

Reconciling Dimensional and Categorical Models of Autism Heterogeneity: a Brain Connectomics & Behavioral Study

Siyi Tang^{1,2*}, Nanbo Sun^{1*}, Dorothea L. Floris^{3,4}, Xiuming Zhang⁵, Adriana Di Martino⁶, B.T. Thomas Yeo^{1,7,8,9}

¹Department of Electrical and Computer Engineering, Clinical Imaging Research Centre, N.1 Institute for Health, National University of Singapore, Singapore; ²Department of Electrical Engineering, Stanford University, Stanford, CA, USA; ³Donders Center for Brain, Cognition and Behavior, Radboud University Nijmegen, Nijmegen, The Netherlands; ⁴Department for Cognitive Neuroscience, Radboud University Medical Center Nijmegen, Nijmegen, The Netherlands; ⁵Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA; ⁶Autism and Social Cognition Center, Child Mind Institute, NY, NY, USA; ⁷Centre for Cognitive Neuroscience, Duke-NUS Medical School, Singapore; ⁸NUS Graduate School for Integrative Sciences and Engineering, Singapore; ⁹Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

*Equal contribution.

Address correspondence to:

B.T. Thomas Yeo
ECE, CIRC, N.1 & MNP
National University of Singapore
Email: thomas.yeo@nus.edu.sg

Running title: ASD factors with distinct RSFC-behavioral patterns

Keywords: ASD Heterogeneity, Behavioral deficits, Resting state functional connectivity, Bayesian modeling, Phenotypes, Default network

Abstract word count: 250

Main text word count (excluding references and figure captions): 3991

Number of figures: 5

Number of tables: 1

Number of supplemental files: 1

Intro: 506 words; **Methods:** 1382 words; **Results:** 977 words; **Discussion:** 1064 words; **Conclusion:** 62 words. **Total:** 3991 words.

Abstract

Background: Heterogeneity in autism spectrum disorder (ASD) has hindered the development of biomarkers, thus motivating subtyping efforts. Most subtyping studies divide ASD individuals into non-overlapping (categorical) subgroups. However, continuous inter-individual variation in ASD suggests the need for a dimensional approach.

Methods: A Bayesian model was employed to decompose resting-state functional connectivity (RSFC) of ASD individuals into multiple abnormal RSFC patterns, i.e., categorical subtypes henceforth referred to as “factors”. Importantly, the model allowed each individual to express one or more factors to varying degrees (dimensional subtyping). The model was applied to 306 ASD individuals (age 5.2-57 years) from two multisite repositories. Posthoc analyses associated factors with symptoms and demographics.

Results: Analyses yielded three factors with dissociable whole-brain hypo/hyper RSFC patterns. Most participants expressed multiple (categorical) factors, suggestive of a mosaic of subtypes within individuals. All factors shared abnormal RSFC involving the default network, but the directionality (hypo/hyper RSFC) differed across factors. Factor 1 was associated with core ASD symptoms, while factor 2 was associated with comorbid symptoms. Older males preferentially expressed factor 3. Factors were robust across multiple control analyses and not associated with IQ, nor head motion.

Conclusions: There exist at least three ASD factors with dissociable patterns of whole-brain RSFC, behaviors and demographics. Heterogeneous default network hypo/hyper RSFC across the factors might explain previously reported inconsistencies. The factors differentiated between core ASD and comorbid symptoms - a less appreciated domain of heterogeneity in ASD. These factors are co-expressed in ASD individuals with different degrees, thus reconciling categorical and dimensional perspectives of ASD heterogeneity.

Introduction

A major challenge in developing biomarkers for Autism Spectrum Disorder (ASD) is the high heterogeneity among ASD individuals. This encompasses core ASD symptoms (1), cognitive skills (2), comorbid psychiatric/medical conditions (3), brain atypicalities (4,5), and genetics (6). Consequently, there have been significant efforts in defining ASD subtypes. Most studies have focused on variability of behavioral or cognitive characteristics (7-10). Studies focusing on brain features are emerging (11-14). Here, we propose a Bayesian framework to decompose whole-brain resting-state functional connectivity (RSFC) patterns in ASD individuals into multiple hypo/hyper RSFC patterns, which we will refer to as “factors” (Figure 1A). This approach allows an individual to express one or more factors (categorical subtypes) to varying degrees (continuous), thus potentially reconciling dimensional (13-15) and categorical (11,12,17) models of ASD heterogeneity.

This approach is motivated by two important considerations. First, most previous ASD subtyping studies assumed that each participant belonged to a single (categorical) subtype. By contrast, the term “spectrum” in ASD suggests continuous variation across individuals (18). This is observed at varying degrees across multiple symptom domains (19,20). In parallel, evidence from genetics and neurobiology suggests that autism results from the combination of multiple factors underlying distinct pathways (21,22). Thus, ASD inter-individual variability may reflect different degrees of expression of such factors and related mechanisms (6,23). Together, these observations motivate a mosaic approach to ASD subtyping that incorporates categorical and dimensional features (Figure 1A). Our model allows each ASD individual to express more than one latent factor. For example, the hypo/hyper RSFC pattern of an ASD individual might be explained by 90% factor 1 and 10% factor 2, while the hypo/hyper RSFC pattern of another ASD individual might be explained by 40% factor 1 and 60% factor 2.

Second, early resting-state fMRI (rs-fMRI) investigations supported models of ASD as a dysconnection syndrome (23-26). Although these early studies focused on *a priori* regions/networks of interest in small to moderately sized samples, these observations have been extended by more recent whole-brain investigations of larger samples showing that multiple functional networks subserving the wide range of processes impaired in ASD are affected (26). Importantly, recent studies have reconciled previously inconsistent findings of either hypo- or hyper-connectivity in ASD by showing that both patterns co-exist, though affecting distinct

functional circuits (27-29). Nevertheless, these studies rely on traditional case-control analyses, which may miss out on less frequently expressed RSFC patterns due to ASD heterogeneity and/or sampling biases. Thus, our study seeks to provide detailed characterization of the nature and spatial extent of functional dysconnections in ASD, while accounting for significant heterogeneity among ASD individuals.

To address these challenges and estimate latent ASD factors with distinct patterns of whole brain hypo/hyper connectivity, we combined two multisite rs-fMRI data repositories (Autism Brain Imaging Data Exchange-second release; ABIDE-II (31) and the Gender Explorations of Neurogenetics and Development to Advance Autism Research; GENDAAR) (32)). Posthoc analyses were performed to examine common and distinct abnormal RSFC across factors. Furthermore, associations between the latent factors and multiple phenotypic information were examined using multivariate analyses to capture the complexity of ASD.

Methods and Materials

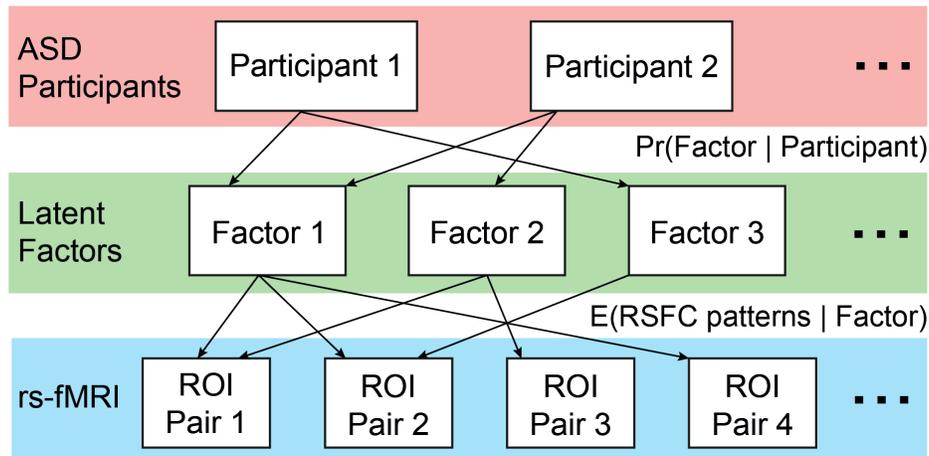
Overview

Our analyses proceeded in four steps (Figure 1B). First, to identify latent ASD factors, we applied a Bayesian model (Figure 1A) to a combined dataset comprising ABIDE-II (31) and GENDAAR (32). We used this combined dataset to maximize sample size with regards to both MRI and non-brain-imaging phenotypic data. Second, we examined the associations between latent factors and ASD participants' phenotypes (i.e., demographic and behavioral symptoms) in the ABIDE-II+GENDAAR combined sample. Third, multiple control analyses were performed to ensure robustness of the results. Lastly, we utilized another independent dataset (ABIDE-I (28)) to explore the drawbacks of case-control analyses, which do not account for ASD heterogeneity.

Participants

MRI data selected from the ABIDE and GENDAAR repositories were analyzed. Following preprocessing and quality control (see "MRI preprocessing" and Supplemental Methods), the resulting sample comprised 242 ASD and 276 neurotypical (NT) participants from ABIDE-II, which were combined with 64 ASD and 72 NT participants from GENDAAR for primary analyses, as well as an independent sample of 166 ASD and 150 NT participants from ABIDE-I (secondary analyses). Age, sex and head motion were matched between ASD and NT participants within each site. Participants' characteristics are summarized in Tables 1 and S1.

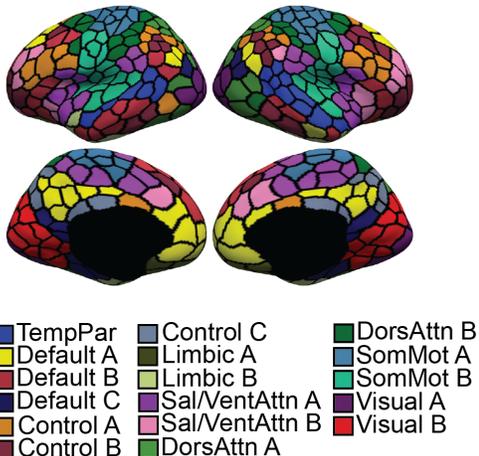
(A)



(B)

Analysis	Sample	N
Identification of latent ASD factors	ABIDE-II + GENDAAR	$N_{ASD} = 306, N_{NT} = 348$
Correlating latent ASD factors with ASD participants' phenotypes	ABIDE-II + GENDAAR	N_{ASD} ranges from 76 to 306, depending on phenotypes (see Methods)
Control analyses	ABIDE-II + GENDAAR	$N_{ASD} = 306, N_{NT} = 348$
Relevance of latent ASD factors for traditional case-control analyses	ABIDE-I	$N_{ASD} = 166, N_{NT} = 150$

(C)



(D)



Figure 1. Bayesian model and overview of analyses in this study. (A) A Bayesian model of ASD participants, latent factors, and resting-state fMRI (rs-fMRI). The model assumes that each ASD participant expresses one or more latent factors, and each factor is associated with distinct, but possibly overlapping patterns of hypo/hyper functional connectivity. Given the RSFC data of

ASD participants and a predefined number of factors K , the model estimates the probability that a participant expresses a latent factor (i.e., factor composition of the participant or $\Pr(\text{Factor} | \text{Participant})$), as well as the expected patterns of hypo/hyper RSFC associated with each factor (i.e., factor-specific hypo/hyper RSFC patterns or $E(\text{RSFC patterns} | \text{Factor})$). **(B)** We first identified latent ASD factors using the ABIDE-II+GENDAAR sample. The latent factors were then correlated with ASD participants' phenotypes (i.e., demographics and behavioral symptoms) in the same sample. Because the multiple sites do not have a uniform set of phenotypes, the sample size varied depending on the phenotype. In addition, we performed several control analyses using the ABIDE-II+GENDAAR sample to ensure the robustness of the latent factors. Finally, we explored the relevance of the latent factors on traditional case-control analyses in the ABIDE-I sample. **(C)** 400 cortical parcels (33). Colors are assigned based on 17 networks widely used in the rs-fMRI literature (34). The 17 networks are divided into eight groups (TempPar, Default, Control, Limbic, Salience/Ventral Attention, Dorsal Attention, Somatomotor and Visual). **(D)** 19 subcortical ROIs (35).

Table 1. Characteristics and behavioral data of participants from ABIDE-II and GENDAAR. Fourteen data collections from ABIDE-II and GENDAAR were included in this study: BNI_1 ($N_{\text{ASD}} = 21$, $N_{\text{NT}} = 15$), ETH_1 ($N_{\text{ASD}} = 6$, $N_{\text{NT}} = 21$), GU_1 ($N_{\text{ASD}} = 34$, $N_{\text{NT}} = 26$), IP_1 ($N_{\text{ASD}} = 10$, $N_{\text{NT}} = 5$), IU_1 ($N_{\text{ASD}} = 17$, $N_{\text{NT}} = 15$), KKI_1 ($N_{\text{ASD}} = 31$, $N_{\text{NT}} = 90$), NYU_1 ($N_{\text{ASD}} = 42$, $N_{\text{NT}} = 23$), OHSU_1 ($N_{\text{ASD}} = 23$, $N_{\text{NT}} = 23$), TCD_1 ($N_{\text{ASD}} = 14$, $N_{\text{NT}} = 18$), UCD_1 ($N_{\text{ASD}} = 15$, $N_{\text{NT}} = 9$), UCLA_1 ($N_{\text{ASD}} = 11$, $N_{\text{NT}} = 11$), USM_1 ($N_{\text{ASD}} = 10$, $N_{\text{NT}} = 13$), U_MIA_1 ($N_{\text{ASD}} = 8$, $N_{\text{NT}} = 7$), GENDAAR ($N_{\text{ASD}} = 64$, $N_{\text{NT}} = 72$). ASD and NT participants were compared with either two-sample t-tests (for continuous measures) or chi-squared tests (for categorical measures). All p-values that survived false discovery rate (FDR) correction ($q < 0.05$) are indicated in bold.

	ASD (N = 306)	NT (N = 348)	p value (two-tailed)
<i>Demographics</i>			
Age, mean (SD), years	14.94 (8.59)	14.89 (8.20)	0.94
Female, No. (%)	70 (22.88)	94 (27.01)	0.26
Full-scale IQ ^a , mean (SD)	106 (16.90)	115 (13.37)	4.43e-13
<i>Head motion</i>			
Mean FD before censoring, mean (SD)	0.10 (0.06)	0.09 (0.07)	0.13
Mean FD after censoring, mean (SD)	0.05 (0.02)	0.05 (0.01)	0.49
<i>Current medication (by target)^b</i>			
GABA, No. (%)	2 (0.97)	0 (0.00)	0.45
Glutamate, No. (%)	3 (1.45)	9 (4.05)	0.18

Serotonin, No. (%)	22 (10.63)	1 (0.45)	8.12e-06
Epinephrine/Adrenaline, No. (%)	1 (0.48)	0 (0.00)	0.97
Norepinephrine/Noradrenaline, No. (%)	50 (24.15)	3 (1.35)	2.13e-12
Dopamine, No. (%)	43 (20.77)	2 (0.90)	5.58e-11
Others, No. (%)	18 (8.70)	9 (4.05)	0.08
<i>Current medication (by class)^c</i>			
Antipsychotic, No. (%)	8 (3.86)	0 (0.00)	9.30e-3
Antidepressant, No. (%)	21 (10.14)	1 (0.45)	1.49e-5
Stimulant, No. (%)	32 (15.46)	2 (0.90)	6.71e-8
SHA, No. (%)	4 (1.93)	0 (0.00)	0.11
Mood stabilizer, No. (%)	4 (1.93)	0 (0.00)	0.11
Others, No. (%)	27 (13.04)	9 (4.05)	1.50e-3

^aN = 297 ASD and 340 NT participants have full-scale IQ scores.

^bMedication was sorted by the primary neurotransmitter system(s) targeted by the medication currently used by participants, based on the Neuroscience-based Nomenclature (NbN-2 (36,37) ; <http://www.nbn2.com/>). Only 207 ASD and 222 NT participants from ABIDE-II+GENDAAR have medication information. As such, percentages were computed based on 207 ASD and 222 NT participants.

^cOnly 207 ASD and 222 NT participants from ABIDE-II+GENDAAR have medication information. As such, percentages were computed based on 207 ASD and 222 NT participants. *Abbreviations:* GABA, Gamma-Aminobutyric acid; SHA, sedatives/hypnotics/anxiolytics.

MRI preprocessing

The neuroimaging data were processed using a previously published pipeline (38-40). See Supplemental Methods for details. Here, we briefly outline the procedure. Rs-fMRI data underwent standard preprocessing including slice time correction, motion correction and alignment with anatomical T1. Frame-wise displacement (FD) and voxel-wise differentiated signal variance (DVARS) were computed using `fsl_motion_outliers` (41,42). Volumes with FD > 0.2 mm or DVARS > 50 were marked as censored frames, together with one frame before and two frames after. Uncensored data segments lasting fewer than five contiguous volumes were also censored (43). Functional runs with more than 50% censored frames were removed (Table S2).

We regressed out 18 nuisance regressors, consisting of six motion parameters, averaged cerebrospinal ventricular signal, averaged white matter signal, global signal and their temporal derivatives. Censored frames were ignored when regression coefficients were computed. Data were interpolated across censored frames using least squares spectral estimation (44). We chose to regress the global signal because of its effectiveness in removing motion-related and

respiratory artifacts (45,46). Recent work has also suggested that global signal regression (GSR) increases the associations between behavior and RSFC (39). Nevertheless, we performed control analyses using an alternative to GSR (see “Control analyses”). Finally, the data were band-pass filtered ($0.009\text{Hz} \leq f \leq 0.08\text{Hz}$), projected onto FreeSurfer fsaverage6 surface space, smoothed with a 6mm kernel and down-sampled onto FreeSurfer fsaverage5.

Resting-state functional connectivity (RSFC)

We utilized a cortical parcellation (33) comprising 400 cortical regions-of-interest (ROIs; Figure 1C) and a subcortical segmentation (35) comprising 19 subcortical ROIs (Figure 1D). RSFC (Pearson’s correlation) was computed among the average time series of 419 brain ROIs (ignoring censored frames), yielding a 419×419 RSFC matrix for each participant. Age, sex, head motion (mean frame-wise displacement (42)) and site differences were regressed out from all participants’ RSFC data with a general linear model (GLM). Regression coefficients were estimated only from NT participants to retain any ASD-specific interactions with participants’ characteristics (e.g., age). Each RSFC entry (i.e., lower triangular entries since the matrices are symmetric) of the ASD participants was z-normalized with respect to the 348 ABIDE-II+GENDAAR NT participants. A z-score larger (or smaller) than zero for a given ROI pair indicates hyper-connectivity (or hypo-connectivity) relative to the NT participants.

Latent factors in ABIDE-II+GENDAAR

Latent ASD factors were identified using the ABIDE-II+GENDAAR dataset. We applied a Bayesian model (Figure 1A) to the z-normalized RSFC of the ASD participants to estimate latent factors. The model is a variant of Bayesian models previously utilized to discover latent atrophy factors in Alzheimer’s Disease (47) and latent components subserving cognitive tasks (48). It assumes that each individual expresses one or more latent factors, associated with distinct, but possibly overlapping patterns of hypo/hyper RSFC. Given the RSFC data and a user-defined number of factors K , we can estimate the factor composition of each participant, i.e., probability that a participant expresses a latent factor (i.e., $\text{Pr}(\text{Factor} \mid \text{Participant})$), as well as the factor-specific hypo/hyper RSFC patterns (i.e., $\text{E}(\text{RSFC patterns} \mid \text{Factor})$). See Supplemental Methods.

We estimated 2, 3 and 4 latent factors. The estimations were robust for 2 and 3 factors, but unstable for 4 factors (Figure S1). Therefore, a larger number of factors was not considered. Furthermore, the two-factor estimates were inconsistent in the control analyses (see “Control analyses”), so we focused on the three-factor estimates in subsequent analyses.

To estimate confidence intervals for the factor-specific hypo/hyper RSFC patterns, we applied a bootstrapping procedure that generated 100 samples from ASD participants’ z-normalized RSFC data. Z-scores were then calculated by dividing factor-specific hypo/hyper RSFC patterns by the bootstrap-estimated standard deviation. To reduce multiple comparisons, the factor-specific hypo/hyper RSFC patterns were averaged across ROI pairs within and between the 17 networks and subcortical structures (Figures 1C-1D), resulting in 18×18 matrices, before computing bootstrapped z-scores. The z-scores were converted to p-values and corrected using false discovery rate ($q < 0.05$) along with other tests (Supplemental Methods).

Associations between participants’ characteristics and latent factors in ABIDE-II+GENDAAR

We applied separate GLM (or logistic regression for binary variables) to the factor compositions and each characteristic (age, sex, FIQ and head motion) of ABIDE-II+GENDAAR ASD participants to investigate potential associations. For each GLM/logistic regression, participants’ characteristic and factor compositions were treated as the dependent and independent variables respectively (Supplemental Methods).

Associations between behavioral symptoms and latent factors in ABIDE-II+GENDAAR

Because ABIDE-II+GENDAAR consisted of datasets across independent sites, not all participants had the same behavioral measures (Table S3). If we considered all available behavioral measures jointly, we would be left with only seven participants. Therefore, available behavioral scores were divided into five groups to maximize the number of participants in each group. For example, Social Responsiveness Scale (SRS) Autistic Mannerism and Repetitive Behaviors Scale-Revised 6 Subscales (RBSR-6) subscales were grouped together because they index aspects of restricted/repetitive behaviors (RRB). Of note, Autism Diagnostic Observation Schedule (ADOS) Stereotyped Behavior subscore was not included into the RRB domain because only 38 participants had ADOS Stereotyped Behavior with SRS Autistic Mannerism, RBSR-6 subscales. The five groups of behavioral scores are shown in Table S4.

We then applied canonical correlation analysis (49) (CCA) between each group of behavioral scores and each factor loading (i.e., $\text{Pr}(\text{Factor} | \text{Participant})$), i.e., fifteen CCAs in total in the case of the three-factor model (see Supplemental Methods). The goal of the CCA was to find an optimal linear combination of the behavioral scores that maximally correlated with the factor loading. Age, sex, head motion and sites were regressed out from both behavioral scores and factor loadings before the CCA. Statistical significance was tested using 10,000 permutations that accounted for different sites. False discovery rate ($q < 0.05$) was utilized to correct for multiple comparisons (Supplemental Methods).

Control analyses in ABIDE-II+GENDAAR

We performed several control analyses to ensure robustness of results. First, to ensure robustness to preprocessing strategies, we applied the Bayesian model to rs-fMRI processed using CompCor (50) instead of GSR. Second, we applied k-means clustering to ABIDE-II+GENDAAR ASD participants' z-normalized RSFC data (processed with GSR or CompCor) to ensure robustness to analysis strategies (k-means versus Bayesian model). Third, we compared behavioral associations of the k-means clusters with those of the latent factors. Lastly, we randomly split the 306 ASD participants in ABIDE-II+GENDAAR into two groups (Table S5) and estimated the latent factors in each group independently. See Supplemental Methods for details.

Drawbacks of traditional case-control analyses in ABIDE-I

In most studies, ASD and NT individuals are compared without accounting for ASD heterogeneity. To explore the drawbacks of case-control analyses, we considered 166 ASD and 150 NT participants from ABIDE-I (Table S1) and inferred their factor compositions using latent factors estimated from ABIDE-II+GENDAAR. These compositions were used to assign each individual to one of three subgroups. This sub-grouping violates the spirit of our hybrid dimensional-categorical approach, but is necessary for comparison with traditional case-control analysis. To ensure robustness, we experimented with two different criteria of assigning ASD participants to subgroups (Tables S6-S7). RSFC differences between each ASD subgroup and demographically-matched NT participants were computed and compared with traditional case-

control analysis (i.e., RSFC differences between ASD and NT participants without subgrouping).
See Supplemental Methods for details.

Results

Latent ASD factors with dissociable hypo/hyper RSFC patterns

We applied the Bayesian model (Figure 1A) to 306 ABIDE-II+GENDAAR ASD participants. An important model parameter is the number of latent factors K . We experimented with $K = 2, 3$ and 4. Four-factor model was unstable (Figure S1), so we did not explore more factors. On the other hand, the two-factor model was sensitive to the preprocessing strategy (Supplemental Results; Table S8). Thus, we focused on the three-factor solution.

Each of the three factor-specific hypo/hyper RSFC patterns among the 400 cortical and 19 subcortical ROIs (Figures 1C-1D) are shown in Figure 2A (unthresholded) and Figure 2B (statistically significant). Figure 2C illustrates the significant RSFC patterns averaged within and between the 17 networks and subcortical structures.

Factor 1 was associated with ASD-related hypo-connectivity (blue in Figure 2) within and between perceptual/motor networks (somatomotor A/B, visual A/B, salience/ventral attention A, dorsal attention A/B). On the other hand, there was ASD-related hyper-connectivity (red in Figure 2) between perceptual/motor and association networks (default, control and salience/ventral attention B), as well as between somatomotor and subcortical regions (caudate and thalamus).

Factor 2 was associated with a pattern of hypo/hyper RSFC almost opposite to factor 1 ($r = -0.57$), but with subtle deviations. For example, regions within default networks A and B were strongly hyper-connected in factor 2, but only weakly hypo-connected in factor 1. Similarly, regions between somatomotor networks and caudate were strongly hyper-connected in factor 1, but did not exhibit any atypical RSFC in factor 2.

Factor 3 was characterized by a complex pattern of hypo/hyper RSFC. For example, there was hyper-connectivity between visual and somatomotor networks. There was also strong hypo-connectivity among regions within default networks A/B, and among regions within the visual networks.

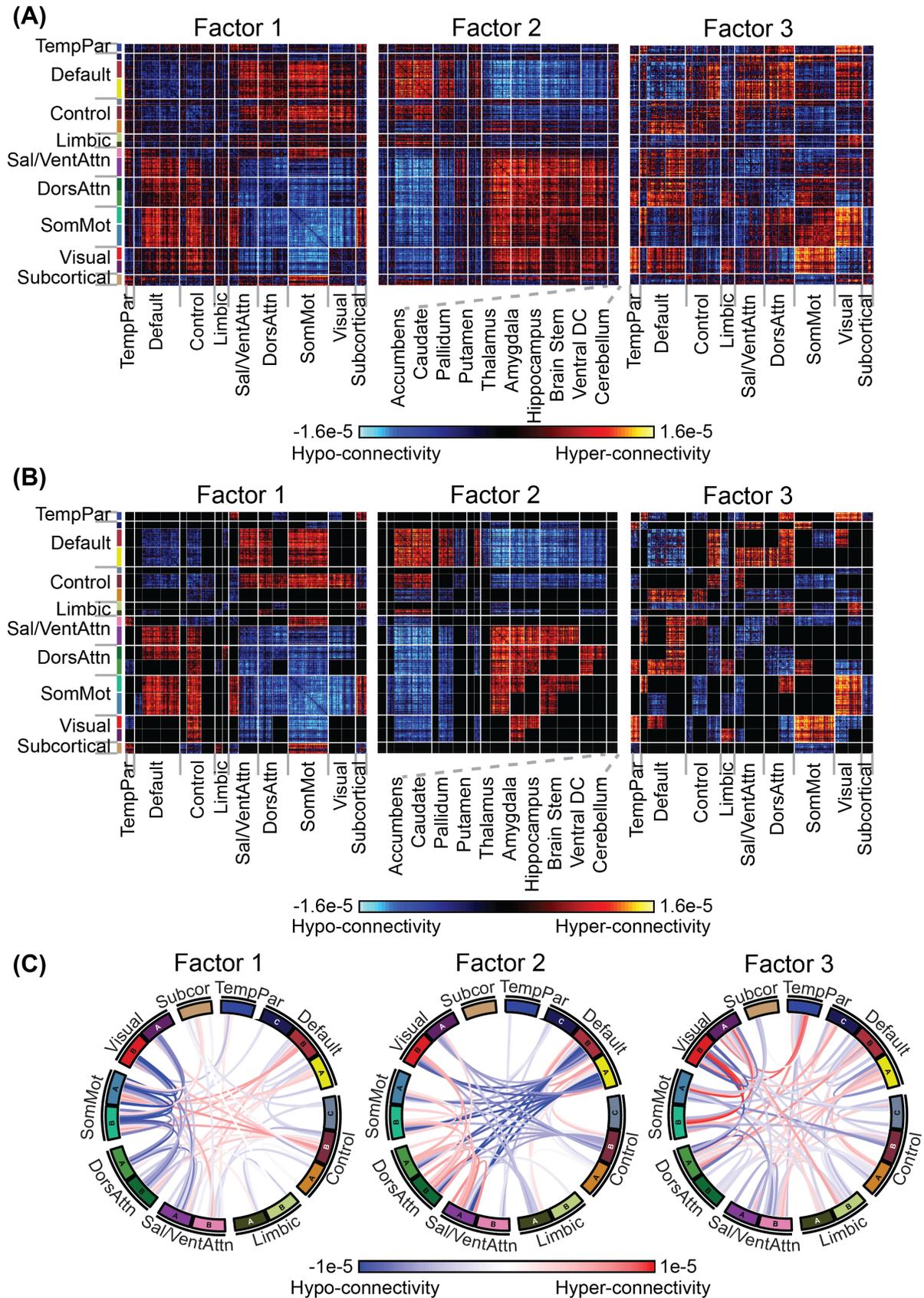


Figure 2. Three latent ASD factors estimated from ABIDE-II+GENDAAR ASD participants. (A) Patterns of hypo/hyper RSFC (unthresholded) associated with each factor. Hot color indicates hyper-connectivity (relative to NT), and cold color indicates hypo-connectivity (relative to NT). (B) Statistically significant patterns of hypo/hyper RSFC associated with each factor. (C) Significant patterns of hypo/hyper RSFC associated with each factor, averaged within and between networks.

Factor compositions of ASD participants in ABIDE-II+GENDAAR

Figure 3 shows the factor compositions of ASD participants in ABIDE-II+GENDAAR. Most participants expressed multiple latent factors rather than a single factor. Critically, no single site showed predominantly one single factor, suggesting that latent factors were not driven by site differences.

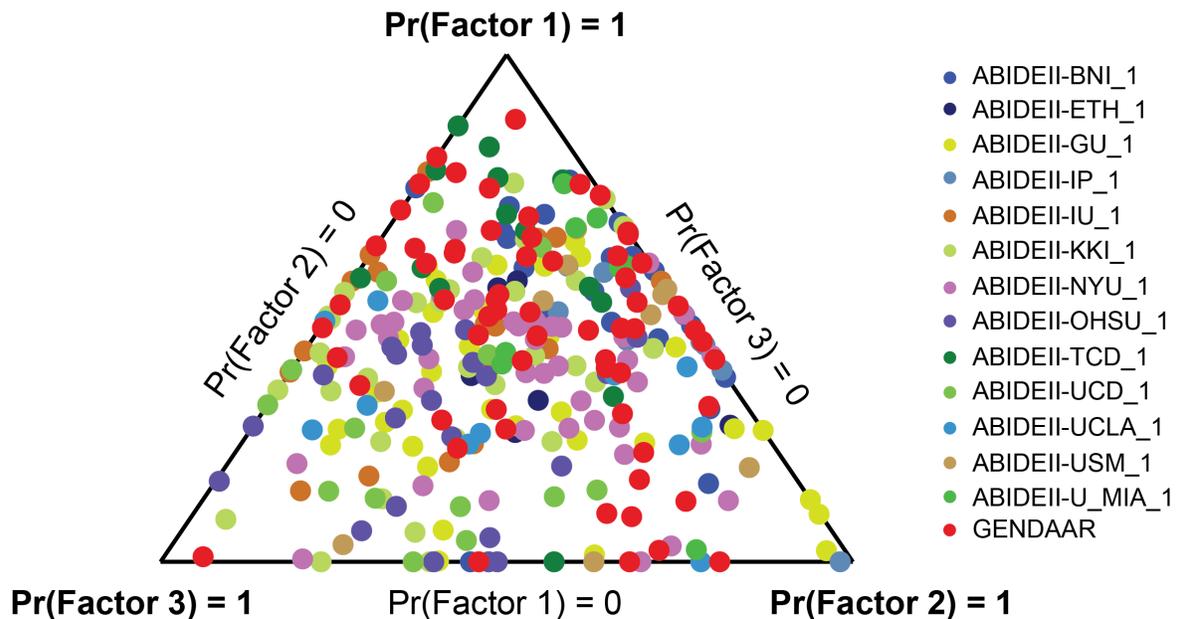


Figure 3. Factor compositions of ASD participants in ABIDE-II+GENDAAR. Each participant corresponds to a dot, with location (in barycentric coordinates) representing the factor composition (i.e., $\text{Pr}(\text{Factor} | \text{Participant})$). Corners of the triangle represent pure factors, and dots closer to the corner indicates higher probability of the respective factor. Most dots are far from the corners, suggesting that most ASD participants expressed multiple factors.

Default network exhibits abnormal connectivity across all three factors

By comparing the RSFC patterns among the three factors (Figure 2B), Figure S2 shows the statistically significant hypo/hyper RSFC unique to each latent factor.

On the other hand, to examine hypo/hyper RSFC patterns that are shared across factors, within-network and between-network blocks with significant bootstrapped z-scores (Figure 2C) were binarized (ignoring directionality of abnormality) and summed across the three factors (Figure 4A). In addition, absolute values of hypo/hyper RSFC patterns that were significant across all three factors (Figure 2B) were summed to obtain the magnitude of hypo/hyper RSFC patterns common across factors (Figure 4B). Altered connectivity within default A and B networks were notable, as well as between default and perceptual/motor networks (somatomotor A, salience/ventral attention A, dorsal attention B). In addition, hypo/hyper connectivity within salience/ventral attention A, within dorsal attention, as well as between somatomotor and control B networks were also common across all three factors.

Lastly, Figure 4C shows the strength of involvement of each ROI obtained by summing the rows of Figure 4B. The strength of default network's involvement was particularly striking. Although atypical default network connectivity was present across all factors, the directionalities were inconsistent. For example, factor 2 exhibited hyper-connectivity within the default network, while factor 1 and 3 exhibited hypo-connectivity (Figure 2B).

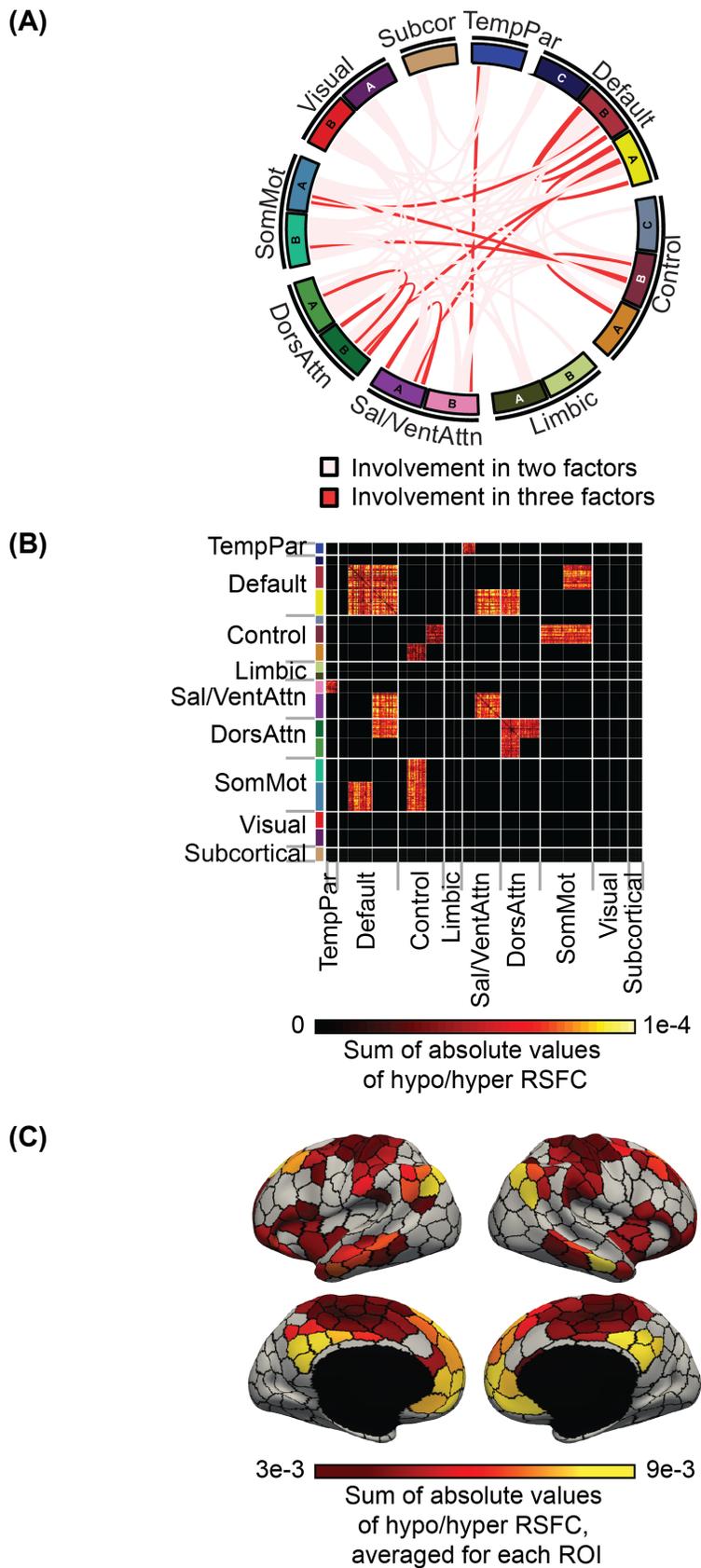


Figure 4. Patterns of hypo/hyper RSFC involved in all three factors. (A) Statistically significant within- and between-network hypo/hyper RSFC patterns (Figure 2C) were binarized, and then summed across the three factors; hypo/hyper RSFC patterns that were significant in only one factor were set to zero. (B) Sum of absolute values of hypo/hyper RSFC (Figure 2B) that were significant across all three factors. (C) The summed absolute values of hypo/hyper RSFC from panel (B) were averaged across the rows for each ROI, and projected onto a surface map for visualization.

Participants' characteristics across latent factors in ABIDE-II+GENDAAR

We used GLM (or logistic regression) to investigate whether ASD participants' characteristics (i.e., age, sex, FIQ, head motion) varied across factors in ABIDE-II+GENDAAR. Factor 3 was more strongly expressed by males relative to factors 2 ($p < 0.001$; Figure S3A). Factor 3 was also associated with older participants compared to factors 1 and 2 ($p = 0.002$ and $p = 0.01$ respectively; Figure S3B). There was no difference in FIQ nor head motion across factors (Figure S3C and S3D).

Associations between latent factors and behavioral symptoms

To examine associations between latent factors and behavioral symptoms in ABIDE-II+GENDAAR ASD participants, we performed CCA between each factor loading and each group of behavioral scores (see Methods). Four sets of CCA analyses remained significant after FDR correction ($q < 0.05$; Figure 5). Higher scores indicated worse symptoms. Therefore, positive values implied that a higher loading on the factor was associated with greater impairment. Factor 1 was associated with both worse RRB (Figure 5A; $r = 0.54$, $p = 0.002$) and social deficits (Figure 5B; $r = 0.27$, $p = 0.004$). Factor 2 was associated with both worse externalizing problems (Figure 5C; $r = 0.43$, $p = 0.004$) and executive dysfunction (Figure 5D; $r = 0.33$, $p = 0.02$). Factor 3 was not associated with any aggregated behavioral symptoms.

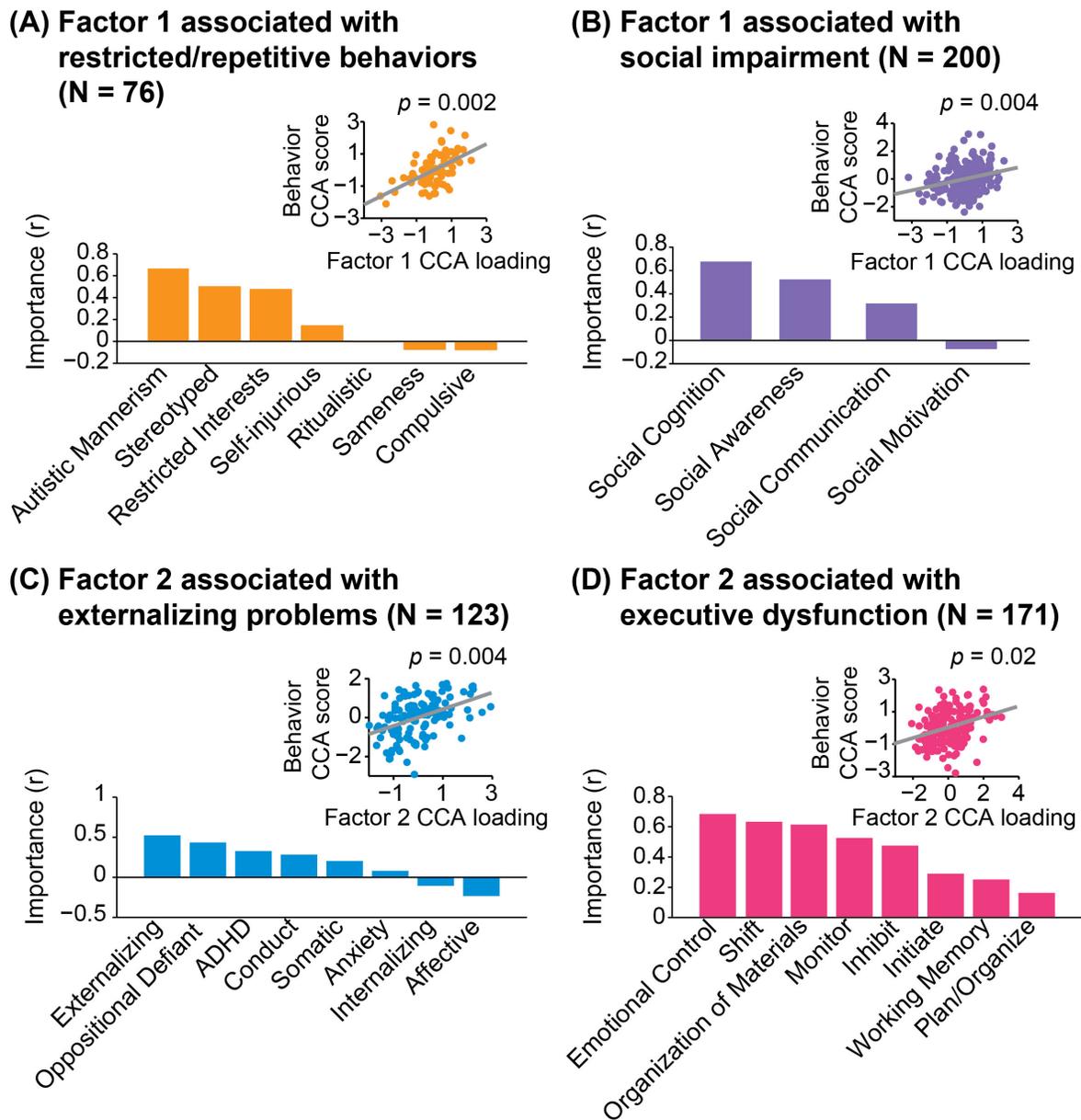


Figure 5. First and second latent ASD factors were associated with distinct behavioral deficits. CCA analyses between the loadings of each factor and five groups of behavioral scores revealed that RSFC factors 1 and 2 were associated with distinct behavioral deficits. Four sets of CCA analyses remained significant after FDR ($q < 0.05$) multiple comparisons correction. **(A)** Associations between factor 1 and RRB measured by SRS Autistic Mannerism subscale and RBSR-6 subscales. **(B)** Associations between factor 1 and social responsiveness measured by SRS subscales (excluding SRS Autistic Mannerism). **(C)** Associations between factor 2 and comorbid psychopathology measured by CBCL-6-18 subscales. **(D)** Associations between factor 2 and executive function measured by BRIEF subscales. The bar plots show the Pearson's correlation between each behavioral score and CCA behavioral loading (more details in Supplemental Methods). Positive correlation suggests that a higher loading on the factor was

associated with greater impairment. The scatterplots show the relationship between CCA behavioral score loading and CCA ASD factor loading, where each dot represents an ASD participant. Factor 1 was associated with worse RRB and social deficits, while factor 2 was associated with worse externalizing problems and executive dysfunction.

Abbreviations: SRS, Social Responsiveness Scale; RBSR-6, Repetitive Behaviors Scale-Revised 6 Subscales; CBCL-6-18, Child Behavior Checklist Ages 6-18; BRIEF, Behavior Rating Inventory of Executive Function.

See Table S4 for behavioral scores entered into the CCA.

Control analyses

Here, we summarize the results of the control analyses (see Supplemental Results for more details). First, factors 1 and 2 from the three-factor model were similar regardless of processing pipeline (GSR or CompCor), but not so for factor 3 (Table S8). On the other hand, all three factors (estimated in primary analyses using GSR; Figure 2) were similar to clusters obtained from k-means regardless of GSR or CompCor (Table S9). Thus, overall, factors 1 and 2, and to some extent factor 3, were relatively robust to preprocessing and analysis strategies.

Compared to the latent factors, k-means clusters showed similar but weaker behavioral associations (Figure S4), suggesting potential advantage of our hybrid dimensional-categorical model. Finally, factors estimated from random splits of ABIDE-II+GENDAAR ASD participants were similar to the factors generated in primary analyses and with each other (Table S10).

Traditional case-control analysis yields smaller effects and misses significant ASD-related RSFC associations

To explore potential drawbacks of traditional case-control analyses, we computed RSFC differences between 166 ASD and 150 demographically-matched NT participants from ABIDE-I. We also computed RSFC differences between ASD and NT participants from each ASD subgroup in ABIDE-I (see Methods). Despite the larger sample size, traditional case-control analysis yielded significantly weaker RSFC differences than the subgroup analyses and missed out on ASD-related RSFC differences (Figures S5-S6). See Supplemental Results for details.

Discussion

In this study, we applied a Bayesian model to a large rs-fMRI cohort of individuals with ASD, revealing three latent factors with dissociable patterns of hypo/hyper RSFC. Each factor was expressed to different degrees across individuals and associated with distinct behavioral and demographic features that are known sources of clinical heterogeneity in ASD, i.e., core ASD impairments, externalizing symptoms, executive dysfunctions, age and sex. Overall, these results suggest that each ASD individual expresses a mosaic of latent factors. This has been missed in prior subtyping approaches that assigned each individual to a single subtype (7-12,51). By contrast, our approach allows each individual's factor composition to be unique, thus retaining inter-individual variability. This is consistent with models suggesting that ASD heterogeneity reflects the contribution of multiple mechanisms with different degrees of expression across subjects (6,23).

Factor 1 loading was the highest when averaged across ASD participants, so unsurprisingly, the hypo/hyper RSFC pattern of factor 1 was the most similar to prior case-control whole-brain comparisons emerging from distinct analytical approaches applied to partially overlapping samples from ABIDE-I (27-29). ASD-related hypo-connectivity within sensory and salience networks, as well as hyper-connectivity between somatomotor and subcortical regions were particularly notable. Likely reflecting its greater prevalence among ASD participants, this factor was more strongly associated with the core ASD symptoms of both social reciprocity and RRB (20). Reflecting their wide range of ASD symptoms, factor 1 involved multiple functional networks previously associated with ASD severity. For example, social skills impairment and hypo-RSFC within default or salience networks have been previously reported (52,53). Although not often explored, hyper-connectivity between thalamus and temporal cortex have also been associated with social reciprocity deficits in ASD (30), and an atypical RSFC balance in cortical striatal circuitry involving limbic somatomotor and frontoparietal networks have been associated with RRB indexed by RBS-R total scores (54).

Interestingly, the factors did not differentiate among core ASD symptoms (e.g., RRB versus social reciprocity), but differentiated between core ASD symptoms (factor 1) and comorbid symptoms (executive dysfunction and externalizing symptoms; factor 2). This finding underscore two important aspects of the ASD phenomenology. One is the strong intercorrelation of ASD symptom domains (20,55) and possibly their partially overlapping biological

underpinnings. While neuroimaging studies with deeper phenotyping may be able to characterize the association between the RSFC patterns in a given factor and specific symptom subdomains, larger scale phenomena reflected in the mosaic of atypical RSFC in a given factor are at play. This is suggested by recent reports of atypical hierarchical network organization in ASD relative to controls, which was also related to global metrics of ASD traits (56,57). The second aspect of the ASD phenomenology underscored by our brain-behavioral findings is that comorbidity critically contributes to ASD heterogeneity. This is a clinical dimension that has been quantitatively and qualitatively overlooked in the neuroimaging literature, albeit with notable exceptions (e.g., 56,57). Our findings suggest that comorbidity should be more consistently accounted for in ASD biomarker efforts.

In contrast with factor 1, factor 2 was characterized by a pattern of hypo/hyper RSFC that included, but was not limited to, hyper-connectivity within the default and attentional networks, as well as hypo-connectivity between the default and attentional networks. As previously mentioned, factor 2 was also associated with executive dysfunction and externalizing symptoms. These findings are consistent with reports that poor executive control might result from abnormalities in attentional networks (60). Greater connectivity within default network has also been linked to worse executive function in ASD (53). Furthermore, externalizing symptoms are frequently observed in large proportions of ASD individuals (59-61). Finally, it remains unknown to what extent factor 2, as well as the other factors and their relationship with behavior are ASD specific or extend across diagnoses. An initial study (13) has suggested that shared factors across ADHD and ASD might exist, but the lack of a shared behavioral battery between the ADHD and ASD samples limited further exploration with specific symptom domains. Another study has shown that distinct clusters of RSFC patterns exist across individuals with and without ASD, and the subtypes revealed unique brain-behavior relations (14). The emergence and availability of deeply-phenotyped transdiagnostic samples (64) will help address this gap.

In our study, factors 1 and 2 were associated with similar age, but distinct behavioral deficits, suggesting that they might not simply reflect disease severity or neurodevelopment stage. On the other hand, factor 3 was more frequently expressed in older participants, and might thus reflect a neurodevelopmental stage. The prevalence of ASD is much higher in males than females (65). Recent studies have suggested that ASD-related sex differences are associated with atypical brain connectivity and mentalization (65-67). Our study suggests that factor 3 was

associated with male participants, but given the small percentage of female participants (23%) in ABIDE-II+GENDAAR, replication with larger sex-balanced datasets is necessary.

Finally, along with factor-specific RSFC signatures, all three factors shared atypical RSFC involving the default network. This is consistent with the larger RSFC ASD literature that have consistently reported structural, and functional ASD-related abnormalities involving the default network (68-70). The directionality of the abnormal default network RSFC varied by factor, e.g., factor 2 exhibited hyper-connectivity within the default network, while factor 1 and 3 exhibited hypo-connectivity (Figure 2B). Similarly, factor 2 exhibited hypo-connectivity between the default and attentional networks, but factors 1 and 3 exhibited hyper-connectivity. The differences between factors may explain some of the inconsistent reports in prior studies on the nature of the DN atypicalities in ASD (50,71-73). As shown by our results from additional analyses on the independent ABIDE-I sample, common case-control comparisons fail to appreciate heterogeneous ASD-related RSFC abnormalities, overall urging for future quantitative ASD subtyping efforts.

A major limitation of this study is the lack of uniform deep phenotyping across all ASD participants because the datasets were gathered post-hoc. Therefore, despite the large sample size, the behavioral associations were separated into five sets of analyses performed on overlapping subsets of ASD participants. A harmonized dataset would also allow the identification of latent factors from RSFC and behavioral deficits simultaneously using a multi-modal variant of the current approach (75). Second, since medication information is limited (e.g., only 8 ASD participants in ABIDEII+GENDAAR were reported to be taking antipsychotic medications), we did not explore the association between factors and medications. Finally, longitudinal data is needed to clarify whether the latent factors reflect ASD heterogeneity or neurodevelopmental stages or complex interactions between heterogeneity and neurodevelopment.

Conclusion

Our study revealed three latent ASD factors with dissociable whole-brain hypo/hyper RSFC patterns. The factors were associated with distinct behavioral symptoms (core ASD versus comorbid symptoms) and demographics. Our approach allows each individual to express multiple latent factors to varying degrees, rather than a single factor. Therefore, each individual's factor composition is unique, which might be potentially useful for future biomarker development.

Disclosures and Acknowledgments

The authors report no competing interests. This manuscript has been posted on bioRxiv.

This work was supported by Singapore MOE Tier 2 (MOE2014-T2-2-016), NUS Strategic Research (DPRT/944/09/14), NUS SOM Aspiration Fund (R185000271720), Singapore NMRC (CBRG/0088/2015), NUS YIA and the Singapore National Research Foundation (NRF) Fellowship (Class of 2017) and partially supported by NIMH R01MH105506-01. Our research also utilized resources provided by the Center for Functional Neuroimaging Technologies, P41EB015896 and instruments supported by 1S10RR023401, 1S10RR019307, and 1S10RR023043 from the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital. Our computational work was partially performed on resources of the National Supercomputing Centre, Singapore (<https://www.nscg.sg>). We thank the investigators and sites who contributed to the ABIDE repositories. Details on the funding sources of each site can be found at http://fcon_1000.projects.nitrc.org/indi/abide/. We also thank the GENDAAR Consortium for their efforts in data collection and sharing. The consortium includes, in alphabetical order, Elizabeth H. Aylward, Raphael A. Bernier, Susan Y. Bookheimer, Mirella Dapretto, Nadine Gaab, Daniel H. Geschwind, Andrei Irimia, Allison Jack, Charles A. Nelson, Kevin A. Pelphrey, Matthew W. State, John D. Van Horn, Pamela Ventola and Sara J. Webb. Data from GENDAAR Consortium were obtained from the NIH-supported National Database for Autism Research (NIMH Data Repositories Study No. 2021; NIMH grant R01 MH100028; Principal Investigator: Kevin A. Pelphrey). National Database for Autism Research (NDAR; <https://ndar.nih.gov/>) is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in autism. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the submitters submitting original data to NDAR.

References

1. Masi A, DeMayo MM, Glozier N, Guastella AJ. An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neurosci Bull.* 2017;33(2):183–93.
2. Lombardo M V, Lai MC, Auyeung B, Holt RJ, Allison C, Smith P, et al. Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Sci Rep.* 2016;6(April):1–15.
3. Matson JL, Williams LW. Differential diagnosis and comorbidity: distinguishing autism from other mental health issues. *Neuropsychiatry (London).* 2013;3(2):233–43.
4. Lenroot RK, Yeung PK. Heterogeneity within Autism Spectrum Disorders: What have We Learned from Neuroimaging Studies? *Front Hum Neurosci.* 2013;7(October):1–16.
5. Floris DL, Lai MC, Auer T, Lombardo M V., Ecker C, Chakrabarti B, et al. Atypically rightward cerebral asymmetry in male adults with autism stratifies individuals with and without language delay. *Hum Brain Mapp.* 2016;37(1):230–53.
6. Betancur C. Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. *Brain Res.* 2011;1380:42–77.
7. Fountain C, Winter AS, Bearman PS. Six Developmental Trajectories Characterize Children With Autism. *Pediatrics.* 2012;129(5):e1112–20.
8. Georgiades S, Szatmari P, Boyle M, Hanna S, Duku E, Zwaigenbaum L, et al. Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. *J Child Psychol Psychiatry Allied Discip.* 2013;54(2):206–15.
9. Munson J, Dawson G, Sterling L, Beauchaine T, Zhou A, Koehler E, et al. Evidence for Latent Classes of IQ in Young Children With Autism Spectrum Disorder. *Am J Ment Retard.* 2008;113(6):439–52.
10. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2016;1(5):433–47.
11. Hrdlicka M, Dudova I, Beranova I, Lisy J, Belsan T, Neuwirth J, et al. Subtypes of autism by cluster analysis based on structural MRI data. *Eur Child Adolesc Psychiatry.* 2005;14(3):138–44.
12. Hong S-J, Valk SL, Di Martino A, Milham MP, Bernhardt BC. Multidimensional

- Neuroanatomical Subtyping of Autism Spectrum Disorder. *Cereb Cortex*. 2017;(November):1–11.
13. Kernbach J, Satterthwaite T, Bassett D, Smallwood J, Margulies D. Shared Endo-phenotypes of Default Mode Dysfunction in Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder. *Transl Psychiatry*. 2018;
 14. Easson AK, Fatima Z, McIntosh AR. Functional connectivity-based subtypes of individuals with and without autism spectrum disorder. *Netw Neurosci*. 2019;3(2):344–62.
 15. Constantino JN. The quantitative nature of autistic social impairment. *Pediatr Res*. 2011;69(5 PART 2):55–62.
 16. Elton A, Di Martino A, Hazlett HC, Gao W. Neural Connectivity Evidence for a Categorical-Dimensional Hybrid Model of Autism Spectrum Disorder. *Biol Psychiatry*. 2016;80(2):120–8.
 17. Sacco R, Lenti C, Sacconi M, Curatolo P, Manzi B, Bravaccio C, et al. Cluster analysis of autistic patients based on principal pathogenetic components. *Autism Res*. 2012;5(2):137–47.
 18. Lombardo M V, Lai M-C, Baron-Cohen S. Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol Psychiatry*. 2019;(September):1487–90.
 19. Grzadzinski R, Huerta M, Lord C. DSM-5 and autism spectrum disorders (ASDs): An opportunity for identifying ASD subtypes. *Mol Autism*. 2013;4(1):2–7.
 20. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (5th ed.)*. 2013;
 21. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*. 2011;474(7351):380–6.
 22. Lombardo M V., Courchesne E, Lewis NE, Pramparo T. Hierarchical cortical transcriptome disorganization in autism. *Mol Autism*. 2017;8(1):1–17.
 23. Jeste SS, Geschwind DH. Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat Rev Neurol*. 2014;10(2):74–81.
 24. Geschwind DH, Levitt P. Autism spectrum disorders: developmental disconnection syndromes. *Curr Opin Neurobiol*. 2007;17(1):103–11.
 25. Minshew NJ, Williams DL. The new neurobiology of autism: Cortex, connectivity, and

- neuronal organization. *Arch Neurol.* 2007;64(7):945–50.
26. Picci G, Gotts SJ, Scherf KS. A theoretical rut: revisiting and critically evaluating the generalized under/over-connectivity hypothesis of autism. *Dev Sci.* 2016;19(4):524–49.
 27. Hutsler JJ, Casanova MF. Review: Cortical construction in autism spectrum disorder: Columns, connectivity and the subplate. *Neuropathol Appl Neurobiol.* 2016;42(2):115–34.
 28. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry.* 2014;19(6):659–67.
 29. Cheng W, Rolls ET, Gu H, Zhang J, Feng J. Autism: Reduced connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. *Brain.* 2015;138(5):1382–93.
 30. Cerliani L, Mennes M, Thomas RM, Di Martino A, Thioux M, Keysers C. Increased functional connectivity between subcortical and cortical resting-state networks in Autism spectrum disorder. *JAMA Psychiatry.* 2015;72(8):767–77.
 31. Di Martino A, O'Connor D, Chen B, Alaerts K, Anderson JS, Assaf M, et al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci Data.* 2017;4:1–15.
 32. Irimia A, Torgerson CM, Jacokes ZJ, Van Horn JD. The connectomes of males and females with autism spectrum disorder have significantly different white matter connectivity densities. *Sci Rep.* 2017;7(March):1–10.
 33. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex.* 2018;28(9):3095–114.
 34. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol.* 2011;106:1125–65.
 35. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron.* 2002;33(3):341–55.
 36. Zohar J, Nutt DJ, Kupfer DJ, Moller H-J, Yamawaki S, Spedding M, et al. A proposal for an updated neuropsychopharmacological nomenclature. *Eur Neuropsychopharmacol J Eur*

- Coll Neuropsychopharmacol. 2014 Jul;24(7):1005–14.
37. Zohar J, Stahl S, Moller H-J, Blier P, Kupfer D, Yamawaki S, et al. A review of the current nomenclature for psychotropic agents and an introduction to the Neuroscience-based Nomenclature. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol*. 2015 Dec;25(12):2318–25.
 38. Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, et al. Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cereb Cortex*. 2019;29(6):2533–2551.
 39. Li J, Kong R, Liégeois R, Orban C, Tan Y, Sun N, et al. Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuroimage*. 2019;196(February):126–41.
 40. Kebets V, Holmes AJ, Orban C, Tang S, Li J. Somatosensory-Motor Dysconnectivity Spans Multiple Transdiagnostic Dimensions of Psychopathology. *Biol Psychiatry*. 2019;1–28.
 41. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23(SUPPL. 1):208–19.
 42. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002;17(2):825–41.
 43. Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb Cortex*. 2016;26(1):288–303.
 44. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*. 2014;84:320–41.
 45. Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, et al. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage*. 2017;154(March):174–87.
 46. Power JD, Plitt M, Gotts SJ, Kundu P, Voon V, Bandettini PA, et al. Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases

- in multiecho data. *Proc Natl Acad Sci.* 2018;201720985.
47. Zhang X, Mormino EC, Sun N, Sperling RA, Sabuncu MR, Yeo BTT. Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. *Proc Natl Acad Sci.* 2016;113(42):E6535–44.
 48. Thomas Yeo BT, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, et al. Functional specialization and flexibility in human association cortex. *Cereb Cortex.* 2015;25(10):3654–72.
 49. Hotelling H. Relations between two sets of variates. *Biometrika.* 1936;28:312–77.
 50. Chai XJ, Castañán AN, Öngür D, Whitfield-Gabrieli S. Anticorrelations in resting state networks without global signal regression. *Neuroimage.* 2012;59(2):1420–8.
 51. Feczko E, Balba NM, Miranda-Dominguez O, Cordova M, Karalunas SL, Irwin L, et al. Subtyping cognitive profiles in Autism Spectrum Disorder using a Functional Random Forest algorithm. *Neuroimage.* 2018;172(April 2017):674–88.
 52. Assaf M, Jagannathan K, Calhoun VD, Miller L, Stevens MC, Sahl R, et al. Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *Neuroimage.* 2010;53(1):247–56.
 53. Abbott AE, Nair A, Keown CL, Datko M, Jahedi A, Fishman I, et al. Patterns of Atypical Functional Connectivity and Behavioral Links in Autism Differ between Default, Salience, and Executive Networks. *Cereb Cortex.* 2016;26(10):4034–45.
 54. Abbott AE, Linke AC, Nair A, Jahedi A, Alba LA, Keown CL, et al. Repetitive behaviors in autism are linked to imbalance of corticostriatal connectivity: A functional connectivity MRI study. *Soc Cogn Affect Neurosci.* 2018;13(1):32–42.
 55. Frazier TW, Ratliff KR, Gruber C, Zhang Y, Law PA, Constantino JN. Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism.* 2014;18(1):31–44.
 56. Hong SJ, de Wael RV, Bethlehem RAI, Larivière S, Paquola C, Valk SL, et al. Atypical functional connectome hierarchy in autism. *Nat Commun.* 2019;10(1):1–13.
 57. Martínez K, Martínez-García M, Marcos-Vidal L, Janssen J, Castellanos FX, Pretus C, et al. Sensory-to-Cognitive Systems Integration Is Associated With Clinical Severity in Autism Spectrum Disorder. *J Am Acad Child Adolesc Psychiatry.* 2019.
 58. Wan B, Wang Z, Jung M, Lu Y, He H, Chen Q, et al. Effects of the Co-occurrence of

- Anxiety and Attention-Deficit/Hyperactivity Disorder on Intrinsic Functional Network Centrality among Children with Autism Spectrum Disorder. *Autism Res.* 2019;(May):1057–68.
59. Di Martino A, Zuo XN, Kelly C, Grzadzinski R, Mennes M, Schvarcz A, et al. Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder. *Biol Psychiatry.* 2013;74(8):623–32.
 60. Keehn B, Müller RA, Townsend J. Atypical attentional networks and the emergence of autism. *Neurosci Biobehav Rev.* 2013;37(2):164–83.
 61. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. *J Am Acad Child Adolesc Psychiatry.* 2008;47(8):921–9.
 62. Panagiotidi M, Overton PG, Stafford T. Co-Occurrence of ASD and ADHD Traits in an Adult Population. *J Atten Disord.* 2017;
 63. Joshi G, Wozniak J, Petty C, Martelon MK, Fried R, Bolfek A, et al. Psychiatric comorbidity and functioning in a clinically referred population of adults with autism spectrum disorders: A comparative study. *J Autism Dev Disord.* 2013;43(6):1314–25.
 64. Alexander R, Langdon PE, Chester V, Barnoux M, Gunaratna I, Hoare S. Heterogeneity within autism spectrum disorder in forensic mental health: the introduction of typologies. *Adv Autism.* 2016;2(4):201–9.
 65. Fombonne E. The changing epidemiology of autism. *J Appl Res Intellect Disabil.* 2005;18(4):281–94.
 66. Ypma RJF, Moseley RL, Holt RJ, Rughooputh N, Floris DL, Chura LR, et al. Default Mode Hypoconnectivity Underlies a Sex-Related Autism Spectrum. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2016;1(4):364–71.
 67. Alaerts K, Swinnen SP, Wenderoth N. Sex differences in autism: A resting-state fMRI investigation of functional brain connectivity in males and females. *Soc Cogn Affect Neurosci.* 2016;11(6):1002–16.
 68. Floris DL, Lai MC, Nath T, Milham MP, Di Martino A. Network-specific sex differentiation of intrinsic brain function in males with autism. *Mol Autism.* 2018;9(1):1–12.

69. Moseley RL, Ypma RJF, Holt RJ, Floris D, Chura LR, Spencer MD, et al. Whole-brain functional hypoconnectivity as an endophenotype of autism in adolescents. *NeuroImage Clin.* 2015;9:140–52.
70. Glerean E, Pan RK, Salmi J, Kujala R, Lahnakoski JM, Roine U, et al. Reorganization of functionally connected brain subnetworks in high-functioning autism. *Hum Brain Mapp.* 2016;37(3):1066–79.
71. Padmanabhan A, Lynch CJ, Schaer M, Menon V. The Default Mode Network in Autism. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2017;2(6):476–86.
72. Weng S-J, Wiggins JL, Peltier SJ, Carrasco M, Risi S, Lord C, et al. Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders. *Brain Res.* 2010;1313(May):202–14.
73. Yerys BE, Gordon EM, Abrams DN, Satterthwaite TD, Weinblatt R, Jankowski KF, et al. Default mode network segregation and social deficits in autism spectrum disorder: Evidence from non-medicated children DMN in children with ASD. *NeuroImage Clin.* 2015;9:223–32.
74. Uddin LQ, Supekar K, Lynch CJ, Khouzam A, Phillips J, Feinstein C, et al. Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry.* 2013;70(8):869–79.
75. Sun N, Mormino EC, Chen J, Sabuncu MR, Thomas BT. Multi-modal Latent Factor Exploration of Atrophy , Cognitive and Tau Heterogeneity in Typical Late-Onset Alzheimer ' s Disease. *bioRxiv.* 2018;