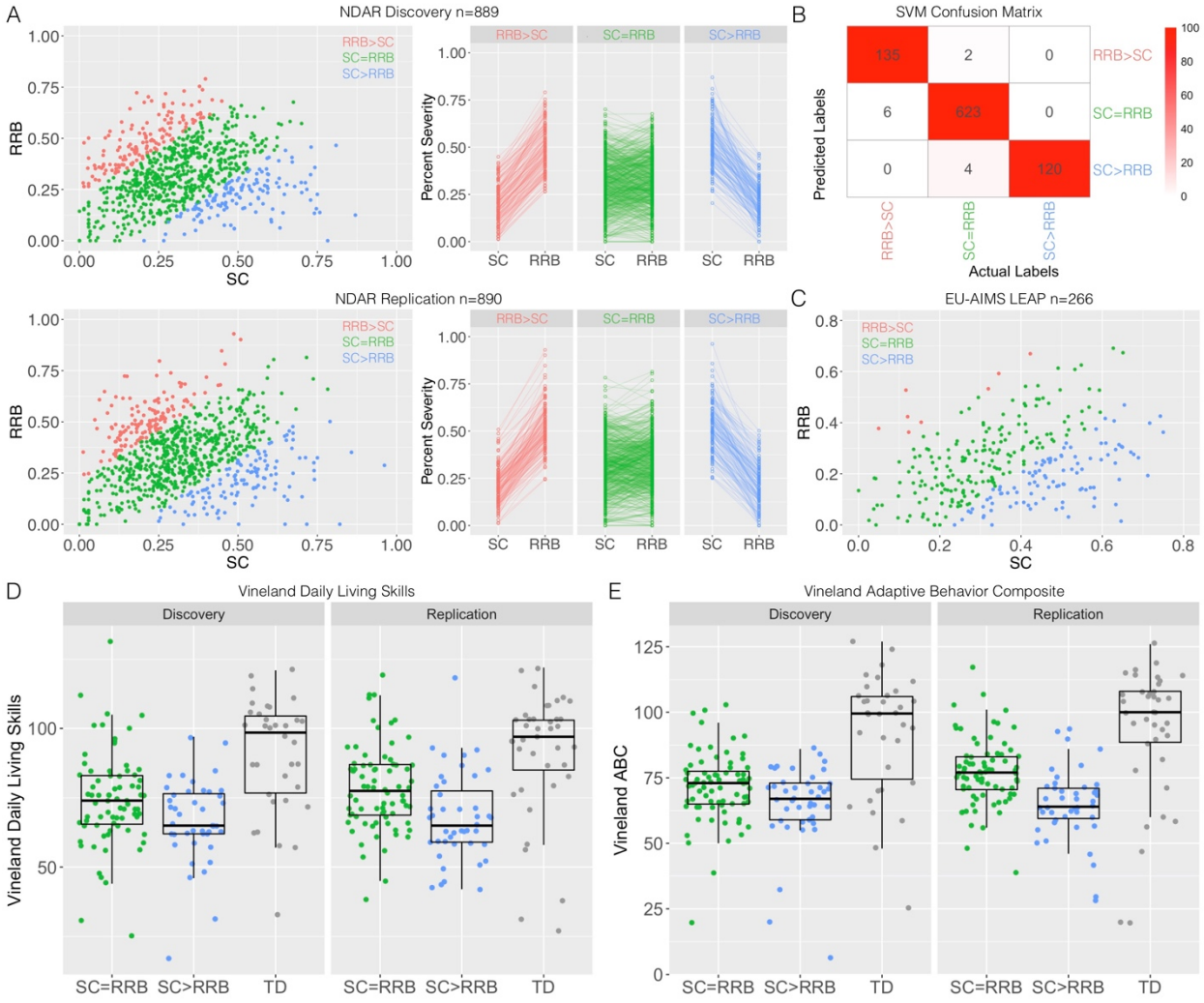


Figure 2: Supervised subtyping of autism by SC-RRB balance. *Panel A shows the subtypes derived from a z-normalized difference score of SC-RRB, with a z-score threshold for cutting the subtypes at z = 1. Red shows the RRB>SC subtype, green shows the SC=RRB subtype, and blue shows the SC>RRB subtype. Panel B shows a confusion matrix with actual subtype labels for the NDAR Replication dataset along with columns and SVM predicted labels along the rows. The colors within the cells indicate the percentage of individuals relative to the actual labels with predicted labels in each cell. The high percentages are indicative of high classification accuracy (98.6% accuracy). Over a range of z-thresholds from 0.5 to 1, the accuracy ranged from 96-98.6% accuracy. Panel D shows boxplots and data points for each individual on the Vineland Daily Living Skills subscale for SC=RRB (green) and SC>RRB (green) subtypes along with the TD group (gray). Panel E shows the same plot as panel D for the Vineland Adaptive Behavior Composite scores.*

To demonstrate how big of a difference the SC>RRB vs SC=RRB effect is with reference to the average autistic individual in the population, we calculated standardized effect sizes (*Cohen's d*) using Vineland mean and standard deviation norms from Chatham et al.,[25], which are computed on over 9,000 autistic individuals. The SC>RRB subtype shows reductions on both the VABS Daily Living Skills subscale and ABC of around half a standard deviation (Daily Living Skills, Discovery: *Cohen's d* = -0.56; Replication: *Cohen's d* = -0.50; ABC, Discovery: *Cohen's d* = -0.52; Replication: *Cohen's d* = -0.53). Thus, not only is this a statistically significant and replicable effect, it is also well above the minimally clinically-important difference suggested by Chatham and colleagues[25]. These results indicate that the SC>RRB subtype may also be a clinically significant subset of individuals with more pronounced difficulty in adaptive behaviors.

Replicable subtype-specific functional connectivity differences

We next evaluated whether such SC-RRB balance subtypes are differentiated from TD comparison groups in rsfMRI functional connectivity. Because subtypes are defined based on thresholding the z-normalized SC-RRB difference score, we identified 'consensus edges' as functional connectivity differences between the autism subtype versus TD that consistently appear across every z-threshold examined. Figure 3 summarizes the consensus edges in each subtype for both the LEAP Discovery and Replication sets. The SC=RRB subtype is characterized by on-average hyperconnectivity between the anterior salience network (IC07) and a medial motor network (IC13) (effect sizes at z=1 threshold: Discovery *Cohen's d* = 0.36; Replication *Cohen's d* = 0.53). In contrast, the SC>RRB subtype is characterized by on-average hypoconnectivity between a bilateral perisylvian temporal network (IC17) and a bilateral somatomotor network (IC12) (effect sizes at z=1 threshold: Discovery *Cohen's d* = -0.38; Replication *Cohen's d* = -0.37). Across each threshold we also counted the number of times that replicably different edges were common across the subtypes. Strikingly, edges were never shared in common across the subtypes, further indicating that when replicable functional connectivity differences appear, they are specific to one subtype and not the other. In contrast to comparing autism subtypes to TD, we also directly compared SC=RRB versus SC>RRB. This analysis did not yield any significant replicable differences, indicating that while these subtypes can differ from a TD comparison group in unique ways, the difference between each other may not be substantially large (effect sizes for z=1 threshold: IC07-IC13, Discovery *Cohen's d* = 0.12, Replication *Cohen's d* = 0.08; IC12-IC17, Discovery *Cohen's d* = 0.23, Replication *Cohen's d* = 0.05). For the full set of statistical results at z=1 threshold across all comparisons see Supplementary Table 2.

Because the subtyping approach uses the difference score between SC and RRB, this metric does not distinguish individuals by overall level of severity. For example, an individual with low SC and RRB severity is treated similarly to an individual with high SC and RRB severity. This leaves open the possibility that degree of severity on a continuum from high to low could also explain variability in functional connectivity. To test this hypothesis, we constructed a dimensional model to predict connectivity strength from SC or RRB severity as a continuous variable. However, there were no instances whereby SC or RRB severity as a continuous measure could replicably predict connectivity strength. Similarly, when using the z-normalized SC-RRB difference score as a continuous variable, we also found no replicable significant effects on connectivity. For the full set of statistics see Supplementary Table 2. These results provide a dimensional model contrast to the categorical subtyping approach and suggests that modeling

continuous SC or RRB variability may be less sensitive as a predictor of functional connectivity compared to SC-RRB balance subtypes.

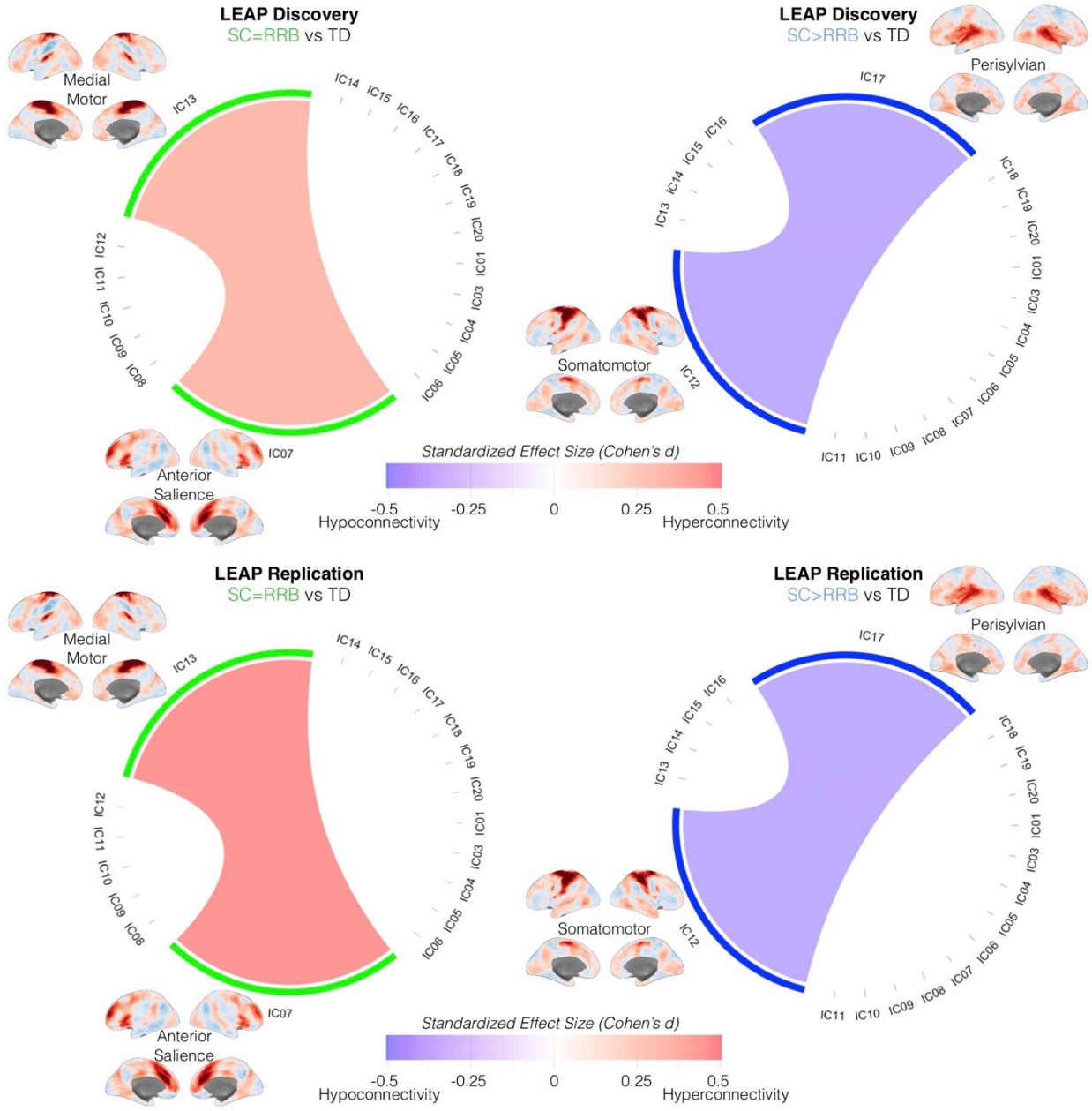


Figure 3: Replicable subtype differences in functional connectivity. *This figure shows chord diagrams of replicable functional connectivity differences between SC=RRB vs TD (left) or SC>RRB vs TD (right), when subtypes are defined at a z-threshold of 1. Edges shown in these diagrams are consensus edges that appear in every analysis of connectivity differences irrespective of the z-threshold used to define the subtypes. Network nodes in the diagram highlighted in green or blue are networks that appear in only one of the autism subtypes. None of the edges in the diagram overlap or share directionality of difference (red edges indicate hyperconnectivity in autism while blue edges indicate hypoconnectivity in autism). Intensity of*

edge color indicates standardized effect size (Cohen's d). The cortical surface renderings of each component are unthresholded z-stat maps. Areas with higher z-stats (dark red) are of primary importance for the IC map. The top rows show effects in the EU-AIMS LEAP Discovery set, while bottom rows show effects in the EU-AIMS LEAP Replication set.

Divergent functional genomic underpinnings of subtype-specific neural circuitry

In the next analysis, we asked if known autism-associated genes are enriched amongst genes known to be highly expressed in these subtype-associated rsfMRI networks. We first identified lists of genes whose spatial expression topology in the Allen Institute Human Brain Atlas[22] is similar to rsfMRI connectivity networks[23] that show replicable subtype differences. Once a set of genes are predicted to underpin such rsfMRI networks, we then asked whether those genes are highly overlapping with known sets of functional genomic mechanisms linked to autism (see Figure 4A for a visual representation of the analysis approach and Supplementary Table 3 for the full set of gene lists used in these analyses). Bilateral somatomotor (IC12) and perisylvian networks (IC17) linked to SC>RRB hypoconnectivity were enriched for gene lists such as highly penetrant de novo protein truncating variants associated with autism, a more broad list of genes associated with autism from the SFARI Gene database, as well as cortically downregulated genes[26] and gene co-expression modules[27], and differentially expressed genes that specifically affect excitatory and inhibitory neurons and astrocytes[28]. We next assessed enrichment for genes that are unique to IC12 and IC17 but not IC07 or IC13, since this gene list is most specific to mechanisms within SC>RRB. The same effects remained significant under this more stringent enrichment test (see the SC>RRB column in Figure 4B). In contrast to these enrichments for SC>RRB, we also examined IC07 and IC13 networks. No enrichments were significant for IC07. However, IC13 was enriched for cortically downregulated genes[26] and co-expression modules[27]. Only the enrichment with cortically downregulated genes remained when selecting genes unique to only IC07 and IC13 but not IC12 or IC17 (see the SC=RRB column in Figure 4B). As an important contrast to autism-associated genes, we also tested for enrichment with genes differentially expressed in cortical tissue in schizophrenia (SCZ) and bipolar disorder (BD), since autism is known to be somewhat genetically correlated with these disorders[26,29]. However, we could not find any enrichments with these SCZ and BD differentially expressed gene sets. This indicates that the signatures we identify here are somewhat specific to autism and not more generalized to other genetically correlated psychiatric conditions.

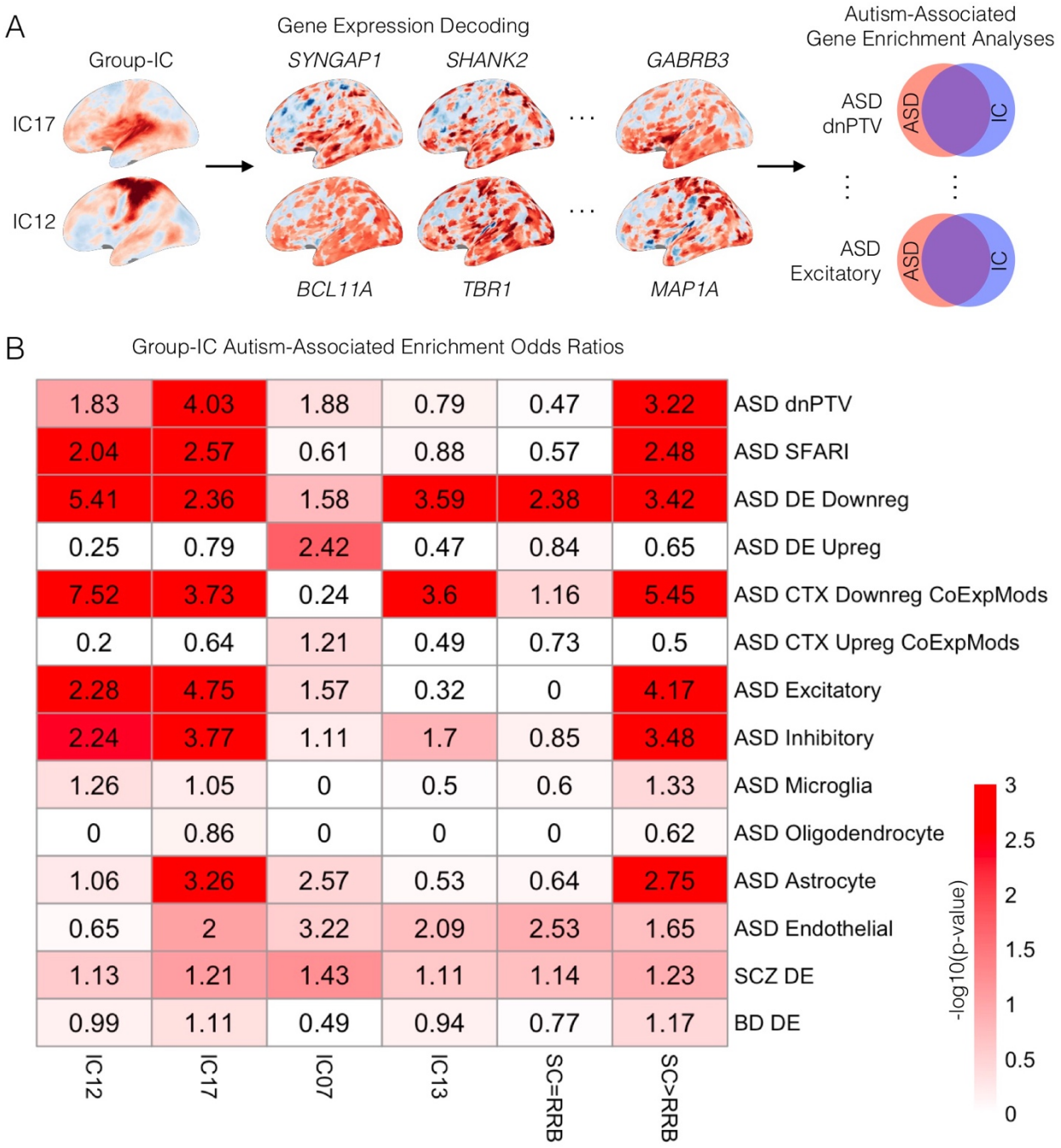


Figure 4: Overlap between genes expressed in functional connectivity networks and genes linked to autism. *In panel A we depict the analysis approach of identifying genes which are highly expressed in similar spatial patterns to the rsfMRI spatial IC maps (i.e. gene expression decoding). Once IC gene lists have been identified, we test these lists for enrichment with known lists of autism-associated functional genomic mechanisms (top left). In panel B we show enrichment odds ratios (numbers in each cell) along with the -log10 p-value (coloring of the cells) for enrichment tests of specific networks (columns) against known lists of autism-associated genomic mechanisms (rows). Cells with darkest red color survive FDR q<0.05. IC12 and IC17 are networks with SC>RRB hypoconnectivity compared to TD. IC07 and IC13 are networks with SC=RRB*

hyperconnectivity compared to TD. The column labeled SC>RRB shows the enrichment results when the gene list under consideration are comprises genes unique to IC12 and IC17, but not IC07 or IC13. The column labeled SC=RRB shows the enrichment results when the gene list under consideration consists of genes unique to IC07 and IC13, but not IC12 or IC17. ASD dnPTV, Autism de novo protein truncating variants; ASD SFARI, SFARI Gene autism associated genes; ASD DE Downreg, autism differentially expressed downregulated genes; ASD DE Upreg, autism differentially expressed upregulated genes; ASD CTX Downreg CoExpMods, autism downregulated cortical co-expression modules; ASD CTX Upreg CoExpMods, autism upregulated cortical co-expression modules; ASD Excitatory, autism differentially expressed genes in excitatory neurons; ASD Inhibitory, autism differentially expressed genes in inhibitory neurons; ASD Microglia, autism differentially expressed genes in microglia; ASD Oligodendrocyte, autism differentially expressed genes in oligodendrocytes; ASD Astrocyte, autism differentially expressed genes in astrocytes; ASD Endothelial, autism differentially expressed genes in endothelial cells; SCZ DE, schizophrenia differentially expressed genes; BD DE, bipolar disorder differentially expressed genes.

Discussion

In this work, we examined how autism SC-RRB balance subtypes are similar or different at the level of macroscale neural circuitry measured with rsfMRI. Prior work has suggested that the core dyad of SC and RRB is fractionable at behavior and neural levels and is underpinned by different genetic mechanisms[4–14]. However, it is unclear whether the road from genome to phenome (e.g., Figure 1) is one that converges on a common pathway or is one of multiple pathways[15]. Here we find evidence more in line with the multiple pathways model. All functional connectome edges (e.g., connections between IC networks) and nodes (e.g., IC networks) that show a consensus for replicable differences compared to TD across z-thresholds were those that were specific to only one subtype.

Autistic individuals within the SC>RRB subtype show on-average hypoconnectivity compared to a matched TD comparison group in cortical circuitry mainly consisting of bilateral somatomotor and perisylvian temporal networks. The bilateral somatomotor network has been implicated in past work on autism. Somatosensory areas have been shown to be some of the most informative regions in prior case-control classifier studies using rsfMRI data[30]. Additionally, prior case-control analyses of the EU-AIMS LEAP dataset find that somatomotor areas show reduced degree centrality and autism-related hyperconnectivity with cerebellar networks[21,31]. However, the lack of identification of hypoconnectivity between somatomotor and perisylvian networks in prior case-control analyses highlights the potential importance and added value of subtyping for revealing more subtle effects that can be masked with case-control contrasts. The identified perisylvian network overlaps with a variety of areas implicated in early development of autism, particularly for auditory processing and language[32–36]. Integration of information processing between this network and others that play a role in embodied somatosensory and social cognitive processing[37,38], such as the somatomotor network (IC12), could be important for explaining the more pronounced difficulties these individuals have within SC compared to RRB. This effect could also point to atypical multisensory integration that has been documented in autism[39,40], particularly with regards to auditory-somatosensory integration[41].

In contrast, the autism SC=RRB subtype was characterized by on-average hyperconnectivity between medial motor and anterior salience networks. This effect could be important given that a subset of individuals with autism show marked motor difficulties[42–44] and because of evidence showing that visual-motor integration is atypical in autism[45–49]. The anterior salience network has also been identified in prior case-control studies. In younger cohorts, anterior salience areas are hyperconnected[50,51], while in older cohorts, hypoconnectivity is observed[52]. While age could be a factor in explaining the discrepant findings from prior work, it likely cannot explain the SC=RRB hyperconnectivity finding. Here we age-matched the groups and additionally included age as a covariate in the statistical model. EU-AIMS LEAP also samples from a wide age range from 6 to 30 years of age, enabling the sample to cover younger and older ages covered by prior work. Thus, age may not be the only explanation for salience network hyperconnectivity. Rather, this work suggests that SC-RRB heterogeneity and the presence of this balanced subtype could also drive such effects in case-control comparisons, particularly if the sample is enriched with this particular subtype.

We also identified autism-relevant genomic underpinnings behind these subtype-specific rsfMRI networks. Genes specific to SC>RRB networks are enriched for a number of genomic mechanisms linked to autism such as genes differentially expressed in excitatory and inhibitory neurons and astrocytes, downregulated co-expression modules, and high-risk genetic mutations associated with autism. These genomic underpinnings suggest that specific neuronal cell types involved in cortical excitation-inhibition balance[53,54] may be especially important for the SC>RRB subtype. This effect also partially corroborates evidence suggesting that excitatory neurons are affected in specific types of autistic individuals that differ in patterns of clinical severity[28]. In contrast, SC=RRB networks lacked similar kinds of enrichments, suggesting that differing functional genomic mechanisms may be linked to this subtype.

This work shows that phenotypically-derived subtypes fractionated based on SC-RRB balance are robust and consistent across the population. While prior work has shown evidence for these subtypes[12,14] derived from complex statistical models, it has been less clear as to how to objectively identify them with simpler models and in new datasets for a priori investigation. Our approach here provides a straightforward solution to this problem by deriving simple autism population norms (e.g., mean and standard deviation) for the difference between SC and RRB severity. These norms are relatively consistent across independent large datasets from NDAR, and this allows for a supervised classification approach to identify multiple subtypes with high levels of accuracy. A considerable strength of this work is that we built the classifier based on large datasets from the NDAR repository that allows for partitioning of datasets into Discovery and Replication sets while retaining sample size on the scale of many hundreds of individuals. The classifier also works well irrespective of the threshold used to define the subtypes. Future work on fractionable phenotypic subtypes could benefit from utilizing our classifier as an a-priori way to identify subtypes. To utilize our approach all that is needed are ADI-R DSM-5-based algorithm items which can be summarized into SC and RRB total scores. By computing z-scores with the population norms we have derived from NDAR, this will allow for designing studies a-priori with these subtypes in mind. As we saw from the EU-AIMS LEAP dataset, it is not always assured that smaller, newly ascertained samples of data will have equally distributed sample sizes across the

subtypes, and thus, the ability to a-priori identify these individuals will significantly help in prospective study designs.

Finally, it is important to underscore that at higher thresholds for defining these SC-RRB subtypes, there are also more pronounced difficulties with adaptive behaviors in the SC>RRB subtype. This effect is consistent with prior work showing a similar reduction in adaptive behavior on the Vineland in an SC>RRB subtype[12], and with prior work showing that more severe early childhood social-communicative severity on the ADI-R is predictive of worse trajectories for adaptive behaviors[55]. This effect is also in line with inferences from a recent paper on the EU-AIMS LEAP dataset showing that higher social impairment, via increasing SRS-2 scores, was associated with lower scores on the Vineland[56]. These findings suggest that the SC>RRB subtype could be a relevant stratifier for clinical trials that are focused on change in real-world outcomes.

There are certain limitations and caveats that need to be discussed. First, the threshold for the z-score cutoff to define subtypes could be viewed as arbitrary. However, to guard against this issue, we re-ran the analysis across a range of thresholds from z=0.5 to z=1 and showed effects that are robust to the threshold used to label the subtypes. Classification accuracy is also high regardless of the threshold. This effect occurs largely because the statistics used for the z-normalization are highly similar across large NDAR Discovery and Replication sets. Thus, while the choice of a threshold may not be well defined, any choice within the range we have analyzed of z=0.5 to z=1, will yield highly consistent results that are not biased by the choice of a threshold. Furthermore, it could be argued that if one's aim was to stratify to obtain clinically meaningful groups that differ on adaptive behavior, the results here would suggest that utilizing higher z-thresholds would be pertinent. Second, the distinctions between these subtypes are not demarcated by large categorical separations. As such, when we applied other unsupervised clustering methods to the data, such methods were not able to consistently identify the same subtypes in independent datasets (Supplementary Figures 2-3). The lack of consistently identifiable subtypes with other complex unsupervised methods like clustering indicate that a simpler approach may be necessary and that large distinctions between the boundaries for different types of patients are not obvious and would thus necessitate a more nuanced and theory-driven approach. Third, direct comparisons of functional connectivity between SC=RRB and SC>RRB did not yield large differences. Thus, while there are unique consensus edges that appear when the autism subtypes are compared to TD, this result should not be taken to imply that the subtypes themselves are also highly different from each other. A likely reason for why these differences manifest when compared to TD but not when subtypes are compared directly may likely be due to effects driven by further subsets of individuals nested within the larger SC=RRB and SC>RRB subtypes. These individuals at the extremes of the functional connectivity distributions likely drive the on-average differences from TD. Future work that digs further into more granular divisions of the population may likely identify much larger differences when autism subtypes are compared directly. Fourth, we also discovered that dimensional models using continuous SC and RRB severity did not uncover any replicable associations with functional connectivity strength. This result could suggest that dimensional models that use continuous severity from the ADI-R are less effective than the subtyping approach. However, it could also be that dimensional models might be more sensitive with other measures of symptomatology (e.g., ADOS, SRS). Fifth, the subtyping here is based on the ADI-R. ADI-R is a commonly used gold standard diagnostic instrument used to help aid clinical judgment regarding diagnosis. However, other gold standard measures such as the ADOS could also have

been used. For our purposes in this study, we chose to utilize the ADI-R over the ADOS due to the fact that participants come from a wide age range, and the ADOS would assess current symptomatology of the individual. If age has an effect on symptomatology[57–59], then this could potentially bias the subtyping approach depending on the composition of the sample. On the other hand, because the ADI-R is a snapshot of early developmental symptomatology, we cannot know how the individual might have changed from that time point to the current assessment. Additionally, it may be that current measures of symptomatology have a stronger association with current measures of functional connectivity than early childhood snapshots of severity provided by the ADI-R. Future work that looks at how these ADI-R subtypes might change over time would be important for future research. It would also be important to investigate how observational measures such as the ADOS might perform as measures of symptomatology, especially if conducted within a restricted age range in order to guard against biases due to large age ranges.

In conclusion, we have shown that SC-RRB balance can point to different macroscale functional connectivity phenotypes and potentially different genomic mechanisms that may underpin such phenotypes. While the divisions between these subtypes at the phenotypic level are not dramatically evident as categorical differences, at the level of macroscale neural circuitry, there is evidence to suggest that these SC-RRB subtypes are different when compared to the TD population. Future work that begins to study these fractionable subtypes in an a-priori fashion will benefit from the use of our simple and supervised subtyping model and will further facilitate our understanding of how heterogeneity in autism manifests in a multi-scale fashion from genome to phenome.

Methods

NDAR Datasets

SC-RRB balance subtyping analyses were conducted on ADI-R data available through the National Database for Autism Research (NDAR). The ADI-R was chosen as the symptomatology measure for subtyping as it assesses early childhood symptoms around 4-5 years of age rather than current symptom severity. The rationale for assessing early childhood symptoms over current symptoms was to be able to sample/estimate symptomatology at similar developmental time points in life across individuals, rather than have those estimates vary depending on the age of the individual. Different experiences over the lifespan could potentially affect symptom presentation and to guard against this issue possibly biasing the subtyping, we opted for the ADI-R.

On December 13, 2019 we conducted a search of NDAR to extract all datasets utilizing the ADI-R[60]. This resulted in 60 independent datasets totaling 2,628 unique individuals. From here, we filtered for all individuals who had data for the verbal items (e.g., acquisition of words, phrases, social verbalization, chit-chat, reciprocal conversation) leaving a total of 1,781 individuals across 57 independent datasets. Within each of these 57 datasets, we randomly split the dataset in half to achieve independent Discovery and Replication sets that are balanced across the 57 datasets and by sex (Discovery n=889, mean age = 8.91 years, SD age = 5.26 years, 77% male; Replication n=890, mean age = 8.89 years, SD age = 5.37 years, 77% male). See Table 1 for characteristics of the NDAR Discovery and Replication sets. ADI-R item-level data was used to rescore algorithm

totals according to the DSM-5[61] symptom dyad of social-communication (SC) and restricted repetitive behavior (RRB) domains. SC is comprised of 3 subscales (A1, A2, A3), while RRB is comprised of 4 subscales (B1, B2, B3, and B4). See Supplementary Table 1 for how items break down into each domain and subscale within a domain. Only item scores of 0 to 3 (indicating increasing SC or RRB symptom severity) were utilized, while scores of 6 to 9 (dummy scores, not indicating symptom severity) were not used. Scores of 3 were kept as is (i.e. not converted to 2 as would typically occur when scoring the ADI-R algorithm) in order to retain information about severity conveyed by the difference between a score of 2 versus 3. Because the number of items in each subscale can vary depending on a person's age (see Supplementary Table 1) and by the number of items with possible scores of 0 to 3, we used percentage scores in order to ensure that the estimates of severity on each subscale are on a comparable scale across individuals. These percentage scores for each domain subscale were then summed and scaled by number of subscales to achieve the final domain total percentage severity.

Subtyping Analyses

To label subtypes by SC-RRB balance, we first computed difference scores between SC and RRB to estimate the level of SC-RRB balance, whereby values above 0 indicate an individual with higher SC versus RRB (SC>RRB), whereas values below 0 indicate the reverse (RRB>SC). These SC-RRB difference scores were then z-normalized using the mean and standard deviation estimated separately for Discovery and Replication sets. A z-score cutoff was used to derive subtype labels. Individuals falling above the z-cutoff (e.g., $z>1$) were labeled as SC>RRB, while individuals falling below the negative value of the z-cutoff (e.g., $z<-1$) were labeled as RRB>SC. All individuals between the z-cutoffs were considered SC=RRB. Because the choice of a z-cutoff is arbitrary, we ran all analyses across a range of z-thresholds from $z=0.5$ to $z=1$, in steps of 0.1. This approach allows us to report results across thresholds rather than using only one arbitrarily defined threshold. For the later functional connectivity analyses, this approach also allowed us to identify a consensus result which is consistent irrespective of the z-threshold used to label subtypes.

Our subtyping approach was embedded within a supervised classification approach in order to evaluate how well this subtyping procedure generalizes to independent datasets. To achieve this aim, we used a multiclass Support Vector Machine (SVM) classifier to train on the SC>RBB, SC=RRB, and RRB>SC subtype distinctions in the NDAR Discovery set. The multiclass SVM was trained using 2 features - percent severity SC and RRB totals. This multiclass SVM model was then tested on the NDAR Replication set, where the subtype labels were determined separately based on the norms estimated from the Replication set. Out of sample classification accuracy on the held-out test set (NDAR Replication) were computed and confusion matrices were built to show how predictions from SVM were made on the test set. We also evaluated how well the classifier performs when the subtype labels were randomly shuffled and this permutation procedure was repeated 10,000 times in order to generate a null distribution of accuracy values. From this null distribution a p-value was computed as the proportion of times an accuracy value was as high or higher than the actual accuracy value with unpermuted labels. To make subtype predictions in the EU-AIMS LEAP dataset, we combined both NDAR Discovery and Replication datasets into one large dataset. From this dataset, norms for the mean and standard deviation were computed (mean = 0.01045243, SD = 0.19482749) and used for the z-scoring procedure. SC-RRB

difference z-scores were then computed and a z-threshold was applied to the EU-AIMS LEAP dataset to generate subtype labels.

In addition to this SC-RRB difference z-score subtyping approach, we also used other unsupervised clustering methods for identifying subtypes. These methods utilize agglomerative hierarchical clustering with Euclidean distance and the ward.D2 method. The optimal number of clusters was determined by a majority vote of 23 different metrics for determining the optimal number of clusters (e.g., using the *NbClust* library in *R*)[62]. With another approach, we ran the same hierarchical clustering analyses, but cut dendrograms to define subtypes using a dynamic hybrid tree cut algorithm, as utilized in past work[63,64].

EU-AIMS LEAP Dataset

The EU-AIMS LEAP data comes from a large multisite European initiative with the aim of identifying biomarkers for ASD[20]. In this study, EU-AIMS LEAP data is utilized to examine how SC-RRB balance subtypes may differ in intrinsic functional connectomic organization using rsfMRI data. rsfMRI data from EU-AIMS LEAP has been analyzed for case-control differences in prior work[21,31]. EU-AIMS LEAP recruited 437 individuals with ASD and 300 TD individuals, both male and female, aged between 6 and 30 years. Participants underwent comprehensive clinical, cognitive, and MRI assessment at one of the following five centers: Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom; Autism Research Centre, University of Cambridge, United Kingdom; Radboud University Nijmegen Medical Centre, the Netherlands; University Medical Centre Utrecht, the Netherlands; and Central Institute of Mental Health, Mannheim, Germany. The study was approved by the local ethical committees of participating centers, and written informed consent was obtained from all participants or their legal guardians (for participants <18 years). For further details about the study design, we refer to Loth et al.,[19], and for a comprehensive clinical characterization of the LEAP cohort, we refer to Charman et al.,[20]. In the present study, we selected all participants for whom structural and rsfMRI data were available. However, n=120 participants had to be excluded from the analysis because of missing ADI-R item-level data (n=64), missing IQ data (n=3), or because preprocessing could not completed for a variety of reasons (e.g., registration/normalization errors because of poor quality MPRAGE data, poor anatomical coverage, or large anatomical deviance such as large ventricles (n=39), incomplete rsfMRI data (n=3), errors in convergence of the ME-ICA algorithm (n=11)). The final sample size was n=266 autistic and n=243 TD participants. This final sample was split into independent Discovery and Replication sets (balanced for sex and age) for the purpose of identifying functional connectivity differences that are replicable. As an example of sample sizes once split into autism subtypes at a z-threshold of 1, within the Discovery set there were n=80 SC=RRB, n=47 SC>RRB, and n=121 TD individuals. Within the Replication set there were n=85 SC=RRB, n=47 SC>RRB, and n=122 TD individuals. N=7 (n=6 Discovery, n=1 Replication) were classified as RRB>SC and because the sample sizes were too small, we did not analyze this subtype further for functional connectivity differences. We tested subtypes on a variety of different phenotypic measures including the ADOS-2, Social Responsiveness Scale (SRS-2), Repetitive Behavior Scale (RBS-R), Short Sensory Profile (SSP) and the Vineland Adaptive Behavior Scales (VABS). See Table 2 for participant characteristics.

EU-AIMS LEAP fMRI Data Acquisition

MRI data were acquired on 3T scanners: General Electric MR750 (GE Medical Systems, Milwaukee, WI, USA) at Institute of Psychiatry, Psychology and Neuroscience, King's College London, United Kingdom (KCL); Siemens Magnetom Skyra (Siemens, Erlangen, Germany) at Radboud University Nijmegen Medical Centre, the Netherlands (RUMC); Siemens Magnetom Verio (Siemens, Erlangen, Germany) at the University of Cambridge, United Kingdom (UCAM); Philips 3T Achieva (PhilipsHealthcare Systems, Best, The Netherlands) at University Medical Centre Utrecht, the Netherlands (UMCU); and Siemens Magnetom Trio (Siemens, Erlangen, Germany) at Central Institute of Mental Health, Mannheim, Germany (CIMH). Procedures were undertaken to optimize the MRI sequences for the best scanner-specific options, and phantoms and travelling heads were employed to assure standardization and quality assurance of the multi-site image-acquisition[20]. Structural images were obtained using a 5.5 minute MPRAGE sequence (TR=2300ms, TE=2.93ms, T1=900ms, voxels size=1.1x1.1x1.2mm, flip angle=9°, matrix size=256x256, FOV=270mm, 176 slices). An eight-to-ten minute resting-state fMRI (rsfMRI) scan was acquired using a multi-echo planar imaging (ME-EPI) sequence [65,66]; TR=2300ms, TE~12ms, 31ms, and 48ms (slight variations are present across centers), flip angle=80°, matrix size=64x64, in-plane resolution=3.8mm, FOV=240mm, 33 axial slices, slice thickness/gap=3.8mm/0.4mm, volumes=200 (UMCU), 215 (KCL, CIMH), or 265 (RUMC, UCAM). Participants were instructed to relax, with eyes open and fixate on a cross presented on the screen for the duration of the rsfMRI scan.

EU-AIMS LEAP fMRI Preprocessing

Multi-echo rsfMRI data were preprocessed with the multi-echo independent components analysis (ME-ICA) pipeline, implemented with the *meica* python library (v3.2) (https://github.com/ME-ICA/me-ica). ME-ICA implements both basic fMRI image preprocessing and decomposition-based denoising that is specifically tailored for multi-echo EPI data. For the processing of each subject, first the anatomical image was skull-stripped and then warped nonlinearly to the MNI anatomical template using AFNI 3dQWarp. The warp field was saved for later application to functional data. For each functional dataset, the first TE dataset was used to compute parameters of motion correction and anatomical-functional coregistration, and the first volume after equilibration was used as the base EPI image. Matrices for de-obliquing and six-parameter rigid body motion correction were computed. Then, 12-parameter affine anatomical-functional coregistration was computed using the local Pearson correlation (LPC) cost functional, using the gray matter segment of the EPI base image computed with AFNI 3dSeg as the LPC weight mask. Matrices for de-obliquing, motion correction, and anatomical-functional coregistration were combined with the standard space nonlinear warp field to create a single warp for functional data. The dataset of each TE was then slice-time corrected and spatially aligned through application of the alignment matrix, and the total nonlinear warp was applied to the dataset of each TE. No time series filtering was applied in the preprocessing phase. No spatial smoothing was applied during preprocessing.

The preprocessed multi-echo time-series datasets were then used by the ME-ICA pipeline to leverage information in the multiple echoes to compute an optimal weighting of TE at each voxel[67], producing an 'optimally combined' time-series dataset. This optimal combination procedure has been shown to double temporal signal-to-noise ratio (tSNR) over traditional single

echo EPI data[68]. This preprocessed optimally combined time-series dataset was then fed into a denoising procedure based on independent components analysis (ICA) and scoring components by ρ and κ pseudo-F statistics that indicate degree of TE-independence or TE-dependence. Components with high ρ and low κ are components high in non-BOLD related contrast (i.e. non-BOLD artefact signal), while components with high κ and low ρ indicate components high in BOLD-related contrast. ME-ICA identifies in an automated fashion high ρ and low κ non-BOLD components and removes them from the optimally combined time-series dataset to produce the final multi-echo denoised dataset. This procedure has been shown to be very effective in removing various types of non-BOLD artefact from rsfMRI data, including head motion artefact, flattens DVARS traces induced by head motion, and increases tSNR by a factor of 4 over and above traditional single echo EPI data[65,66,68,69]. The final multi-echo denoised datasets were used in further connectivity analyses. Head motion estimates and DVARS were estimated in order to show the impact of denoising on reducing non-BOLD artefact due to head motion (see Supplementary Figure 1 for examples). In the EU-AIMS LEAP data, groups did not differ in mean FD (see Table 2).

EU-AIMS LEAP Functional Connectivity Analyses

To assess large-scale intrinsic functional organization of the brain we input the multi-echo denoised data into a group-ICA analysis. Dual regression was then utilized to back-project spatial maps and individual time series for each component and subject. Both group-ICA and dual regression were implemented with FSL's MELODIC and Dual Regression tools (www.fmrib.ox.ac.uk/fsl). For group-ICA, we constrained the dimensionality estimate to 30. Of the 30 final components, 11 were discarded after visual examination of spatial maps indicated that they did not correspond to well-known rsfMRI networks and instead resembled white matter or other artefacts[70].

Time courses for each subject and each independent component (IC) were used to model between-component connectivity. This was achieved by constructing a partial correlation matrix amongst all 19 components using Tikhonov-regularization (i.e. ridge regression, rho=1) as implemented within the *nets_netmats.m* function in the FSLNets MATLAB toolbox (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets). The aim of utilizing partial correlations was to estimate direct connection strengths in a more accurate manner than can be achieved with full correlations, which allow more for indirect connections to influence connectivity strength[71–73]. Partial correlations were then converted into Z-statistics using Fisher's transformation for further statistical analyses. The lower diagonal of each subject's partial correlation matrix was extracted for a total of 171 separate component-pair comparisons.

To identify replicable subtype effects on functional connectivity, we partitioned the EU-AIMS LEAP dataset into Discovery and Replication sets. This was achieved via a random half split of the subtypes within each scanning site and balancing for sex. TD comparison groups for Discovery and Replication sets were also selected via a random split balancing sex and achieving an age-match (achieved using the *MatchIt* library in R with the default method of nearest neighbor matching). Models implementing the main hypothesis tests of subtype differences were computed as linear mixed effect models (*lme* function from the *nlme* library in R), whereby connectivity was the dependent variable, and subtype, sex, and age were used as fixed effect independent variables

and site was modeled with random intercepts as a random effect. These models were computed separately for the Discovery and Replication set. Connectivity pairs were deemed as showing replicable subtype differences if the Discovery set showed an effect at $p<0.05$ and the replication Bayes Factor statistic[74] computed on t-statistics from Discovery and Replication sets was greater than 10 (*repBF*>10), indicating strong evidence in favor of replication.

Because subtyping depends on the choice of a z-threshold, we ran the connectivity analyses across a range of z-thresholds from z=0.5 to z=1, moving up in steps of 0.1. Across all these z-thresholds, we identified 'consensus edges', defined as replicable subtype connectivity differences that appear at all z-thresholds. These edges are focused since they are the robust subtype connectivity differences that are not dependent on a particular z-threshold for labeling the subtypes. For each threshold, we also counted up the number of edges that are common across subtypes and with similar directionality in order to estimate how often subtypes show similar functional connectivity differences.

To contrast the subtyping approach to a more dimensional approach where z-normalized SC-RRB differences scores are left continuous, we also ran similar mixed effect models where these continuous scores are the primary independent variable of interest rather than the subtype variable. Because the z-normalized difference score does not capture overall severity level well (e.g., an individual with low SC and RBB has a difference near 0 just like an individual with high SC and RRB), we also ran models whereby continuous SC or RRB scores were used as independent variables rather than the z-normalized difference score. This allowed for another contrast to test if overall level of severity within each domain could explain connectivity strength. In each of these dimensional models, the same criteria for identifying replicable effects in the subtype models was used (e.g., $p<0.05$ in the Discovery set and a *repBF* > 10).

Gene expression decoding analyses

To identify genes whose spatial expression pattern is similar to subtype-specific ICs, we used the Gene Expression Decoding feature embedded within Neurosynth[23] to identify genes that are statistically similar in their expression profile in a consistent manner across all 6 donor brains within the Allen Institute Human Brain Atlas[22]. The analysis first utilizes a linear model to compute similarity between the observed rsfMRI IC map and spatial patterns of gene expression for each of the six brains in the Allen Institute dataset. The slopes of these subject-specific linear models encode how similar each gene's spatial expression pattern is with our rsfMRI IC maps. These slopes were then subjected to a one-sample t-test to identify genes whose spatial expression patterns are consistently of high similarity across the donor brains to the rsfMRI IC maps we input. This analysis was restricted to cortical tissue since all of the networks being analyzed are primarily cortical. The resulting list of genes was then thresholded for multiple comparisons and only the genes surviving FDR $q < 0.05$ and also had a positive t-statistic value were considered.

Enrichment analyses with autism-associated gene lists

To test if network-associated genes were enriched for different classes of autism-associated genes we first curated a list of genes known at genetic and transcriptomic levels to be associated with autism. At the genetic level, we utilized the list of 102 genes reported by Satterstrom et al.,[75]

that are rare de novo protein truncating variants that are associated with a diagnosis of autism (ASD dnPTV). A second list of autism-associated genes (ASD SFARI) at the genetic level was the list curated by SFARI Gene (https://gene.sfari.org/). We utilized the entire list of genes in categories S, 1, 2, and 3 for these analyses (downloaded on January 14, 2020). At the transcriptomic level we used several lists. First, we used a list of differentially expressed genes in autism post-mortem frontal and temporal cortex tissue from Gandal et al.,[26] and this list was further split by genes that were downregulated (ASD DE Downreg) or upregulated (ASD DE Upreg) in autism. To contrast these enrichments with other psychiatric diagnoses that are genetically correlated with autism, we also use differentially expressed genes in schizophrenia (SCZ DE) and bipolar disorder (BD DE), identified from the same Gandal et al., study[26]. To go beyond differentially expressed genes in bulk tissue samples, we also examined autism differentially expressed genes identified in specific cell types - particularly, excitatory (ASD Excitatory) and inhibitory (ASD Inhibitory) neurons, microglia (ASD Microglia), astrocytes (ASD Astrocyte), oligodendrocytes (Oligodendrocyte), and endothelial (ASD Endothelial) cells[28]. Beyond differentially expressed genes, we utilized all genes identified in frontal and temporal cortical tissue that were members of co-expression modules identified to be downregulated (ASD CTX Downreg CoExpMods) or upregulated (ASD CTX Upreg CoExpMods) in autism[27]. All tests of enrichment were conducted with custom code written in R that computes enrichment odds ratios and p-values based on hypergeometric tests. The background total for these enrichment tests was set to 20,787, which is the total number of genes considered by the gene expression decoding analysis in Neurosynth. FDR was computed amongst all of the enrichment tests done and only tests that survived FDR $q < 0.05$ were interpreted further as statistically significant enrichments.

Data and code availability

Tidy data and reproducible analysis code for this study is available at https://github.com/landiit/adir_subtyping.

Acknowledgments

We thank all participants and their families for participating in this study. We also acknowledge the contributions of all members of the EU-AIMS LEAP group. Members of the EU-AIMS LEAP group are as follows: Jumana Ahmad, Sara Ambrosino, Bonnie Auyeung, Tobias Banaschewski, Simon Baron-Cohen, Sarah Baumeister, Christian F. Beckmann, Sven Bölte, Thomas Bourgeron, Carsten Bours, Michael Brammer, Daniel Brandeis, Claudia Brogna, Yvette de Bruijn, Jan K. Buitelaar, Bhismadev Chakrabarti, Tony Charman, Ineke Cornelissen, Daisy Crawley, Flavio Dell'Acqua, Guillaume Dumas, Sarah Durston, Christine Ecker, Jessica Faulkner, Vincent Frouin, Pilar Garcés, David Goyard, Lindsay Ham, Hannah Hayward, Joerg Hipp, Rosemary Holt, Mark H. Johnson, Emily J. H. Jones, Prantik Kundu, Meng-Chuan Lai, Xavier Liogier D'ardhuy, Michael V. Lombardo, Eva Loth, David J. Lythgoe, René Mandl, Andre Marquand, Luke Mason, Maarten Mennes, Andreas Meyer-Lindenberg, Carolin Moessnang, Nico Mueller, Declan G. M. Murphy, Bethany Oakley, Laurence O'Dwyer, Marianne Oldehinkel, Bob Oranje, Gahan Pandina, Antonio M. Persico, Barbara Ruggeri, Amber N. V. Ruigrok, Jessica Sabet, Roberto Sacco, Antonia San José Cáceres, Emily Simonoff, Will Spooren, Julian Tillmann, Roberto Toro, Heike Tost, Jack Waldman, Steve C. R. Williams, Caroline Wooldridge, and Marcel P. Zwiers.

Disclosures

JKB has been a consultant to, advisory board member of, and a speaker for Janssen Cilag BV, Eli Lilly, Shire, Lundbeck, Roche, and Servier. He is not an employee of any of these companies and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents, or royalties. CFB is director and shareholder in SBGneuro Ltd. SB discloses that he has in the last 5 years acted as an author, consultant, or lecturer for Shire/Takeda, Medice, Roche, Eli Lilly, and Prima Psychiatry. He receives royalties for textbooks and diagnostic tools from Huber/Hogrefe, Kohlhammer, and UTB. TC has received consultancy from Roche and Servier and received book royalties from Guildford Press and Sage. DGMM has been a consultant to, and advisory board member, for Roche and Servier. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. AML has received consultant fees from Boehringer Ingelheim, Elsevier, Brainsway, Lundbeck Int. Neuroscience Foundation, Lundbeck A/S, The Wolfson Foundation, Bloomfield Holding Ltd, Shanghai Research Center for Brain Science, Thieme Verlag, Sage Therapeutics, v Behring Röntgen Stiftung, Fondation FondaMental, Janssen-Cilag GmbH, MedinCell, Brain Mind Institute, Agence Nationale de la Recherche, CISSN (Catania Internat. Summer School of Neuroscience), Daimler und Benz Stiftung and American Association for the Advancement of Science. Additionally he has received speaker fees from Italian Society of Biological Psychiatry, Merz-Stiftung, Forum Werkstatt Karlsruhe, Lundbeck SAS France, BAG Psychiatrie Oberbayern, Klinik für Psychiatrie und Psychotherapie Ingolstadt, med Update GmbH, Society of Biological Psychiatry and Siemens Healthineers. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. JT is a consultant to Roche. The other authors report no biomedical financial interests or potential conflicts of interest.

Funding

This work was supported by an ERC Starting Grant (ERC-2017-STG; 755816) to MVL. This work was also supported by EU-AIMS and EU AIMS-2-TRIALS, which both received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No. 115300 and the Innovative Medicines Initiative 2 Joint Undertaking under Grant Agreement No. 777394, the resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (Grant No. FP7/2007-2013), from the European Federation of Pharmaceutical Industries and Associations companies' in-kind contributions, and from Autism Speaks, Autistica and the Simons Foundation for Autism Research Initiative. This work was also supported by the Netherlands Organization for Scientific Research through Vidi grants (Grant No. 864.12.003 [to CFB]); from the FP7 (Grant Nos. 602805) (AGGRESSOTYPE) (to JKB), 603016 (MATRICS), and 278948 (TACTICS); and from the European Community's Horizon 2020 Programme (H2020/2014-2020) (Grant Nos. 643051 [MiND] and 642996 (BRAINVIEW). This work received funding from the Wellcome Trust UK Strategic Award (Award No. 098369/Z/12/Z) and from the National Institute for Health Research Maudsley Biomedical Research Centre (to DGMM). M-CL was supported by the Academic Scholars Award from the Department of Psychiatry, University of Toronto, the Slaight Family Child and Youth Mental Health Innovation Fund from the CAMH Foundation, the Ontario Brain Institute via the Province of Ontario Neurodevelopmental Disorders (POND) Network (IDS-I l-02), the Canadian Institutes of Health Research (PJT 159578), and the Innovation Fund of the Alternative Funding Plan for the Academic Health Sciences Centres of Ontario (CAM-20-004). R.A.I.B acknowledges research support by the Autism Research Trust and a British Academy Fellowship (PF2\180017). MHJ, TC, and EJHJ acknowledge support from a UK MRC Programme Grant.

References

1. Lai, M.-C., Lombardo, M. V. & Baron-Cohen, S. Autism. *Lancet* **383**, 896–910 (2014).
2. Lord, C. *et al.* Autism spectrum disorder. *Nat Rev Dis Primers* **6**, 5 (2020).
3. Lombardo, M. V., Lai, M.-C. & Baron-Cohen, S. Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol. Psychiatry* **24**, 1435–1450 (2019).
4. Happé, F. & Ronald, A. The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychol Rev* **18**, 287–304 (2008).
5. Graybiel, A. M. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31**, 359–387 (2008).
6. Langen, M., Durston, S., Kas, M. J. H., van Engeland, H. & Staal, W. G. The neurobiology of repetitive behavior: …and men. *Neurosci Biobehav Rev* **35**, 356–365 (2011).
7. Kennedy, D. P. & Adolphs, R. The social brain in psychiatric and neurological disorders. *Trends Cogn. Sci. (Regul. Ed.)* **16**, 559–572 (2012).
8. Ronald, A., Happe, F. & Plomin, R. The genetic relationship between individual differences in social and nonsocial behaviours characteristic of autism. *Developmental Sci* **8**, 444–458 (2005).
9. Ronald, A., Happé, F., Price, T. S., Baron-Cohen, S. & Plomin, R. Phenotypic and Genetic Overlap Between Autistic Traits at the Extremes of the General Population. *Journal of the American Academy of Child & Adolescent Psychiatry* **45**, 1206–1214 (2006).
10. Ronald, A. *et al.* Genetic Heterogeneity Between the Three Components of the Autism Spectrum: A Twin Study. *Journal of the American Academy of Child & Adolescent Psychiatry* **45**, 691–699 (2006).
11. Warrier, V. *et al.* Social and non-social autism symptoms and trait domains are genetically dissociable. *Commun Biol* **2**, 328 (2019).
12. Georgiades, S. *et al.* Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modeling approach. *J Child Psychol Psychiatry* **54**, 206–215 (2013).
13. Hu, V. W. & Steinberg, M. E. Novel clustering of items from the Autism Diagnostic Interview-Revised to define phenotypes within autism spectrum disorders. *Autism Res* **2**, 67–77 (2009).
14. Cholemkery, H., Medda, J., Lempp, T. & Freitag, C. M. Classifying Autism Spectrum Disorders by ADI-R: Subtypes or Severity Gradient? *J Autism Dev Disord* **46**, 2327–2339 (2016).
15. Happé, F. & Frith, U. Annual Research Review: Looking back to look forward - changes in the concept of autism and implications for future research. *J Child Psychol Psychiatry* **61**, 218–232 (2020).
16. Richiardi, J. *et al.* BRAIN NETWORKS. Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–1244 (2015).
17. Hawrylycz, M. *et al.* Canonical genetic signatures of the adult human brain. *Nat. Neurosci.* **18**, 1832–1844 (2015).
18. Fornito, A., Arnatkevičiūtė, A. & Fulcher, B. D. Bridging the Gap between Connectome and Transcriptome. *Trends Cogn. Sci. (Regul. Ed.)* **23**, 34–50 (2019).
19. Loth, E. *et al.* The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. *Mol Autism* **8**, 24 (2017).

20. Charman, T. *et al.* The EU-AIMS Longitudinal European Autism Project (LEAP): clinical characterisation. *Mol Autism* **8**, 27 (2017).
21. Oldehinkel, M. *et al.* Altered Connectivity Between Cerebellum, Visual, and Sensory-Motor Networks in Autism Spectrum Disorder: Results from the EU-AIMS Longitudinal European Autism Project. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **4**, 260–270 (2019).
22. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
23. Gorgolewski, K. J. *et al.* Tight fitting genes: finding relations between statistical maps and gene expression patterns. *F1000 Posters* **5**, 1607 (2014).
24. Romero-Garcia, R., Warrier, V., Bullmore, E. T., Baron-Cohen, S. & Bethlehem, R. A. I. Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. *Mol. Psychiatry* **24**, 1053–1064 (2019).
25. Chatham, C. H. *et al.* Adaptive behavior in autism: Minimal clinically important differences on the Vineland-II. *Autism Res* **11**, 270–283 (2018).
26. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, eaat8127 (2018).
27. Parikshak, N. N. *et al.* Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* **540**, 423–427 (2016).
28. Velmeshev, D. *et al.* Single-cell genomics identifies cell type–specific molecular changes in autism. *Science* **364**, 685–689 (2019).
29. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
30. Chen, C. P. *et al.* Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. *NeuroImage: Clinical* **8**, 238–245 (2015).
31. Holiga, Š. *et al.* Patients with autism spectrum disorders display reproducible functional connectivity alterations. *Sci. Transl. Med.* **11**, eaat9223 (2019).
32. Lombardo, M. V. *et al.* Different functional neural substrates for good and poor language outcome in autism. *Neuron* **86**, 567–577 (2015).
33. Lombardo, M. V. *et al.* Large-scale associations between the leukocyte transcriptome and BOLD responses to speech differ in autism early language outcome subtypes. *Nat. Neurosci.* **21**, 1680–1688 (2018).
34. Redcay, E. & Courchesne, E. Deviant Functional Magnetic Resonance Imaging Patterns of Brain Activity to Speech in 2–3-Year-Old Children with Autism Spectrum Disorder. *Biological Psychiatry* **64**, 589–598 (2008).
35. Eyler, L. T., Pierce, K. & Courchesne, E. A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* **135**, 949–960 (2012).
36. Dinstein, I. *et al.* Disrupted Neural Synchronization in Toddlers with Autism. *Neuron* **70**, 1218–1225 (2011).
37. Adolphs, R., Damasio, H., Tranel, D., Cooper, G. & Damasio, A. R. A Role for Somatosensory Cortices in the Visual Recognition of Emotion as Revealed by Three-Dimensional Lesion Mapping. *J. Neurosci.* **20**, 2683–2690 (2000).
38. Keysers, C., Kaas, J. H. & Gazzola, V. Somatosensation in social perception. *Nat Rev Neurosci* **11**, 417–428 (2010).

39. Stevenson, R. A. *et al.* Multisensory Temporal Integration in Autism Spectrum Disorders. *J. Neurosci.* **34**, 691–697 (2014).
40. Foss-Feig, J. H. *et al.* An extended multisensory temporal binding window in autism spectrum disorders. *Exp Brain Res* **203**, 381–389 (2010).
41. Russo, N. *et al.* Multisensory processing in children with autism: high-density electrical mapping of auditory-somatosensory integration. *Autism Res* **3**, 253–267 (2010).
42. Bhat, A. N., Landa, R. J. & Galloway, J. C. Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Phys Ther* **91**, 1116–1129 (2011).
43. Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N. & Cauraugh, J. H. Motor coordination in autism spectrum disorders: a synthesis and meta-analysis. *J Autism Dev Disord* **40**, 1227–1240 (2010).
44. Green, D. *et al.* Impairment in movement skills of children with autistic spectrum disorders. *Dev Med Child Neurol* **51**, 311–316 (2009).
45. Nebel, M. B. *et al.* Intrinsic Visual-Motor Synchrony Correlates With Social Deficits in Autism. *Biol. Psychiatry* **79**, 633–641 (2016).
46. Glazebrook, C., Gonzalez, D., Hansen, S. & Elliott, D. The role of vision for online control of manual aiming movements in persons with autism spectrum disorders. *Autism* **13**, 411–433 (2009).
47. Dowd, A. M., McGinley, J. L., Taffe, J. R. & Rinehart, N. J. Do planning and visual integration difficulties underpin motor dysfunction in autism? A kinematic study of young children with autism. *J Autism Dev Disord* **42**, 1539–1548 (2012).
48. Crippa, A., Forti, S., Perego, P. & Molteni, M. Eye-hand coordination in children with high functioning autism and Asperger's disorder using a gap-overlap paradigm. *J Autism Dev Disord* **43**, 841–850 (2013).
49. Marko, M. K. *et al.* Behavioural and neural basis of anomalous motor learning in children with autism. *Brain* **138**, 784–797 (2015).
50. Uddin, L. Q. *et al.* Salience Network–Based Classification and Prediction of Symptom Severity in Children With Autism. *JAMA Psychiatry* **70**, 869 (2013).
51. Green, S. A., Hernandez, L., Bookheimer, S. Y. & Dapretto, M. Salience Network Connectivity in Autism Is Related to Brain and Behavioral Markers of Sensory Overresponsivity. *Journal of the American Academy of Child & Adolescent Psychiatry* **55**, 618-626.e1 (2016).
52. Di Martino, A. *et al.* The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry* **19**, 659–667 (2014).
53. Rubenstein, J. L. R. & Merzenich, M. M. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* **2**, 255–267 (2003).
54. Sohal, V. S. & Rubenstein, J. L. R. Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol. Psychiatry* **24**, 1248–1257 (2019).
55. Bal, V. H., Kim, S.-H., Cheong, D. & Lord, C. Daily living skills in individuals with autism spectrum disorder from 2 to 21 years of age. *Autism* **19**, 774–784 (2015).
56. Tillmann, J. *et al.* Investigating the factors underlying adaptive functioning in autism in the EU-AIMS Longitudinal European Autism Project. *Autism Res* **12**, 645–657 (2019).
57. Lord, C., Bishop, S. & Anderson, D. Developmental trajectories as autism phenotypes. *Am J Med Genet C Semin Med Genet* **169**, 198–208 (2015).

58. Georgiades, S., Bishop, S. L. & Frazier, T. Editorial Perspective: Longitudinal research in autism - introducing the concept of 'chronogeneity'. *J Child Psychol Psychiatry* **58**, 634–636 (2017).
59. Kim, S. H. *et al.* Variability in Autism Symptom Trajectories Using Repeated Observations From 14 to 36 Months of Age. *J Am Acad Child Adolesc Psychiatry* **57**, 837-848.e2 (2018).
60. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* **24**, 659–685 (1994).
61. Huerta, M., Bishop, S. L., Duncan, A., Hus, V. & Lord, C. Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *Am J Psychiatry* **169**, 1056–1064 (2012).
62. Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set. *J. Stat. Soft.* **61**, (2014).
63. Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2008).
64. Lombardo, M. V. *et al.* Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Sci Rep* **6**, 35333 (2016).
65. Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M. & Bandettini, P. A. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage* **60**, 1759–1770 (2012).
66. Kundu, P. *et al.* Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. *Neuroimage* **154**, 59–80 (2017).
67. Posse, S. *et al.* Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magn Reson Med* **42**, 87–97 (1999).
68. Kundu, P. *et al.* Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16187–16192 (2013).
69. Lombardo, M. V. *et al.* Improving effect size estimation and statistical power with multi-echo fMRI and its impact on understanding the neural systems supporting mentalizing. *Neuroimage* **142**, 55–66 (2016).
70. Griffanti, L. *et al.* Hand classification of fMRI ICA noise components. *Neuroimage* **154**, 188–205 (2017).
71. Smith, S. M. *et al.* Functional connectomics from resting-state fMRI. *Trends Cogn. Sci. (Regul. Ed.)* **17**, 666–682 (2013).
72. Smith, S. M. *et al.* Network modelling methods for FMRI. *Neuroimage* **54**, 875–891 (2011).
73. Marrelec, G. *et al.* Partial correlation for functional brain interactivity investigation in functional MRI. *Neuroimage* **32**, 228–237 (2006).
74. Verhagen, J. & Wagenmakers, E.-J. Bayesian tests to quantify the result of a replication attempt. *J Exp Psychol Gen* **143**, 1457–1475 (2014).
75. Satterstrom, F. K. *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* (2020) doi:10.1016/j.cell.2019.12.036.

Tables

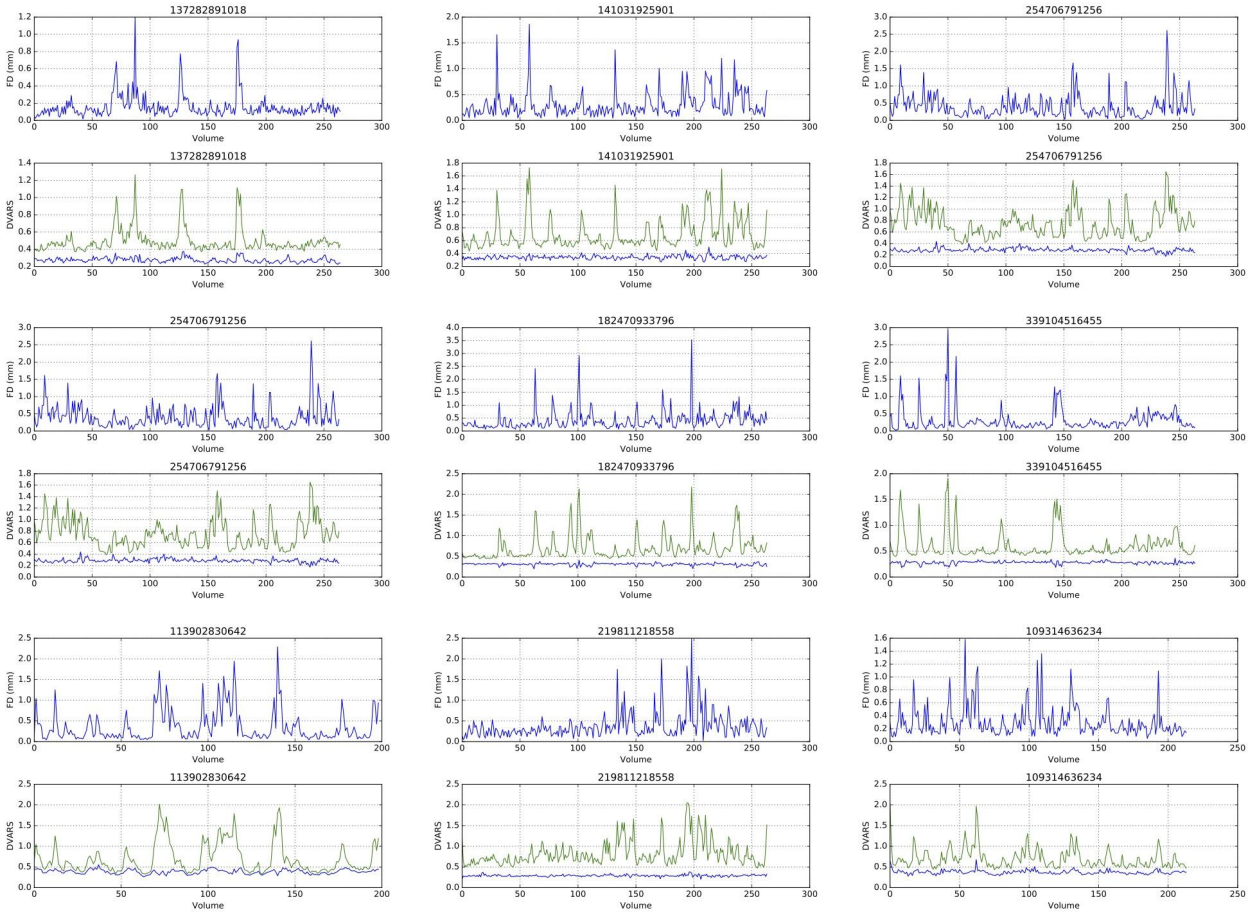
	Discovery				Replication			
	SC>RRB	**SC=RRB**	**RRB>SC**	**F (p-val)**	**SC>RRB**	**SC=RRB**	**RRB>SC**	**F (p-val)**
N (male)	137 (109)	611 (468)	141 (115)	-	124 (97)	629 (482)	137 (115)	-
Age in years	7.24 (5.50)	9.03 (5.38)	10.05 (4.02)	10.61 (2.77e-5)	6.92 (5.14)	9.10 (5.54)	9.76 (4.35)	10.80 (2.31e-5)
ADI-R SC	0.51 (0.13)	0.31 (0.14)	0.21 (0.10)	212.57 (2.2e-16)				191.67 (2.2e-16)
ADI-R RRB	0.21 (0.10)	0.32 (0.14)	0.49 (0.12)	174.19 (2.2e-16)				184.74 (2.2e-16)
ADOS SA CSS[i]	7.46 (1.60)	6.92 (2.01)	6.11 (2.49)	2.85 (0.06)	6.96 (1.73)	6.93 (2.04)	6.62 (2.39)	1.83 (0.16)
ADOS RRB CSS[i]	8.37 (1.31)	7.68 (2.30)	7.37 (2.29)	0.09 (0.91)	7.42 (2.55)	7.40 (2.30)	6.62 (1.06)	0.43 (0.64)
FIQ[i]	106.83 (17.10)	103.89 (18.97)	102.69 (14.59)	0.32 (0.72)	117.27 (13.76)	105.44 (17.96)	106.92 (15.88)	2.38 (0.09)

Table 1: Participant characteristics from the NDAR datasets. *At a z-threshold of 1, this table shows sample sizes and descriptive statistics (mean and standard deviation) for age and ADOS social affect (SA) and restricted repetitive behaviors (RRB) calibrated severity scores. The final column on the right shows the F-statistic and p-value from an ANOVA testing for an effect of group.* [i] *Sample sizes: ADOS (Discovery, RRB>SC n=19, SC=RRB n=99, SC>RRB n=35; Replication RRB>SC n=8, SC=RRB n=119, SC>RRB n=26); FIQ (Discovery, RRB>SC n=39, SC=RRB n=142, SC>RRB n=18; Replication RRB>SC n=40, SC=RRB n=135, SC>RRB n=11). Abbreviations: FIQ = full-scale IQ; ADI-R = Autism Diagnostic Interview Revised; ADOS = Autism Diagnostic Observation Schedule; SC = social-communication; RRB = restricted repetitive behaviors; SA = social affect; CSS = calibrated severity score.*

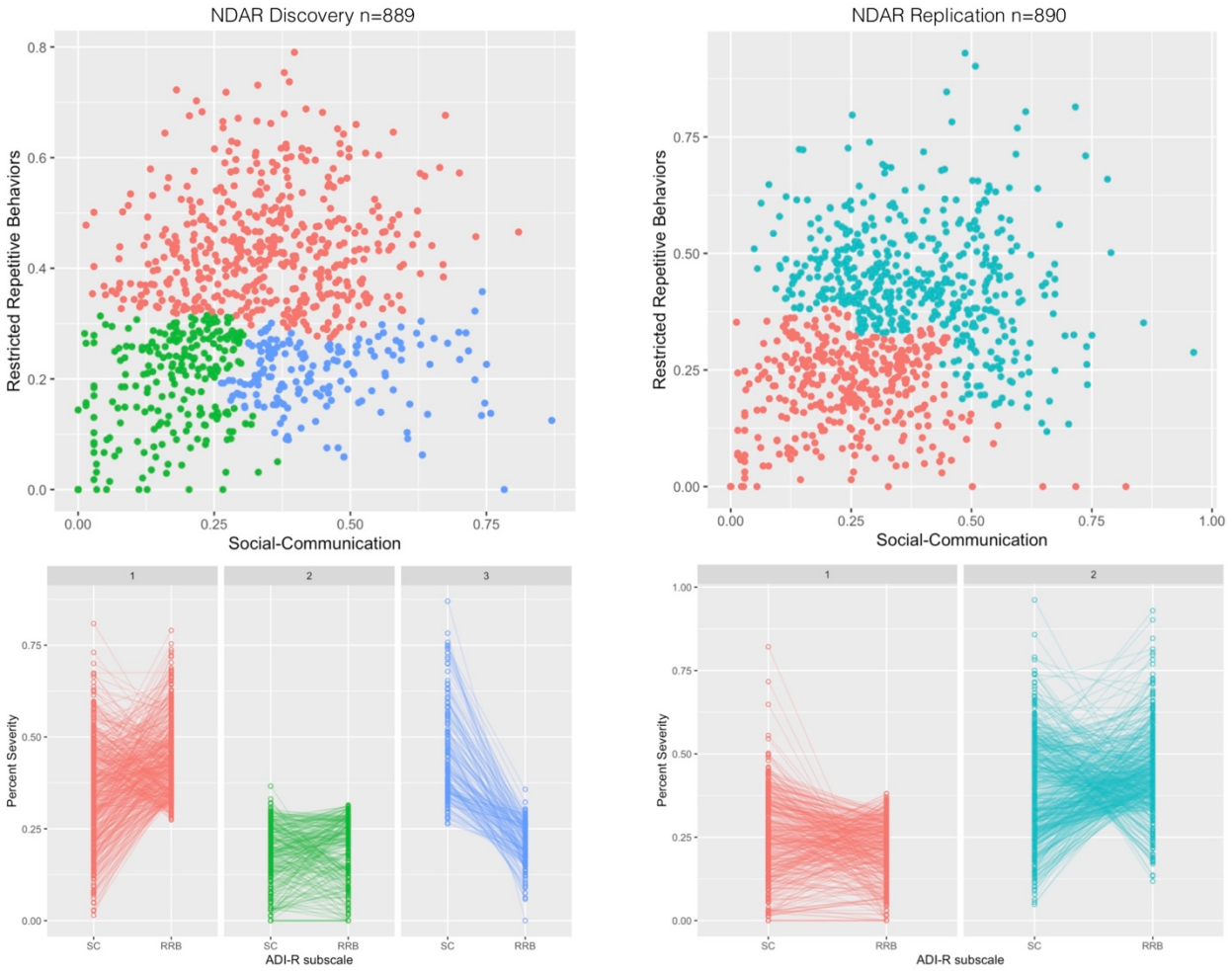
	Discovery					Replication				
	SC>RRB	**SC=RRB**	**RRB>SC**	**TD**	**F (p-val)**	**SC>RRB**	**SC=RRB**	**RRB>SC**	**TD**	**F (p-val)**
N (male)	47 (34)	80 (61)	6 (3)	121 (80)	-	47 (34)	85 (62)	1 (1)	122 (75)	-
Age in years	15.69 (4.50)	16.62 (5.91)	18.08 (7.13)	16.83 (5.23)	0.78 (0.50)	16.00 (5.06)	16.68 (5.73)	11.45 (N/A)	16.86 (6.07)	0.77 (0.50)
FIQ	97.27 (18.34)	99.50 (18.80)	102.67 (14.63)	105.72 (18.33)	2.45 (0.06)	94.46 (20.00)	103.62 (16.50)	148.00 (N/A)	104.94 (17.14)	4.78 (0.002)
Mean FD	0.24 (0.25)	0.28 (0.45)	0.29 (0.42)	105.72 (18.33)	2.03 (0.10)	0.24 (0.22)	0.25 (0.27)	0.38 (N/A)	0.23 (0.46)	0.09 (0.96)
ADI-R Soc	19.32 (5.75)	16.43 (7.01)	15.83 (8.33)	-	6.20 (0.01)	19.68 (5.77)	14.91 (5.88)	16.00 (N/A)	-	21.48 (8.72e-6)
ADI-R Comm	15.02 (4.73)	13.74 (5.94)	14.67 (6.31)	-	2.05 (0.14)	15.83 (4.41)	11.81 (5.39)	7.00 (N/A)	-	20.49 (1.35e-5)
ADI-R RRB	3.15 (1.93)	4.89 (2.68)	9.00 (1.26)		18.07 (4.17e-5)	3.98 (2.54)	3.88 (2.30)	6.00 (N/A)		0.04 (0.84)
ADOS SA CSS[i]	6.44 (2.55)	6.15 (2.58)	3.83 (3.37)	-	0.44 (0.50)	6.27 (3.01)	5.65 (2.50)	5.00 (N/A)	-	2.14 (0.14)
ADOS RRB CSS[i]	4.73 (2.82)	4.84 (2.81)	4.83 (4.22)	-	0.16 (0.68)	4.66 (2.79)	4.82 (2.56)	1.00 (N/A)	-	0.70 (0.40)
SRS[i]	74.57 (9.25)	71.26 (13.00)	71.33 (16.65)	47.84 (9.40)	0.005 (0.25)	77.67 (9.68)	66.85 (11.63)	60.00 (N/A)	47.23 (9.34)	0.48 (0.48)
RBS[i]	15.37 (13.26)	18.30 (16.42)	17.33 (10.69)	2.15 (4.74)	1.23 (0.52)	22.56 (15.58)	13.08 (11.28)	13.00 (N/A)	3.08 (11.54)	15.14 (1.72e-4)
SSP[i]	137.66 (28.21)	136.62 (31.32)	143.00 (33.94)	177.86 (12.71)	0.38 (0.69)	133.19 (23.72)	143.16 (26.38)	153.00 (N/A)	174.93 (19.38)	3.47 (0.06)
VABS Comm[i]	69.09 (16.07)	73.64 (17.57)	86.50 (20.68)	91.97 (25.44)	1.95 (0.16)	68.95 (14.28)	82.16 (13.77)	99.00 (N/A)	92.74 (25.45)	24.68 (2.37e-6)
VABS DLS[i]	66.86 (14.78)	74.39 (17.03)	71.00 (9.63)	90.74 (20.28)	6.07 (0.01)	67.84 (15.74)	78.36 (15.20)	74.00 (N/A)	91.10 (22.65)	13.14 (4.33e-4)
VABS Soc[i]	65.63 (16.06)	71.01 (15.55)	70.75 (10.34)	96.21 (23.67)	3.34 (0.07)	62.98 (15.78)	76.00 (15.14)	76.00 (N/A)	98.90 (27.04)	19.83 (1.96e-5)
VABS ABC[i]	64.93 (15.25)	71.57 (13.22)	74.00 (12.03)	92.06 (23.04)	6.12 (0.01)	64.81 (13.67)	77.18 (12.77)	81.00 (N/A)	93.00 (25.39)	25.55 (1.65e-6)

Table 2: Participant characteristics from the EU-AIMS LEAP dataset. *At a z-threshold of 1, this table shows sample sizes and descriptive statistics alongside F-statistic and p-value from an ANOVA testing for an effect of group. For ADI-R, ADOS, SRS, RBS, SSP, and VABS, the F-statistic and p-value refer to a group difference between SC=RRB vs SC>RRB, while for age, mean FD, and FIQ, the F-statistic and p-value refer to a model that takes into account all groups.* [i] *Sample sizes: ADOS (Discovery, RRB>SC n=6, SC=RRB n=79, SC>RRB n=45; Replication RRB>SC n=1, SC=RRB n=83, SC>RRB n=44); SRS (Discovery, RRB>SC n=3, SC=RRB n=69, SC>RRB n=42, TD n=68; Replication RRB>SC n=1, SC=RRB n=78, SC>RRB n=39, TD n=65); RBS (Discovery, RRB>SC n=3, SC=RRB n=66, SC>RRB n=41, TD n=68; Replication RRB>SC n=1, SC=RRB n=75, SC>RRB n=39, TD n=63); SSP (Discovery, RRB>SC n=2, SC=RRB n=45, SC>RRB n=32, TD n=59; Replication RRB>SC n=1, SC=RRB n=51, SC>RRB n=31, TD n=54); Vineland (Discovery, RRB>SC n=4, SC=RRB n=75, SC>RRB n=43, TD n=34; Replication RRB>SC n=1, SC=RRB n=76, SC>RRB n=43, TD n=39). Abbreviations: FD = framewise displacement; FIQ = full-scale IQ; ADI-R = Autism Diagnostic Interview Revised; ADOS = Autism Diagnostic Observation Schedule; SC = social-communication; RRB = restricted repetitive behaviors; SA = social affect; CSS = calibrated severity score; SRS = Social Responsiveness Scale; RBS = Repetitive Behavior Scale; SSP = Short Sensory Profile; VABS = Vineland Adaptive Behavior Scales; Comm = Communication; DLS = Daily Living Skills; Soc = Socialization; ABC = Adaptive Behavior Composite.*

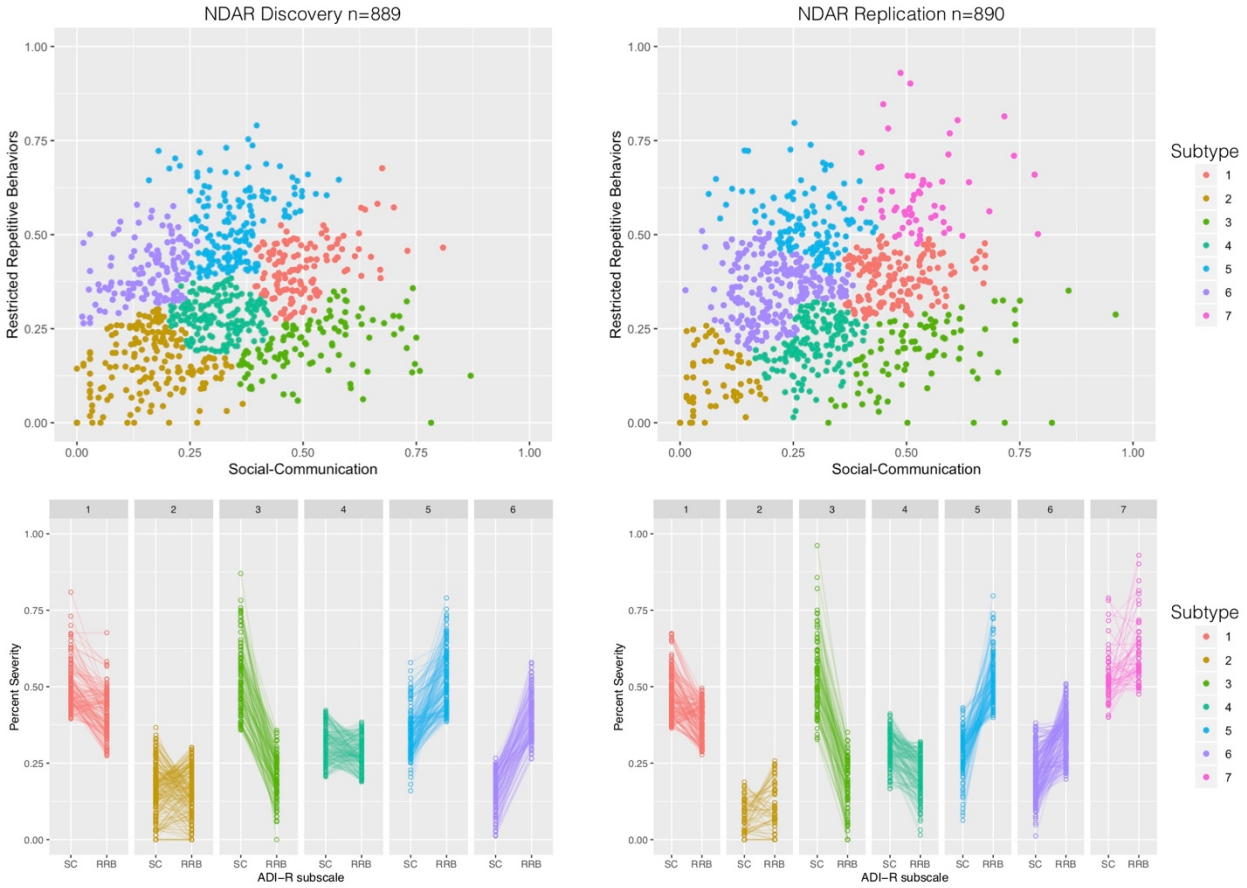
Supplementary Figures



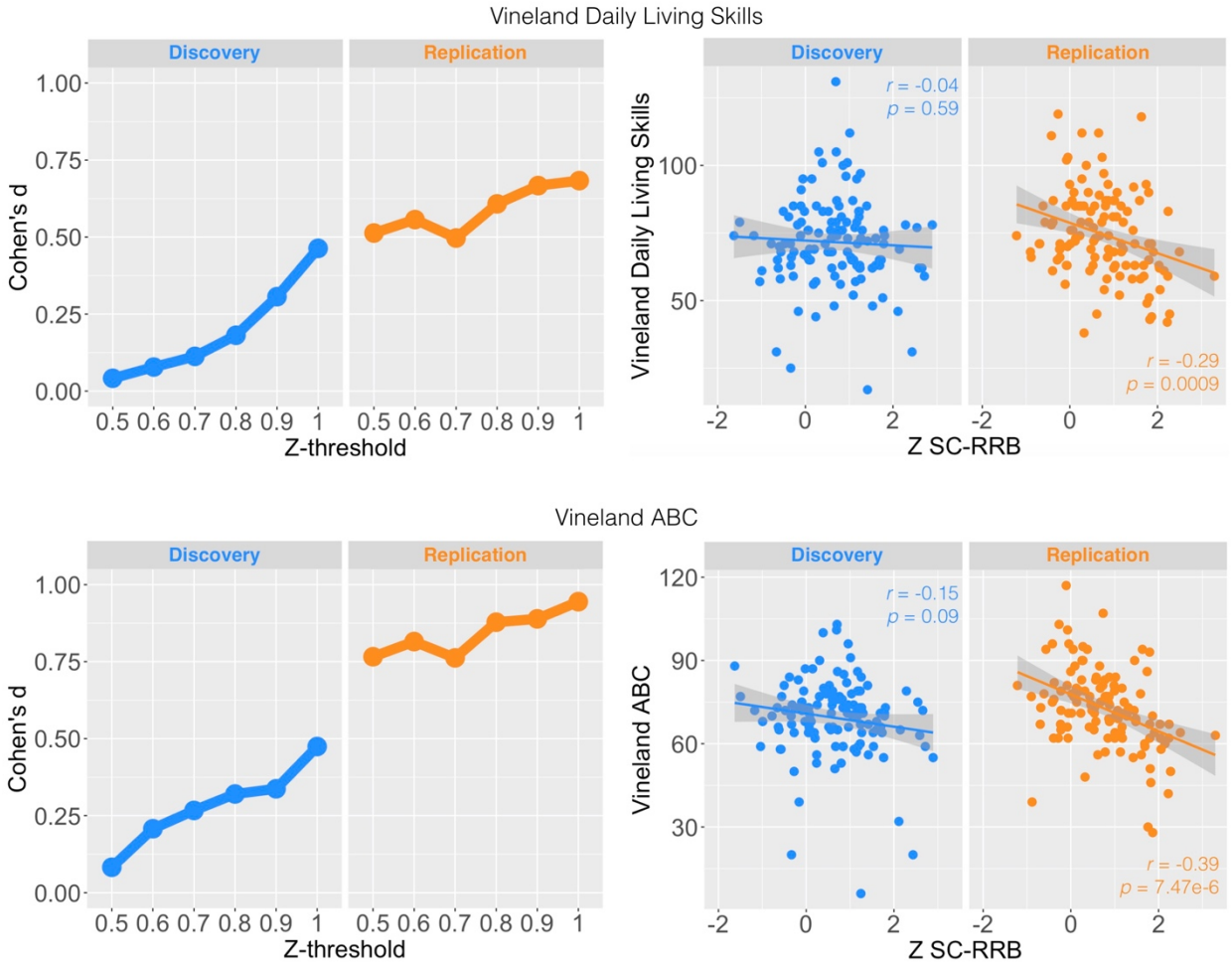
Supplementary Figure 1: Examples of head-motion derived artefact that the effect of ME-ICA at flattening DVARS. *This figure shows 9 example subjects, each with two plots. The top plot always shows framewise displacement (in mm) as an indicator of how much head motion is exhibited at each successive volume. The bottom plot shows DVARS traces from the optimally combined time-series before ME-ICA denoising (green) and after ME-ICA denoising (blue). DVARS traces before ME-ICA (green) closely follow the same pattern of framewise displacement and showcases how head motion can induce non-BOLD artefact. However, after ME-ICA denoising these DVARS traces are heavily flattened out, as a large proportion of this non-BOLD head motion artefact is isolated and removed as part of the denoising process.*



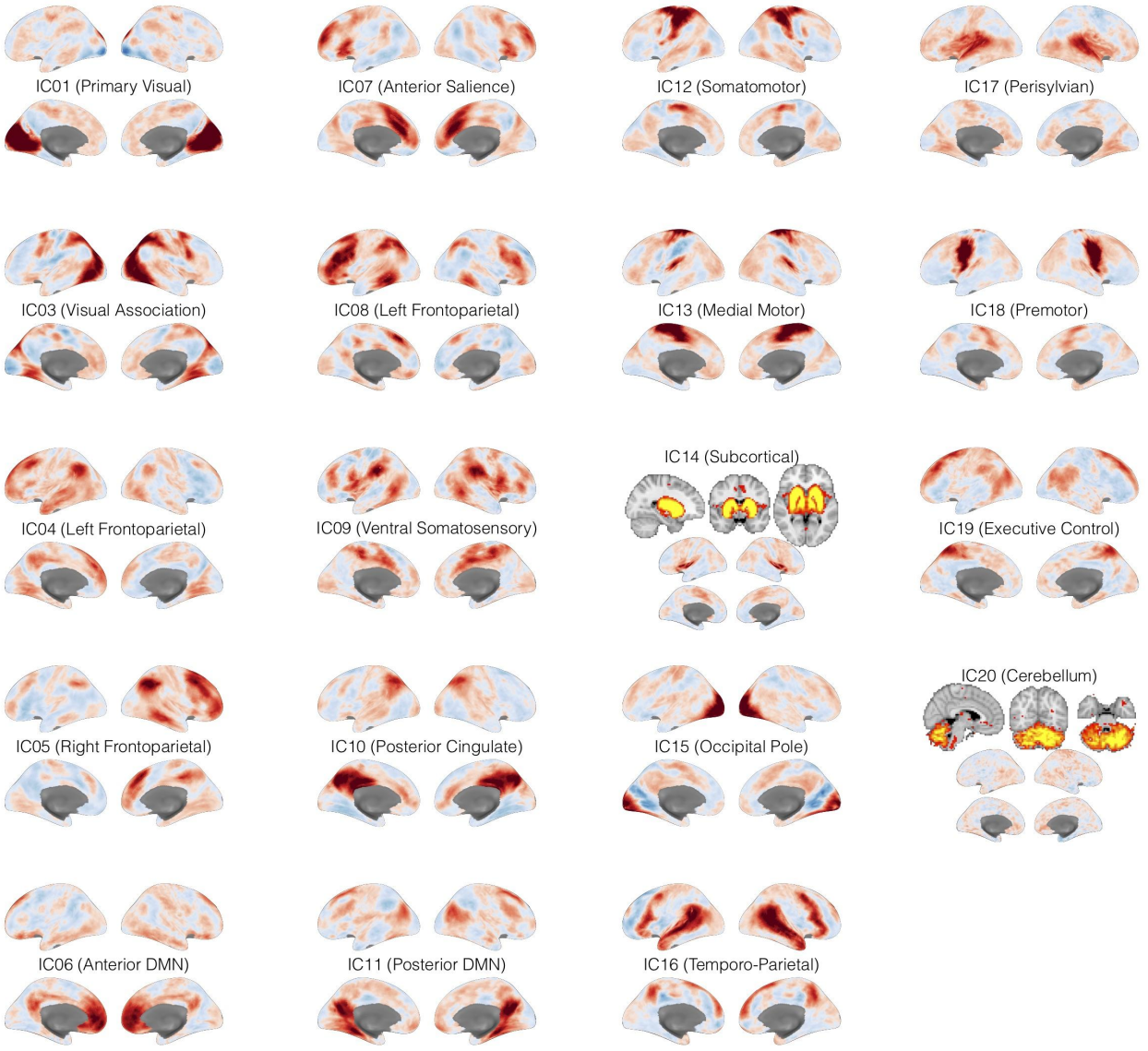
Supplementary Figure 2: Unsupervised stratification with agglomerative hierarchical clustering approaches. *Supplementary Figure 1 shows clustering results from NDAR Discovery and Replication datasets when the optimal number of clusters is selected by majority vote with NbClust.*



Supplementary Figure 3: Unsupervised stratification with agglomerative hierarchical clustering approaches. *Supplementary Figure 2 shows subtypes that emerge from clustering when the number of clusters is automatically selected with a dynamic hybrid tree cutting algorithm.*



Supplementary Figure 4: Effect sizes for SC>RRB vs SC=RRB subtype differences or dimensional association of z-normalized SC-RRB difference scores with Vineland Daily Living Skills and Adapative Behavior Composite. *In the upper left panel we show standardized effect size (Cohen's d) for a SC>RRB vs SC=RRB subtype difference on the Vineland Daily Living Skills subscale as a function of the z-threshold used to define the subtypes. These effect sizes are shown separately for Discovery (blue) and Replication (orange) sets. The bottom left panel shows the same plot, except for the Vineland Adaptive Behavior Composite score. The upper right panel shows scatterplots of the z-normalized SC-RRB difference score (x-axis) against the Vineland Daily Living Skills subscale (y-axis). Pearson's r correlations and p-values are shown for Discovery (blue) and Replication (orange) sets. The bottom right panel shows the same plot, except for the Vineland Adaptive Behavior Composite score.*



Supplementary Figure 5: Visual depiction of the 19 components identified with group-ICA. *Z-statistics are indicated with the blue-to-red color scale. Increasingly dark red regions are the regions of primary importance for the component. For components with mostly subcortical or cerebellar regions of importance (i.e. IC14 and IC20), these regions are highlighted in bright orange in axial, sagittal, and coronal planes.*

Supplementary Table 1: ADI-R items to use in DSM-5 scoring. *This scoring scheme is identical to that reported by Huerta et al., [61]. Subscales A1-A3 are within the social-communication (SC) domain while subscales B1-B4 are within the restricted repetitive behavior (RRB) domain. All items in A1-A3 utilize the Current scores, while all items in B1-B4 utilize the Ever scores.*

	< 4 years	**4 - 10 years**	**> 10 years**	**Description**
A1	*COM34,* *COM35,* *COM31,* *SOCIAL61,* *SOCIAL52,* *SOCIAL54,* *SOCIAL55,* *COM46,* *SOCIAL51*	*COM34,* *COM35,* *COM31,* *SOCIAL61,* *SOCIAL52,* *SOCIAL54,* *SOCIAL55,* *SOCIAL51*	*COM34,* *COM35,* *COM31,* *SOCIAL52,* *SOCIAL54,* *SOCIAL55,* *SOCIAL51*	*Social verbalization and chat,* *Reciprocal conversation,* *Use of other's body to communicate,* *Imitative social play,* *Showing and directing attention,* *Seeking to share his/her enjoyment with others,* *Offering comfort,* *Attention to voice,* *Social smiling*
A2	*SOCIAL50,* *COM42,* *COM43,* *COM44,* *COM45,* *SOCIAL57,* *SOCIAL56*	 *COM42,* *COM43,* *COM44,* *COM45,* *SOCIAL57,* *SOCIAL56*	 *COM42,* *COM43,* *COM44,* *COM45,* *SOCIAL57,* *SOCIAL56*	*Direct gaze,* *Pointing to express interest,* *Nodding,* *Head shaking,* *Conventional/instrumental gestures,* *Range of facial expressions to communicate,* *Quality of social overtures*
A3	*COM36,* *SOCIAL58,* *SOCIAL53,* *SOCIAL59,* *SOCIAL62,* *SOCIAL63,*	*COM36,* *SOCIAL58,* *SOCIAL53,* *SOCIAL59,* *SOCIAL62,* *SOCIAL63,* *COM49,* *SOCIAL64,* *SOCIAL65,* *SOCIAL66*	*COM36,* *SOCIAL58,* *SOCIAL53,* *SOCIAL59,* *SOCIAL65,* *SOCIAL66*	*Inappropriate questions or statements,* *Inappropriate facial expressions,* *Offering to share,* *Appropriateness of social responses,* *Interest in children,* *Responses to approaches of other children,* *Imaginative play with peers,* *Group play with peers,* *Friendships,* *Social disinhibition*
B1	*COM33,* *COM37,* *COM38,* *RRB69,* *RRB77,* *RRB78*	*COM33,* *COM37,* *COM38,* *RRB69,* *RRB77,* *RRB78*	*COM33,* *COM37,* *COM38,* *RRB69,* *RRB77,* *RRB78*	*Stereotyped utterances and delayed echolalia,* *Pronominal reversal,* *Neologisms/idiosyncratic language,* *Repetitive use of objects or interest in parts of objects,* *Hand and finger mannerisms,* *Other complex mannerisms or stereotyped body movements*
B2	*COM39,* *RRB70,* *RRB74,* *RRB75*	*COM39,* *RRB70,* *RRB74,* *RRB75*	*COM39,* *RRB70,* *RRB74,* *RRB75*	*Verbal rituals,* *Compulsions/rituals,* *Difficulties w/ minor changes in routines or personal environment* *Resistance to trivial changes in the environment (not directly affecting the subject)*
B3	*RRB67,* *RRB68,* *RRB76*	*RRB67,* *RRB68,* *RRB76*	*RRB67,* *RRB68,* *RRB76*	*Unusual preoccupations,* *Circumscribed interests,* *Unusual attachment to objects*

B4	*RRB72,* *RRB73,* *RRB71*	*RRB72,* *RRB73,* *RRB71*	*RRB72,* *RRB73,* *RRB71*	*Undue general sensitivity to noise,* *Abnormal idiosyncratic negative response to specific sensory stimuli,* *Unusual sensory interests*

Supplementary Table 2: Statistics from all functional connectivity comparisons.

Supplementary Table 3: Gene lists used for enrichment analyses.