Widespread Associations between Behavioral Metrics and Brain Microstructure in ASD Suggest Age Mediates Subtypes of ASD

Abbreviated Title: *Microstructural Associations with Behavior in ASD*

Haylee J. Ressa[1]*, Benjamin T. Newman[1,2]*, Zachary Jacokes[3], James C. McPartland[4], Natalia M. Kleinhans[5], T. Jason Druzgal[2], Kevin A. Pelphrey[2,1], & John Darrell Van Horn[1,3*]

For the GENDAAR Research Consortium

[1]Department of Psychology, University of Virginia, Gilmer Hall, Charlottesville, VA 22903

[2] Department of Neurology, University of Virginia, School of Medicine, Gilmer Hall, Charlottesville, VA 2290

[3]School of Data Science, University of Virginia, Elson Building, Charlottesville, VA 22903

[4]Yale Child Study Center and the Yale Center for Brain and Mind Health, Yale School of Medicine, Sterling Hall of Medicine, New Haven, CT 06520

[5]University of Washington Integrated Brain Imaging Center, Eunice Kennedy Shriver Intellectual and Developmental Disabilities Research Center, Seattle, WA 98195

*These authors are co-first authors of this manuscript

[#]Email: jdv7g@virginia.edu

Words in Abstract: 239; **Words in Body:** 6,641; **Figures:** 7; **Tables:** 1

Conflicts

James C. McPartland consults with Customer Value Partners, Bridgebio, Determined Health, Apple, and BlackThorn Therapeutics, has received research funding from Janssen Research and Development, serves on the Scientific Advisory Boards of Pastorus and Modern Clinics, and receives royalties from Guilford Press, Lambert, Oxford, and Springer.

Acknowledgements

This work was performed on behalf of the GENDAAR Consortium (NIH R01 MH100028), and we thank all of our collaborating colleagues, the study participants, and their family members. We would like to specifically acknowledge the contributions of Anna Kresse, MPH; Megha

Santhosh, MHA; Désirée Lussier-Lévesque, PhD; Emily Neuhaus, PhD; Katy Ankenman, MSW; Jessica Benton, MA; and Rachel Fung, BS. The grant sponsors had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract:

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social communication and repetitive behaviors. Our lab has previously found that g-ratio, the proportion of axon width to myelin diameter, and axonal conduction velocity, which is associated with the capacity of an axon to carry information, are both decreased in ASD individuals. By associating these differences with performance on cognitive and behavioral tests, this study aims to first associate a broad array of behavioral metrics with neuroimaging markers of ASD, and to explore the prevalence of ASD subtypes using a neuroimaging driven perspective. Analyzing 273 participants (148 with ASD) ages 8 to 17 through an NIH-sponsored Autism Centers of Excellence network (MH100028), we observe widespread associations between behavioral and cognitive evaluations of autism and between behavioral and microstructural metrics, alongside different directional correlations between different behavioral metrics. Stronger associations with individual subcategories from each test rather than summary scores suggest that different neuronal profiles may be masked by composite test scores. Machine learning cluster analyses applied to neuroimaging data reinforce the association between neuroimaging and behavioral metrics and suggest that age-related maturation of brain metrics may drive changes in ASD behavior. This suggests that if ASD can be definitively subtyped, these subtypes may show different behavioral trajectories across the developmental period. Clustering identified a pattern of restrictive and repetitive behavior in some participants and a second group that was defined by high sensory sensitivity and language performance.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with a prevalence of approximately 1 in 36 children (Maenner et al., 2023). In a clinical setting, a diagnosis of ASD is made by a clinician according to criteria established by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) that describe persistent social or communication deficits, repetitive behaviors or interests, and atypical response to sensory information (*Diagnostic and Statistical Manual of Mental Disorders*, 2013). Recently, to counteract variability in clinical decision-making, there has been an increased emphasis on assessing performance on one or more behavioral and cognitive tests (Constantino & Charman, 2016). Several tests have been developed and normed, but the relationship between different test metrics and the neural structure underlying ASD is still unclear. The highly heterogeneous nature of behaviors in ASD suggests that individuals exist on a spectrum from neurotypical to those with high support needs due to severe autism. Concurrently research may suggest that there is not a singular form of autism but rather multiple forms of "autisms" (Masi et al., 2017). Large datasets have played a critical role in advancing autism research by offering a comprehensive means to study variability across diverse aspects, including genetic factors, brain structure, and behavioral traits (Lombardo et al., 2019).

ASD and the Brain: Research highlighting altered functional connectivity in individuals with ASD has given rise to the underconnectivity theory, suggesting that individuals with ASD experience reduced functional connectivity in both frontal and posterior brain regions, attributed to limited cortical bandwidth (Just et al., 2012). However recent studies have challenged this theory by demonstrating that patterns of atypical neural connectivity in autism cannot be consistently explained by distance or connectivity (Picci et al., 2016). Furthermore, idiosyncratic connectivity patterns in which individuals with ASD show greater inter-individual variability have been observed (Hong et al., 2017). Underlying cellular differences, such as changes in neuronal microstructure, may contribute to these connectivity variations. Diffusion MRI provides a means to analyze axonal characteristics like diameter and anisotropy at a sub-voxel level (Afzali et al., 2021). Deficiencies in axonal structure can impair the speed or efficiency of action potentials. The thickness of the myelin sheath surrounding white matter in the cortex is associated with the inner diameter of axons, which influences conduction velocity (Liewald et

al., 2014). Additionally, effective long-range signaling depends on axons with larger diameters to support signal transmission across greater distances (Das & Gilbert, 1995).

Diffusion Microstructure and ASD: Structural differences in the axons of ASD participants have been observed in previous diffusion studies. In particular, ASD individuals exhibited abnormal white matter microstructural patterns in the splenium of the corpus callosum (Zhao et al., 2022). Findings suggest ASD participants have lower fractional anisotropy and higher diffusivity in white matter tracts associated with behaviors commonly disrupted in ASD individuals. Fixel-based analysis of white matter tracts demonstrated that ASD individuals have lower fiber density in the splenium, corresponding with greater social impairments (Dimond et al., 2019). Alternatively, neurite orientation dispersion and density imagining (NODDI) has found that ASD participants had higher extracellular free-water levels and lower neurite density. These differences were mainly present in long-range association tracts that guide ASD behaviors (Andica et al., 2021).

Conduction Velocity, G-ratio, and Extracellular Water: A recent paper (Newman et al., 2024) demonstrated architectural differences in the axonal microstructure of ASD participants. G-ratio is a proportion of axon to myelin diameter with an optimal range of around 0.6-0.7 (Rushton, 1951). As axon and myelination thickness play a role in signal transduction speed and efficiency, g-ratio is related to conduction velocity, a measure associated with an axon's capacity to carry information sensitive to myelin and axonal development differences. This study found that in ASD participants, extracellular water, g-ratio, and conduction velocity, was altered in the brains of ASD participants compared to non-ASD participants. These microstructural differences were present throughout the cortex, subcortex, and white matter skeleton. Decreases in conduction velocity and g-ratio, in particular, suggests deficits in long-range connections that rely on larger axon and myelination diameters (Newman et al., 2024).

To continue this work, we leveraged a uniquely rich dataset on a cohort of autistic and non-autistic individuals from the NIH-sponsored (MH100028) Autism Centers of Excellence (ACE) cohort. This cohort represents a very large and carefully evaluated group of age, sex, and diagnosis-matched individuals that enabled us to explore the relationship between brain microstructure and behavioral patterns across a spectrum of development, ranging from typical development to moderately autistic profiles. By focusing on microstructural features, we aimed to uncover how variability in brain organization relates to the diverse symptom presentations

observed in ASD. One of the central goals of this work is to better understand the neural correlates of ASD subtypes. Rather than attempting to divide autism into rigid subcategories, we adopted a spectrum-based view, recognizing that microstructural features and their behavioral correlates may not align neatly with traditional diagnostic boundaries. Instead, these features may represent dimensions of variability that are shared across typical and atypical development, offering a more nuanced framework for interpreting heterogeneity in autism. Ultimately, this study seeks to address the critical challenge of ASD variability by identifying brain-based markers that could predict symptom patterns and inform future classifications of autism. This approach holds the potential to bridge the gap between brain and behavior, advancing our understanding of ASD heterogeneity and paving the way for more personalized interventions and predictive biomarkers.

Methods:

Participants:

Two-hundred seventy-three (mean age = 154.3 months ±35.21 S.D., age range = 96-216 months; 133 female [49%]) participants from Wave 1 of an NIH-sponsored Autism Centers of Excellence Network were included in this study. The study cohort included 148 individuals diagnosed with ASD (mean age = 150.8 months ±34.31 S.D., 70 female [47%]) and 124 neurotypical participants (mean age = 154.3 months ±35.21 S.D., 62 female [50%]). Data used in this study is available from the NIH National Database for Autism Research (NDAR).

Behavioral and Cognitive Assessments:

Participants were administered the battery of behavioral and cognitive tests summarized in Table 1 at each participating site (Harvard University, Yale University, Seattle Children's Hospital with the University of Washington, and the University of California, Los Angeles). All participants within the ASD cohort were validated via the administration of the ADI-R and ADOS, including separate social and behavioral subcomponents, by a clinician. Participants family members were also asked about age of language acquisition. The social, behavioral, overall, and composite ADOS scores were included for ASD participants only. Participants not meeting ASD criteria in the ASD cohort were excluded from the study. Participants in the final ASD cohort were paired using age and sex-matched non-autistic participants. The behavioral tests utilized in this study are as follows:

The Clinical Evaluation of Language Fundamentals—Fourth Edition (CELF-4) is a test given to participants ages 5 through 21 to evaluate language ability through expressive and receptive language-based subtests that measure phonology, morphology, syntax, semantics, and working memory. CELF-4 is administered by a professional, typically a speech pathologist. Language skills vary greatly among children with ASD (Salem et al., 2021). While children with ASD tend to have impairments in both expressive and receptive language skills, deficits in receptive language are usually more significant (Mody et al., 2012).

The Behavior Rating Inventory of Executive Functions (BRIEF) is a behavior rating scale used to screen for executive function deficits in children ages 5 to 18. Parents and teachers complete the BRIEF questionnaire and ask how the child behaves in everyday situations, particularly those that require problem-solving. Children with ASD have been found to have significantly elevated BRIEF scores compared to neurotypical children, corresponding to executive function ability deficits (Blijd-Hoogewys et al., 2014).

The Autism Diagnostic Observation Schedule (ADOS), used by many clinicians as the gold standard of ASD diagnosis, is a semi-structured assessment of social communication, social interaction, and repetitive behaviors in individuals suspected to have autism spectrum disorder (ASD). It is administered by trained clinicians who select the appropriate module based on the individual's developmental and language level, with metrics targeting areas such as quality of social responses, reciprocity, and restricted or repetitive behaviors (Kamp-Becker et al., 2018).

Autism Diagnostic Interview-Revised (ADI-R), a caregiver interview conducted by clinicians, assess behaviors associated with autism. ADI-R measures social communication, restricted and repetitive behaviors, and developmental milestones, providing information on symptoms through detailed, structured questions (Snow et al., 2009).

The Repetitive Behavior Scale-Revised (RBS-R) evaluates an array of restrictive, repetitive behaviors (RRBs) an individual with ASD may exhibit. RRBs, a common presentation of ASD, are divided into subscales of stereotypic behaviors, self-injurious behaviors, compulsions, ritualized behaviors, insistence on sameness, and restricted interests (Lam & Aman, 2007). Caregivers self-report the RBS-R questionnaire. As RRBs are a hallmark of ASD, children with ASD are expected to score higher on subscales of the RBS-R.

The Adolescent/Adult Sensory Profile (AASP) measures sensory processing and individual sensory preferences that lead to behaviors in participants 11 and older. AASP is a self-

reported questionnaire that scores components such as sensory sensitivity and avoidance for the five senses. Children with ASD have been found to have abnormal sensory processing behaviors that match in intensity but differ in processing patterns from neurotypical children (Crane et al., 2009).

The Child Behavior Checklist (CBCL) component of the Achenbach System of Empirically Based Assessment assesses a range of behavioral and emotional syndromes, including anxiety, depression, aggression, and defiant behavior problems. The syndromes are grouped into internally and externally focused behaviors and emotions. The CBCL questionnaire is administered to parents of children ages 6 to 18. Children with ASD have been found to have higher scores on CBCL subscales for depression, social problems, thought problems, and attention problems compared to neurotypical children (Arias et al., 2022).

The Social Responsiveness Scale version 2 (SRS-II) is a rating scale measuring behavioral associated with ASD and can be completed by raters with at least 1 month of experience with the rated individual. Different rating forms are available for age groups, including a self-report form for individuals aged 19 and up. The SRS-2 focuses on social differences and each item is responded with a 4-point Likert scale rating(Bruni, 2014; Constantino & Gruber, 2012).

The Vineland Adaptive Behavior Scales – II (Vineland-II) is utilized for the assessment of social and adaptive functions, also termed social competency. Vineland-II is sometimes used as a substitute for traditional intelligence testing in situations where the participant's verbal ability is inadequate or there is active psychopathology due to the shortness and ease of administration(Doll, 1935).

The Differential Ability Scales (DAS) is a cognitive battery designed to test a range of abilities with a narrower and more specific domain than general intelligence tests. The DAS is intended to provide a profile of specific cognitive strengths and weaknesses to educators in order to tailor interventions. The various subtests are altered depending on participant age and ability and subdivisions cover verbal and nonverbal domains(Elliott et al., 1990, 2007).

[Insert Table 1 Here]

Statistical Analysis:

All behavioral and cognitive test results were compared to the mean microstructural values in each ROI using general linear models. All models featured sex-as-assigned at birth, participant age, full scale IQ, scanner/evaluation site, and intracranial volume as control terms. All resulting aggregate g-ratio and conduction velocity p-values were corrected for multiple comparisons using the Benjamini & Hochberg (Benjamini & Hochberg, 1995) method across all 214 ROIs. Test scales were not batch altered or z-scored for linear model testing as several tests are highly bimodally distributed in scoring by design, with non-autistic individuals frequently scoring at or near 0.

Image Acquisition:

Diffusion, T1-weighted, and T2-weighted images were acquired from each subject. Diffusion images were acquired with an isotropic voxel size of 2x2x2mm^3, 64 non-colinear gradient directions at b=1000 s/mm^2, and 1 b=0, TR=7300ms, TE=74ms. T1-weighted MPRAGE images with a FOV of 176x256x256 and an isotropic voxel size of 1x1x1mm^3, TE=3.3; T2-weighted images were acquired with a FOV of 128x128x34 with a voxel size of 1.5x1.5x4mm^3, TE=35.

Image Data Processing:

All image data was processed per the protocol described in Newman et al. (Newman et al., 2024) to generate aggregate g-ratio and aggregate conduction velocity maps. In brief, preprocessing was performed following prior work (Newman, Dhollander, et al., 2020); diffusion images were denoised (Veraart et al., 2016), corrected for Gibbs ringing artifacts (Kellner et al., 2016), and corrected for inhomogeneity fields using FSL's *topup* and *eddy* commands utilizing outlier detection and replacement (Andersson et al., 2003, 2016; Andersson & Sotiropoulos, 2016), The final preprocessed diffusion images were up-sampled to an isotropic voxel size of 1.3x1.3x1.3mm^3 (Greenspan, 2009). WM, GM, and CSF tissue response functions were generated using the Dhollander algorithm (Dhollander et al., 2016), and single-shell 3-tissue-constrained spherical deconvolution were used to generate the WM fiber orientation distribution (FODs) and GM and CSF representations. 3-Tissue Constrained Spherical Deconvolution (Dhollander et al., 2017; Kelly et al., 2022; Mito et al., 2018, 2020) was used to calculate the voxel-wise maps of the fraction of signal arising from each of 3 compartments: an intracellular anisotropic, intracellular isotropic, and extracellular isotropic freely diffusing water compartment by setting the sum of all FOD coefficients equal to unity. WM-FODs were then used to create a cohort-specific template

with a subset of 40 individuals counterbalanced between sex and diagnosis (D. Raffelt et al., 2012). All subject's WM-FODs were registered to this template using an affine non-linear transform warp, and then the template was registered to a b-value matched template in stereotaxic MNI space (Hsu et al., 2015; Newman, Untaroiu, et al., 2020). A fixel-based morphometry (FBM) (D. Raffelt et al., 2012; D. A. Raffelt et al., 2017) approach was used to estimate the intra-axonal cross-sectional area within each voxel to be used as an apparent axonal volume fraction (AVF). Each subject's AVF maps were then registered to MNI space using the ANTs SyN nonlinear registration technique by aligning each to the 11-year-old adolescent template developed by Richards et al. Note that a template approximately one standard deviation below the mean age of this study was used to better register the comparatively smaller younger subjects (Richards et al., 2016; Richards & Xie, 2015). T1w and T2w images were processed as described in the MICA-MNI pipeline (Cruces et al., 2022), including N4-bias correction (Avants et al., 2009), rescaling both images from 0-100, co-registration using a rigid transform, and subsequently non-linear ANTs SyN registration to the same Richards et al., template as the diffusion-based images (Avants et al., 2014; Richards et al., 2016). While there are noted shortcomings to using T1w/T2w ratio to measure myelin in white matter regions (Sandrone et al., 2023), the method has also been shown to correlate well with myelin in the cortex (Glasser & Van Essen, 2011; Sandrone et al., 2023). No calibration or adjustments were performed because g-ratio values are generally not well established in adolescents, and there is a desire not to alter or introduce additional error to g-ratio measurements before proceeding to aggregate conduction velocity calculation.

Code Accessibility:

All work was performed using publicly available software, and additional code for calculating g-ratio and conduction velocity from WM FODs and T1w/T2w ratio images can be found here: https://github.com/btn6sb/Conduction_Velocity

Aggregate G-Ratio and Conduction Velocity Calculation:

Both metrics were calculated according to previously published methods (Newman et al., 2024). The aggregate g-ratio was calculated on a voxel-wise basis according to Stikov et al. and was used according to Mohammadi & Callaghan, as displayed in Equation 1(Campbell et al., 2018; Mohammadi & Callaghan, 2021; Stikov et al., 2011, 2015). As a measure of intra-axonal volume, the fiber density cross section was used as the AVF (D. Raffelt et al., 2012). As a metric of myelin density, the T1w/T2w ratio was used as the myelin volume fraction (MVF). These metrics

represent the total sums of each respective compartment across the volume of the voxel and are a volume-based equivalent to the original formulation of g as the ratio of axon diameter (d) to fiber diameter (D).

$$g = \frac{d}{D} = \sqrt{1 - \frac{MVF}{MVF+AVF}}$$

(1)

Aggregate conduction velocity was calculated based on the calculations of Rushton (Rushton, 1951) and Berman et al. (Berman et al., 2019), reiterating Rushton's calculation that conduction velocity (θ) is proportional to the length of each fiber segment (l) and that this is roughly proportional to *D*, which in turn can be defined as the ratio between *d* and the g-ratio (g). Furthering the considerations of Rushton, Berman et al. show that a value proportional to conduction velocity can be calculated using axon diameter and the g-ratio as in equation 2 (Berman et al., 2019):

$$\theta \propto l \propto Dg\sqrt{-\ln(g)} \propto d\sqrt{-\ln(g)}$$

(2)

Aggregate g-ratio and conduction velocity were averaged across 214 ROIs from the JHU-ICBM WM atlas (48 ROIs) (Mori et al., 2005) and the Destrieux Cortical Atlas (164 ROIs) (Destrieux et al., 2010). Additionally, two composite ROIs were included, one of all 48 JHU ROIs and one of 150 neocortical regions from the Destrieux Atlas.

Cluster Analyses:

To further investigate the relationship between the behavioral and cognitive metrics and brain microstructure beyond the ROI level, we employed two separate machine learning cluster analysis techniques in an exploratory manner. These clustering techniques allow for examining the relationship between participants and brain imaging metrics by grouping participants, brain areas, and behavioral and cognitive metrics based on similarity across metrics. In performing these analyses, we aimed to explore whether groups or clusters of participants features exist that describe subtypes or groups within the cohort. We also demonstrated the utility of cluster analysis for

analyzing structural neuroimaging data. We applied Clustering Hierarchy Optimization by Iterative Random Forests (CHOIR) to examine similarities across participants. This clustering analyses was performed using the respective R packages for each method.

CHOIR is a repeated iterative random forest and statistical permutation method based on distinctive features for clustering data. Initially designed for single-cell analysis, where, for example, a cluster might denote a specific biologically distinct cell type or population, here we substitute participants for cells and use the mean aggregate g-ratio and conduction velocity in each ROI as features. CHOIR generates robust clusters through an iterative random permutation testing procedure that merges clusters that fail the prediction testing (Petersen et al., 2024). After the final clusters are generated, we can recover the original identity of participants and observe if the final clusters correspond with demographic, behavioral, or cognitive variables of interest. As the clustering is performed using only neuroimaging metrics, the correspondence of clusters to other variables suggests that particular groupings of neuroimaging results across ROIs are associated with these variables.

Results:

Behavioral Metrics:

Figure 1 presents the distribution of summary scores and individual subtest results across different measures for individuals with ASD (in red) and neurotypical individuals (in blue). These results represent one of the most comprehensive cognitive and behavioral batteries ever assembled for assessing autism spectrum disorder (ASD), including multiple commonly applied child behavior assessments, providing an analysis of the relationships between diverse domains such as sensory processing, executive function, language, and social behaviors. We present data from all subscales assessed as well as summary total scores. Compared to individual subscale scores, the summary scores exhibit narrower interquartile ranges and smaller variances compared to the individual subtest scores, which show a broader spread and higher variability. This contrast is consistent across most measures, with subtest results reflecting a greater range of values within each diagnostic group. On the BRIEF, CBCL, RBS-R, SRS-2, the ASD participants had higher mean scores on each metric compared to the non-ASD participants while on the Vineland-II and CELF the ASD participants had lower mean scores on each metric

compared to the non-ASD participants. The AASP was mixed with the ASD participants having a higher score in 10 metrics and non-ASD participants having a higher mean score in 3 metrics.

The variation in sub-scales that measure unique behavior patterns is potentially masked by summary scores that determine diagnostic criteria. Individual subtest scores demonstrate greater variability compared to the more narrowly distributed summary scores suggesting that the nuances in specific behaviors—such as those measured by the BRIEF, CBCL, RBS-R, SRS-2, Vineland-II, CELF, and AASP—may not be fully represented by aggregate measures often used in diagnostic contexts. Similarly, the mixed results in the AASP indicate that even within a single diagnostic group, there is considerable variability in response to specific sensory-related items. This complexity is potentially lost when only summary scores are used, as they may conceal these behavior-specific variations, resulting in a less specific understanding of the individual's profile.

[Insert Figure 1 Here]

Correlation Analysis:

Results from all tests across all subjects were analyzed using a simple Pearson correlation analysis to observe cross test associations (Fig. 1). There were widespread significant correlations between individual subscales, particularly within tests (for column labels, p-values, and correlation coefficients that compose this figure see Supplementary Table 1). Interestingly there was widespread positive correlations between the CBCL, RBS-R and BRIEF, and a second group of less positive correlations between the AASP and CELF-4 with these two groups significantly negatively correlated from each other, suggesting a split between sensory and language processing and behavioral and social domains.

The Adolescent/Adult Sensory Profile (AASP) as well as the Clinical Evaluation of Language Fundamentals (CELF) displayed strong internal correlations among their own submetrics. However, the CELF did not show significant correlations with any other metrics, including the BRIEF (Behavior Rating Inventory of Executive Function). This lack of correlation suggests that the CELF captures unique information about language abilities that is not aligned with the constructs measured by other assessments. The subscales measuring social metrics, such

as those within the AAS (Autism Spectrum-related Social Metrics), were poorly correlated with other domains, indicating a distinct construct that may not overlap substantially with sensory, executive function, or language domains. Interestingly, the summary scores of the ADOS (Autism Diagnostic Observation Schedule), considered the gold standard for autism diagnosis, did not correlate significantly with any other measures in the analysis, suggesting that the ADOS integrates individuals with several different behavioral profiles that may not share scores on individual domains, further supporting that ASD subtypes may appear in the behavioral results.

[Insert Figure 2 Here]

Linear Modeling:

When examining data from all participants, conduction velocity was significantly associated with 47 different subscales in at least 1 ROI. Subscales with significant associations came from the BRIEF, RBS-R, CELF-4, CBCL, SRS-2, and Vineland-II tests. The BRIEF test, in particular, had significant subscales in many ROIs (Fig. 3). For example, the monitor subscale, which measures an individual's ability to monitor plans, thoughts, and emotions, was significant for conduction velocity in 168 ROIs after multiple comparison corrections. Significant ROIs were located across a wide range of cortical ROIs but were especially prominent in the superior parietal and frontal cortex and subcortical gray matter (Fig. 4).

When considering the same data set, g-ratio was not as widely nor strongly associated across ROIs with behavioral metrics. Only seven different subscales were significantly associated with at least one ROI. Significant subscales for g-ratio were from the DAS-School Age, SRS-2, BRIEF, ADOS, and Language Acquisition tests. The most significant associations were found in the deep WM in the BRIEF and DAS, associated with 19 different ROIs. Across their different subscales, the CELF-4 was positively associated with conduction velocity, while RBS-R was negatively associated with conduction velocity.

[Insert Figure 3 Here]

[Insert Figure 4 Here]

While conduction velocity had a much stronger pattern of significance when considering all the data, when associations were considered exclusively within the autistic participants, g-ratio was more strongly associated with behavioral metrics across several ROIs. In ASD-only data evaluations, conduction velocity was significant for only three subscales from the Vineland-II and BRIEF tests. The composite total metric of Vineland-II, which was significant in 33 ROIs when analyzing data from all participants, was only significant for 1 ROI when considering solely ASD participants. G-ratio, conversely, was significant for nine different subscales derived from the BRIEF, ADOS, CBCL, Language Acquisitions, SRS-2, and Vineland-II tests. In particular, g-ratio was more strongly associated with the CBCL test than conduction velocity, and the total score of the CBCL subscales was significant for g-ratio in 37 ROIs. G-ratio relationships were primarily located in the motor cortex and WM. Subscales of the SRS-2 demonstrated powerful associations across analyses with multiple significant areas in g-ratio tests when considering all participants and only those with ASD.

CHOIR Cluster Analysis:

CHOIR analysis of the 414 total ROIs resulted in a total of 6 distinct clusters that were generally associated with a gradient between higher aggregate g-ratio and lower conduction velocity at one pole (lower microstructural maturity) and lower aggregate g-ratio and higher conduction velocity at the other pole (higher microstructural maturity; Fig. 5). This pattern was broadly similar across all the ROIs used to generate the clusters (Supplementary File 1). Demographically, the clusters did not appear to be predictive of subject sex or ASD diagnosis (Fig. 6). However, there was a gradient for subject age, with clusters 2, 4, and 5 having a higher mean subject age than the overall mean and clusters 1, 3, and 6 having a lower mean subject age than the overall mean. Interestingly, this pattern was slightly different for brain volume, with clusters 1, 2, and 5 having higher mean subject brain volume than the overall mean and clusters 3, 4, and 6 having a lower mean subject brain volume than the overall mean. This allows the division of the clusters into four general patterns based on demographics: a lower maturity, younger cluster with low brain volume (clusters 3 & 6), a higher maturity, older cluster with high

brain volume (clusters 2 & 5), a cluster with mixed microstructural maturity, younger, with high brain volume (cluster 1), and a cluster with mixed microstructural maturity, older, with low brain volume (cluster 4).

When the behavioral and cognitive test results were projected onto the clusters, the pattern generally appeared to follow the age-related poles defined earlier, with subscales having a high/low gradient of scoring matching the younger/older divide for clusters 3 & 6 and 2 & 5. However, within the same test, like the CELF and AASP Sensory Profile, this polarity was switched between different items (Fig. 6). Age, rather than brain volume, appeared to be more indicative of which pole the intermediate clusters (1 & 4) were aligned to. The complete list of all behavioral results projected onto the CHOIR cluster is available in Supplementary File 2.

A final CHOIR analysis was performed on only ASD participants, both clustered using behavioral results, neuroimaging results, and using all the metrics available both behavioral and neuroimaging (Fig. 7). While the first two analyses delineated 3 clusters of ASD subjects, the combined analysis delineated 2 larger clusters. In each case, age was a vital factor that polarized the groups, with separate clusters for younger subjects in particular. In each case, a younger group with high scores on restricted and repetitive behavior was separated from other participants that had lower performance scores on language skills and higher scores on sensory sensitivity (The CHOIR results from all behavioral and neuroimaging metrics are available in Supplementary File 3). This suggests that age polarizes two distinct clusters of ASD participants with common behavioral features. The same pattern is observed when neurotypical participants are included in Fig. 6 where higher CELF language scale scores are paired with higher sensory sensitivity along an age polarized axis.

[Insert Figure 5 Here]

[Insert Figure 6 Here]

[Insert Figure 7 Here]

Discussion

This study identifies multiple regions associated with particular behavioral and cognitive autism evaluations. Several tests, particularly the CELF-4 and BRIEF, showed excellent associations between brain metrics and individual performance. Cluster analysis identified age as a defining axis for neuronal and behavioral metrics, and identified two distinct and inversely correlated groups of individuals, with restrictive and repetitive behavior in some participants and a second group that was defined by high sensory sensitivity and language performance. Despite differences in assessments, there was much overlap in brain regions associated with the various metrics, mainly when non-autistic participants were included. Associations in white matter regions demonstrated robust associations for both analysis of all participants and ASD-specific tests. However, when only autistic individuals were evaluated, aggregate g-ratio was significantly associated with more ROIs than conduction velocity. This switch from more significant conduction velocity measurements when all participants were assessed to more g-ratio associations when evaluating exclusively autistic participants suggests that the observed neurological differences between autistic and non-autistic individuals may differ from the neurological correlates of autism severity. This finding confirms the consensus among the ASD clinical community that behavioral tests need to be evaluated for effectiveness at assessing along separate dimensions of diagnosis and severity and that any single test may not assess both components well (Hus & Lord, 2014; Lord et al., 2012). By finding widespread findings between behavioral and cognitive tests and brain structure associated with ASD, this study suggests that clinical evaluations by experts and parental assessments are both able to predict differences in neurological structure to at least some degree. While the linear models showed that many different assessments can be reflective of neuronal structure, the cluster analyses suggested that the breadth of assessments within many major tests is important to fully capture individuals who may present different behavioral profiles depending on age. While several of the tests examined in this study showed relationships between brain structure and total or composite scores across the internal metrics of the study, only the CBCL total score was significantly associated with

multiple brain regions within the ASD only cohort, suggesting that the notion of severity within ASD may manifest in highly different ways along different behavioral vectors. Interestingly the CBCL has been found to have higher validity within subscales, particularly anxiety and depression, than the total score(Pandolfi et al., 2014). While we do not delineate specific groups or subgroups with this study, these findings do suggest that differential brain regions may be involved in different symptom profiles, and lends support for clinical examination of multiple subscale scores, rather than composite scores alone. Conversely, it is possible that some of the associations between various subscales and brain microstructure is not unique or wholly specific to ASD, and it is possible that these tests are capturing behavior not unique to ASD(Havdahl et al., 2016). .

The unequal distribution of significant results across behavioral and cognitive tests indicates that diagnostic assessments vary in their correlation to brain structure. Prior to the DSM-5 recategorization, the BRIEF assessment was reported to be elevated across subgroups of ASD, including autistic disorder, Asperger's syndrome, and pervasive developmental disorder (Blijd-Hoogewys et al., 2014). While these diagnostic labels have been discontinued, the high number of associations between microstructural metrics and BRIEF subscales may demonstrate that BRIEF captures a breadth of behaviors, some of which may not be specific to autism alone, in a way that aligns with what is occurring structurally in the brain. The SRS-2 assessment displayed multiple associations with g-ratio when analyzed in the context of all participants and the ASD subgroup specifically. The SRS is a valid predictive measure of ASD across time points in childhood, indicating that it captures consistent autistic traits within a child over time (Chan et al., 2017). This continuity across development may indicate an underlying neurological difference present from a young age, and thus, it has the potential to be detected early on. Conversely, the lack of associations between microstructural metrics and widely used diagnostic assessments, such as the ADOS, suggests that there may be a disjunction between behaviors emphasized in diagnostic measures and those that stem from neurological differences shared across ASD individuals.

A striking findings is the lack of significant correlations between the summary scores of the Autism Diagnostic Observation Schedule (ADOS), the gold standard of ASD diagnosis, and any other cognitive or behavioral metrics. This lack of correlation suggests that while the ADOS summary scores effectively capture diagnostic criteria for ASD, they may not reflect the

underlying individual traits or symptoms measured by other assessments. This finding could indicate that, when broken down, the components of the ADOS summary scores may not align well with specific metrics across a heterogeneous population.This observation highlights the profound heterogeneity inherent to ASD. The diagnosis of autism encompasses a broad spectrum of presentations, with individuals exhibiting diverse profiles of strengths and challenges. As such, a diagnosis of "autism" alone may not provide detailed information about a person's specific symptomatology or functional abilities. The variability in correlations observed in this study underscores the importance of examining individual dimensions of behavior and cognition rather than relying solely on diagnostic labels to guide intervention or research.

The observed differences in correlations across behavioral tests may also reflect underlying neurobiological diversity within the ASD population. Tests may tap into distinct neuronal networks or functional systems, leading to variable patterns of association depending on the metric being assessed. For example, language skills measured by the CELF may rely on different brain regions than those associated with social metrics captured by the AAS or sensory metrics. These distinct neural profiles likely contribute to the variability in test outcomes and further reinforce the need for multidimensional approaches to understanding ASD. Pinpointing which behavioral metrics align with brain structure across ranges of symptom severity may better enable diagnostic tools to accurately distinguish ASD children from those that are neurotypical or have other disorders.

The cluster analysis performed here was unique in adapting both methods, previously designed for single-cell analysis (CHOIR) and psychometric latent factor analysis (EGA) to neuroimaging data. CHOIR can cluster subjects incorporating multiple features per subject, while EGA can associate features across subjects, providing different but complementary approaches. While conclusions drawn from this novel approach should be limited and done cautiously, CHOIR, in particular, appeared to replicate many demographic and behavioral differences from exclusively microstructural neuroimaging data that suggests that accounting for multiple ROI measurements may have value in neuroimaging analysis. The CHOIR results instead elegantly showcase associations between brain metrics and age as being essential markers of change relative to sex or even ASD diagnosis.

Overlap between age-related clusters and different behavioral metrics suggest that the behavioral profile of autism may be modulated by age. Other studies have suggested that some

ASD behaviors decrease with age during the end of the developmental period, such as restricted and repetitive behaviors(Esbensen et al., 2009), while other studies have found more complex age-related trajectories of behavioral score increase and decrease(Waizbard-Bartov et al., 2022). The EGA technique correctly replicated correlations between metrics and replicated clusters of imaging metrics despite being provided with these metrics as independent features, for example, the purple cluster at the bottom of the conduction velocity figure represents the left and right olfactory cortex, left and right nucleus accumbens, and left and right subcallosal gyrus. This symmetry between left and right ROIs appearing in the same cluster was common throughout the EGA results. Conduction velocity appeared to have more distinct clustering of brain regions compared to the 3 clusters of g-ratio ROIs. This may be due to increased sensitivity to the development of axonal tracts that cross or connect, and are thus shared between, multiple ROIs.

A previous study found that ASD was associated with decreased aggregate g-ratio and conduction velocity but not a significant difference in T1/T2 ratio, demonstrating that neurological differences in ASD may be based on changes in axonal structure and not simply a deficit of myelination (Newman et al., 2024). These changes may cause the observed switch when evaluating ASD participants solely as differences in axonal architecture manifest uniquely in conduction velocity compared to g-ratio. Subtle differences in inner axonal diameter can significantly affect conduction velocity, while g-ratio relies on the balance between axonal diameter and myelin thickness. Based on our observations, ASD participants with more severe behavioral symptoms may have similarly altered ratios between inner axonal diameter and myelin diameter. In contrast, high-performing ASD participants may have a more neurotypical relationship between diameter and myelin. This range of g-ratios within the ASD cohort may underlie the range of behavioral severity seen across the disorder. As g-ratio had more and stronger significant relationships with behavioral metrics than conduction velocity, combined with the lack of significant associations between T1w/T2w ratio, the critical difference affecting g-ratio in ASD participants may be inner axon diameter as opposed to myelin diameter.

These findings are supported by prior post-mortem electron microscopy studies of the corpus callosum of 12 subjects found significantly decreased axon diameters and cross-sectional areas in autistic subjects (Wegiel et al., 2018). Additionally, a decrease in the percentage of large-diameter axons in all five segments of the corpus callosum was observed in autistic individuals. Prior work with the data used in this study found the same result from

microstructural analysis (Newman et al., 2024). Furthermore, the histological study found that autism had a more significant correlation with axon diameter and area than with myelin thickness, agreeing with the explanation that the axonal diameter element of g-ratio has a predominating effect on autism development compared to myelin diameter. The structural abnormalities in axonal development observed in this study demonstrated deficits in interhemispheric connection specificity, which may underlie dysregulation of velocity and volume of information and issues with information processing. This study reinforces our findings that axonal diameter may inform behavior patterns associated with ASD and provides a basis for how this property may alter long-range connections.

A recent structural MRI study in toddlers and preschoolers found that age-related increases in cortical myelination in TD participants were absent in children with ASD, indicating that myelin trajectory may follow a different timeline in those with ASD (Chen et al., 2022). A significant association was not present between cortical myelin and autism symptoms when analyzing results from ASD participants, aligning with our observations that myelination may not heavily contribute to ASD behaviors as development continues. The disruption of myelination in young children found by this study may reflect that the microstructural abnormalities underlying ASD are not constant until the brain is further developed, explaining the changes in symptom severity that can be present as a child with ASD gets older.

Interestingly, these microstructural and behavioral patterns uncovered by this study do not respect the traditional boundaries of an ASD diagnosis. Instead, they highlight the need to study autism as a spectrum where overlapping features and dimensions provide insights into individual variability. By focusing on the relationships between brain features and behavioral subtypes, research can move toward identifying meaningful subgroups within the spectrum. This approach may provide valuable insights into how unique neurobiological profiles underlie specific behavioral patterns, enabling the development of more targeted interventions. Ultimately, this framework shifts the focus from strict diagnostic categories to understanding ASD as a multidimensional condition defined by shared and distinct traits across individuals.

While other MRI studies have found associations between abnormal or decreased axonal growth and autism, this study correlates these microstructural differences to behavioral severity. It evaluates a wide range of commonly administered ASD evaluations. As the symptomatic behaviors of ASD can span a wide range, determining differences within ASD subgroups is

crucial to understanding the neurobiological factors driving these diverse behaviors. Further research is needed to track the relationships between microstructural metrics and behavioral subscales and investigate if the associations between g-ratio and behavior remain as the individuals in this study continue to develop. In particular, observing the changes in these associations as participants reach and pass pubescent markers will provide evidence that microstructural metrics remain abnormal in the long term. A longitudinal framework will also be crucial to elucidate sex differences in ASD behaviors observed in score distributions. Insight into how microstructural metrics are related to behavior across the lifespan has the potential to serve as the basis of a biomarker of ASD and provide a neurobiological understanding of the disorder. Moreover, having a neurological explanation of behavioral severity in ASD can help inform the effect of psychological treatment and identify individuals at high risk of ASD before behavioral tests in infancy and early childhood can evaluate them.

Works Cited:

Afzali, M., Pieciak, T., Newman, S., Garyfallidis, E., Özarslan, E., Cheng, H., & Jones, D. K. (2021). The sensitivity of diffusion MRI to microstructural properties and experimental factors. *Journal of Neuroscience Methods*, *347*, 108951. https://doi.org/10.1016/j.jneumeth.2020.108951

Andersson, J. L., Graham, M. S., Zsoldos, E., & Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage*, *141*, 556–572.

Andersson, J. L., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *Neuroimage*, *20*(2), 870–888.

Andersson, J. L., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage*, *125*, 1063–1078.

Andica, C., Kamagata, K., Kirino, E., Uchida, W., Irie, R., Murata, S., & Aoki, S. (2021). Neurite orientation dispersion and density imaging reveals white matter microstructural alterations in adults with autism. *Molecular Autism*, *12*(1), Article 1.

Arias, A. A., Rea, M. M., Adler, E. J., Haendel, A. D., & Van Hecke, A. V. (2022). Utilizing the Child Behavior Checklist (CBCL) as an Autism Spectrum Disorder Preliminary Screener and Outcome Measure for the PEERS® Intervention for Autistic Adolescents. *Journal of Autism and Developmental Disorders*, *52*(5), 2061–2074. https://doi.org/10.1007/s10803-021-05103-8

Avants, B. B., Tustison, N. J., Stauffer, M., Song, G., Wu, B., & Gee, J. C. (2014). The Insight ToolKit image registration framework. *Frontiers in Neuroinformatics*, *8*.

Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTS). *Insight j*, *2*(365), 1–35.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Berman, S., Filo, S., & Mezer, A. A. (2019). Modeling conduction delays in the corpus callosum using MRI-measured g-ratio. *Neuroimage*, *195*, 128–139.

Blijd-Hoogewys, E. M. A., Bezemer, M. L., & Van Geert, P. L. C. (2014). Executive Functioning in Children with ASD: An Analysis of the BRIEF. *Journal of Autism and Developmental Disorders*, *44*(12), 3089–3100. https://doi.org/10.1007/s10803-014-2176-9

Bruni, T. P. (2014). Test review: Social responsiveness scale–Second edition (SRS-2). *Journal of Psychoeducational Assessment*, *32*(4), 365–369.

Campbell, J. S., Leppert, I. R., Narayanan, S., Boudreau, M., Duval, T., Cohen-Adad, J., Pike, G. B., & Stikov, N. (2018). Promise and pitfalls of g-ratio estimation with MRI. *Neuroimage*, *182*, 80–96.

Chan, W., Smith, L. E., Hong, J., Greenberg, J. S., & Mailick, M. R. (2017). Validating the social responsiveness scale for adults with autism. *Autism Research*, *10*(10), 1663–1671. https://doi.org/10.1002/aur.1813

Chen, B., Linke, A., Olson, L., Kohli, J., Kinnear, M., Sereno, M., Müller, R., Carper, R., & Fishman, I. (2022). Cortical myelination in toddlers and preschoolers with autism spectrum disorder. *Developmental Neurobiology*, *82*(3), 261–274. https://doi.org/10.1002/dneu.22874

Constantino, J. N., & Charman, T. (2016). Diagnosis of autism spectrum disorder: Reconciling the syndrome, its diverse origins, and variation in expression. *The Lancet Neurology*, *15*(3), Article 3.

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2*.

Crane, L., Goddard, L., & Pring, L. (2009). Sensory processing in adults with autism spectrum disorders. *Autism*, *13*(3), 215–228. https://doi.org/10.1177/1362361309103794

Cruces, R. R., Royer, J., Herholz, P., Larivière, S., De Wael, R. V., Paquola, C., Benkarim, O., Park, B., Degré-Pelletier, J., & Nelson, M. C. (2022). Micapipe: A pipeline for multimodal neuroimaging and connectome analysis. *NeuroImage*, *263*, 119612.

Das, A., & Gilbert, C. D. (1995). Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature*, *375*(6534), Article 6534.

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, *53*(1), 1–15.

Dhollander, T., Raffelt, D., & Connelly, A. (2017). Towards interpretation of 3-tissue constrained spherical deconvolution results in pathology. *Proc. Intl. Soc. Mag. Reson. Med*, *25*, 1815.

Diagnostic and statistical manual of mental disorders: DSM-5 (Fifth edition). (2013). American Psychiatric Association.

Dimond, D., Schuetze, M., Smith, R. E., Dhollander, T., Cho, I., Vinette, S., Ten Eycke, K., Lebel, C., McCrimmon, A., & Dewey, D. (2019). Reduced white matter fiber density in autism spectrum disorder. *Cerebral Cortex*, *29*(4), Article 4.

Doll, E. A. (1935). A genetic scale of social maturity. *American Journal of Orthopsychiatry*, *5*(2), 180.

Elliott, C. D., Murray, G. J., & Pearson, L. S. (1990). Differential ability scales. *San Antonio, Texas*. https://books.google.com/books?hl=en&lr=&id=I6Up5n5vfdEC&oi=fnd&pg=PA65&dq=differential+ability+scales&ots=86s5WAZrmQ&sig=caQcqis59mFqSIbYf4VJNtQu3IA

Elliott, C. D., Salerno, J. D., Dumont, R., & Willis, J. O. (2007). Differential ability scalesSecond edition. *San Antonio, TX*.

Esbensen, A. J., Seltzer, M. M., Lam, K. S., & Bodfish, J. W. (2009). Age-related differences in restricted repetitive behaviors in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 57–66.

Glasser, M. F., & Van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. *Journal of Neuroscience*, *31*(32), 11597–11616.

Greenspan, H. (2009). Super-resolution in medical imaging. *The Computer Journal*, *52*(1), 43–63.

Havdahl, K. A., von Tetzchner, S., Huerta, M., Lord, C., & Bishop, S. L. (2016). Utility of the child behavior checklist as a screener for autism spectrum disorder. *Autism Research*, *9*(1), 33–42.

Hong, A. L., Newman, B. T., Khalid, A., Teter, O. M., Kobe, E. A., Shukurova, M., Shinde, R., Sipzner, D., Pignolo, R. J., & Udupa, J. K. (2017). Feasibility of fabricating personalized 3D-printed bone grafts guided by high-resolution imaging. *Medical Imaging 2017: Imaging Informatics for Healthcare, Research, and Applications*, *10138*, 101380O.

Hsu, Y.-C., Lo, Y.-C., Chen, Y.-J., Wedeen, V. J., & Isaac Tseng, W.-Y. (2015). NTU-DSI-122: A diffusion spectrum imaging template with high anatomical matching to the ICBM-152 space. *Human Brain Mapping*, *36*(9), 3528–3541.

Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: Revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, *44*, 1996–2012.

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., & Varma, S. (2012). Autism as a neural systems disorder: A theory of frontal-posterior underconnectivity. *Neuroscience & Biobehavioral Reviews*, *36*(4), Article 4.

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., Poustka, L., Weber, L., Schmidt, H., Smidt, J., Stehr, T., Roessner, V., Kucharczyk, K., Wolff, N., & Stroth, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child & Adolescent Psychiatry*, *27*(9), 1193–1207. https://doi.org/10.1007/s00787-018-1143-y

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, *76*(5), 1574–1581.

Kelly, C., Dhollander, T., Harding, I. H., Khan, W., Beare, R., Cheong, J. L., Doyle, L. W., Seal, M., Thompson, D. K., & Inder, T. E. (2022). Brain tissue microstructural and free-water composition 13 years after very preterm birth. *Neuroimage*, *254*, 119168.

Lam, K. S. L., & Aman, M. G. (2007). The Repetitive Behavior Scale-Revised: Independent Validation in Individuals with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *37*(5), 855–866. https://doi.org/10.1007/s10803-006-0213-z

Liewald, D., Miller, R., Logothetis, N., Wagner, H.-J., & Schüz, A. (2014). Distribution of axon diameters in cortical white matter: An electron-microscopic study on three human brains and a macaque. *Biological Cybernetics*, *108*, 541–557.

Lombardo, M. V., Lai, M.-C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular Psychiatry*, *24*(10), 1435–1450. https://doi.org/10.1038/s41380-018-0321-0

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M., Ousley, O., Guy, L., Bernier, R., & Gerdts, J. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry*, *69*(3), 306–313.

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … Shaw, K. A. (2023). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *MMWR. Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neuroscience Bulletin*, *33*(2), 183–193. https://doi.org/10.1007/s12264-017-0100-y

Mito, R., Dhollander, T., Raffelt, D., Xia, Y., Salvado, O., Brodtmann, A., Rowe, C., Villemagne, V., & Connelly, A. (2018). Investigating microstructural heterogeneity of white matter hyperintensities in Alzheimer's disease using single-shell 3-tissue constrained spherical deconvolution. *Proc. Intl. Soc. Mag. Reson. Med*, 135.

Mito, R., Dhollander, T., Xia, Y., Raffelt, D., Salvado, O., Churilov, L., Rowe, C. C., Brodtmann, A., Villemagne, V. L., & Connelly, A. (2020). In vivo microstructural heterogeneity of white matter lesions in healthy elderly and Alzheimer's disease participants using tissue compositional analysis of diffusion MRI data. *NeuroImage: Clinical*, *28*, 102479.

Mody, M., MGH/HST Athinoula A. Martinos Center for Biomedical Imaging, Harvard Medical School, Department of Radiology, Charlestown, MA, Belliveau, J. W., & MGH/HST Athinoula A. Martinos Center for Biomedical Imaging, Harvard Medical School, Department of Radiology, Charlestown, MA. (2012). Speech and Language Impairments in Autism: Insights from Behavior and Neuroimaging. *American Chinese Journal of Medicine and Science*, *5*(3), 157. https://doi.org/10.7156/v5i3p157

Mohammadi, S., & Callaghan, M. F. (2021). Towards in vivo g-ratio mapping using MRI: Unifying myelin and diffusion imaging. *Journal of Neuroscience Methods*, *348*, 108990.

Mori, S., Wakana, S., Van Zijl, P. C., & Nagae-Poetscher, L. M. (2005). *MRI atlas of human white matter*. Elsevier.

Newman, B. T., Dhollander, T., Reynier, K. A., Panzer, M. B., & Druzgal, T. J. (2020). Test–retest reliability and long-term stability of three-tissue constrained spherical deconvolution methods for analyzing diffusion MRI data. *Magnetic Resonance in Medicine*, *84*(4), 2161–2173. https://doi.org/10.1002/mrm.28242

Newman, B. T., Jacokes, Z., Venkadesh, S., Webb, S. J., Kleinhans, N. M., McPartland, J. C., Druzgal, T. J., Pelphrey, K. A., Van Horn, J. D., & Consortium, G. R. (2024). Conduction velocity, G-ratio, and extracellular water as microstructural characteristics of autism spectrum disorder. *Plos One*, *19*(4), e0301964.

Newman, B. T., Untaroiu, A., & Druzgal, T. J. (2020). *A novel diffusion registration method with the NTU-DSI-122 template to transform free water signal fraction maps to stereotaxic space. Proceedings of the ISMRM 28th General Meeting*.

Pandolfi, V., Magyar, C. I., & Norris, M. (2014). Validity study of the CBCL 6–18 for the assessment of emotional problems in youth with ASD. *Journal of Mental Health Research in Intellectual Disabilities*, *7*(4), 306–322.

Petersen, C., Mucke, L., & Corces, M. R. (2024). CHOIR improves significance-based detection of cell types and states from single-cell data. *bioRxiv*, 2024.01. 18.576317.

Picci, G., Gotts, S. J., & Scherf, K. S. (2016). A theoretical rut: Revisiting and critically evaluating the generalized under/over-connectivity hypothesis of autism. *Developmental Science*, *19*(4), Article 4.

Raffelt, D. A., Tournier, J.-D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage*, *144*, 58–73.

Raffelt, D., Tournier, J.-D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O., & Connelly, A. (2012). Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *Neuroimage*, *59*(4), 3976–3994.

Richards, J. E., Sanchez, C., Phillips-Meek, M., & Xie, W. (2016). A database of age-appropriate average MRI templates. *Neuroimage*, *124*, 1254–1259.

Richards, J. E., & Xie, W. (2015). Brains for all the ages: Structural neurodevelopment in infants and children from a life-span perspective. *Advances in Child Development and Behavior*, *48*, 1–52.

Rushton, W. A. H. (1951). A theory of the effects of fibre size in medullated nerve. *The Journal of Physiology*, *115*(1), Article 1.

Salem, A. C., MacFarlane, H., Adams, J. R., Lawley, G. O., Dolata, J. K., Bedrick, S., & Fombonne, E. (2021). Evaluating atypical language in autism using automated language measures. *Scientific Reports*, *11*(1), 10968. https://doi.org/10.1038/s41598-021-90304-5

Sandrone, S., Aiello, M., Cavaliere, C., Thiebaut de Schotten, M., Reimann, K., Troakes, C., Bodi, I., Lacerda, L., Monti, S., & Murphy, D. (2023). Mapping myelin in white matter with T1-weighted/T2-weighted maps: Discrepancy with histology and other myelin MRI measures. *Brain Structure and Function*, *228*(2), 525–535.

Snow, A. V., Lecavalier, L., & Houts, C. (2009). The structure of the Autism Diagnostic Interview-Revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry*, *50*(6), 734–742. https://doi.org/10.1111/j.1469-7610.2008.02018.x

Stikov, N., Campbell, J. S., Stroh, T., Lavelée, M., Frey, S., Novek, J., Nuara, S., Ho, M.-K., Bedell, B. J., & Dougherty, R. F. (2015). In vivo histology of the myelin g-ratio with magnetic resonance imaging. *Neuroimage*, *118*, 397–405.

Stikov, N., Perry, L. M., Mezer, A., Rykhlevskaia, E., Wandell, B. A., Pauly, J. M., & Dougherty, R. F. (2011). Bound pool fractions complement diffusion measures to describe white matter micro and macrostructure. *Neuroimage*, *54*(2), 1112–1121.

Veraart, J., Novikov, D. S., Christiaens, D., Ades-Aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *Neuroimage*, *142*, 394–406.

Waizbard-Bartov, E., Ferrer, E., Heath, B., Rogers, S. J., Nordahl, C. W., Solomon, M., & Amaral, D. G. (2022). Identifying autism symptom severity trajectories across childhood. *Autism Research*, *15*(4), 687–701.

Wegiel, J., Kaczmarski, W., Flory, M., Martinez-Cerdeno, V., Wisniewski, T., Nowicki, K., Kuchna, I., & Wegiel, J. (2018). Deficit of corpus callosum axons, reduced axon diameter and decreased area are markers of abnormal development of interhemispheric connections in autistic subjects. *Acta Neuropathologica Communications*, *6*, 1–14.

Zhao, Y., Yang, L., Gong, G., Cao, Q., & Liu, J. (2022). Identify aberrant white matter microstructure in ASD, ADHD and other neurodevelopmental disorders: A meta-analysis of diffusion tensor imaging studies. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *113*, 110477.

Table 1:

Test Name	Focus	Administrator	Subscales	Labels
The Clinical Evaluation of Language Fundamentals – Fourth Edition (CELF-4)	Receptive language-skills including phonology, morphology, syntax, semantics, and working memory	Clinician	· Recalling Sentences	· celf4_rs_scld_scr
			· Formulating Sentences	· celf4_fs_scld_scr
			· Concepts & Following Directions	· celf4_cfd_scld_scr
			· Word Classes (Receptive)	· celf4_wc2r_scld_scr
			· Word Classes (Expressive)	· celf4_wc2e_scld_scr
			· Word Classes (Total)	· celf4_wc2t_scld_scr
			· Word Definition	· celf4_wd_scld_scr
			· Word Structure	· celf4_ws_scld_scr
			· Core Language	· celf4_cl_ss
The Behavior Rating Inventory of Executive Functions (BRIEF)	Executive function deficits	Caregiver	· Planning & Organizing	· brief_p_po
			· Organization of Materials	· brief_p_om
			· Monitor	· brief_p_monitor
			· Inhibit	· brief_p_inhibit
			· Shift	· brief_p_shift
			· Emotional Control	· brief_p_ec
			· Initiate	· brief_p_iniitiate
			· Working Memory	· brief_p_wm
			· Behavioral Regulation	· brief_p_bri

			· Metacognition	· brief_p_mi
			· Global Executive	· brief_p_gec
			· Negativity	· brief_p_negativity
			· Inconsistency	· brief_p_inconsistency
The Autism Diagnostic Observation Schedule (ADOS)	Social and behavioral features associated with ASD	Clinician	· Social Affect	· Socialaffect
			· Social Comparison	· Scoresumm_compscore
			· Behavioral Profile	· behavioraltotal
			· ADOS Total Score	· Scoresumm_overaltotal
Autism Diagnostic Interview-Revised (ADI-R)	Social and behavioral features associated with ASD	Clinician	· Language Acquisition (Words)	· acqorlossoflang_aword
			· Language Acquisition (Phrases)	· acqorlossoflang_aphrase
			· Loss of language	· loslang
			· Communication Profile	· dbaes_atotal
			· Social Behavior Profile	· dbaes_bvtotal
The Repetitive Behavior Scale-Revised (RBS-R)	Specifically examines repetitive behavior	Caregiver	· Stereotyped Behavior	· rbsr_stereo_total
			· Self-injurous Behavior	· rbsr_self_injurous_total
			· Compulsive Behavior	· rbsr_compulsive_total
			· Ritualistic Behavior	· rbsr_ritualistic_total
			· Sameness Behavior	· rbsr_sameness_total
			· Restricted Behavior	· rbsr_restricted_total

			· RBS-R Total Score	· rbsr_overall_total
The Adolescent/Adult Sensory Profile (AASP)	Sensory processing and preferences	Self	· Sensory Seeking	· sensory _c_fg1
			· Emotionally Reactive	· sensory_c_fg2
			· Low Endurance/Tone	· sensory_c_fg3
			· Oral Sensory Sensitivity	· sensory_c_fg4
			· Inattention/Distractibility	· sensory_c_fg5
			· Poor Registration	· sensory_c_fg6
			· Sensory Sensitivity	· sensory_c_fg7
			· Sedentary	· sensory_c_fg8
			· Fine Motor/Perceptual	· sensory_c_fg9
			· Low Registration	· sensory_a_qg1
			· Sensation Seeking	· sensory_a_qg2
			· Sensory Sensitivity	· sensory_a_qg3
			· Sensation Avoiding	· sensory_a_qg4
The Child Behavior Checklist (CBCL)	Behavioral and emotional syndromes, including anxiety, depression, aggression, and defiant	Caregiver	· Anxious Depressed	· cbcl_anxious
			· Somatic Complaints	· cbcl_somatic_c
			· Withdrawn	· cbcl_withdrawn
			· Attention Problems	· cbcl_attention
			· Aggressive Behavior	· cbcl_aggressive
			· Internalizing Problems	· cbcl_internal
			· Externalizing Problems	· cbcl_external
			· Affective Problems	· cbcl_affective

	behavior problems		· Anxiety Problems	· cbcl_anxiety
			· Attention Deficit/Hyperactivity	· cbcl_adhd
			· Oppositional Defiant Problems	· cbcl_oppositional
			· Social Problems	· cbcl_social_p
			· Thought Problems	· cbcl_thought
			· Rule-Breaking Behavior	· cbcl_rulebreak
			· Somatic Problems	· cbcl_somatic_p
			· Conduct Problems	· cbcl_conduct
			· CBCL Total Score	· cbcl_total
Social Responsiveness Scale (SRS-II)	Social behaviors associated with ASD	Caregiver	· Social Awareness	· awr_raw
			· Social Cognition	· cog_raw
			· Social Communication	· com_raw
			· Social Motivation	· mot_raw
			· Restricted Interest & Repetitive Behavior	· rrb_raw
			· Social Awareness (T-scored)	· awr_tscore
			· Social Cognition (T-scored)	· cog_tscore
			· Social Communication (T-scored)	· com_tscore
			· Social Motivation (T-scored)	· mot_tscore

			· Restricted Interest & Repetitive Behavior (T-scored)	· rrb_tscore
			· SRS Total Score	· srs2_rawscore
			· SRS Total Score (T-scored)	· srs2_tscore
Vineland Adaptive Behavior Scales – II (Vineland-II)	Measurement of adaptive behavior skills.	Caregiver	· Communication	· communicationdomain_totalb
			· Living Skills	· livingskillsdomain_totalb
			· Socialization	· socializationdomain_totalb
			· Adaptive Behavior	· composite_totalb
Differential Ability Scales (DAS)	Primary focus on cognitive abilities	Clinician	· Verbal Reasoning	· verbal_ss
			· Nonverbal Reasoning	· nvr_ss
			· Spatial Reasoning	· spatial_ss
			· General Conceptual Ability	· gca_ss
			· DAS Summary Score	· snc_ss

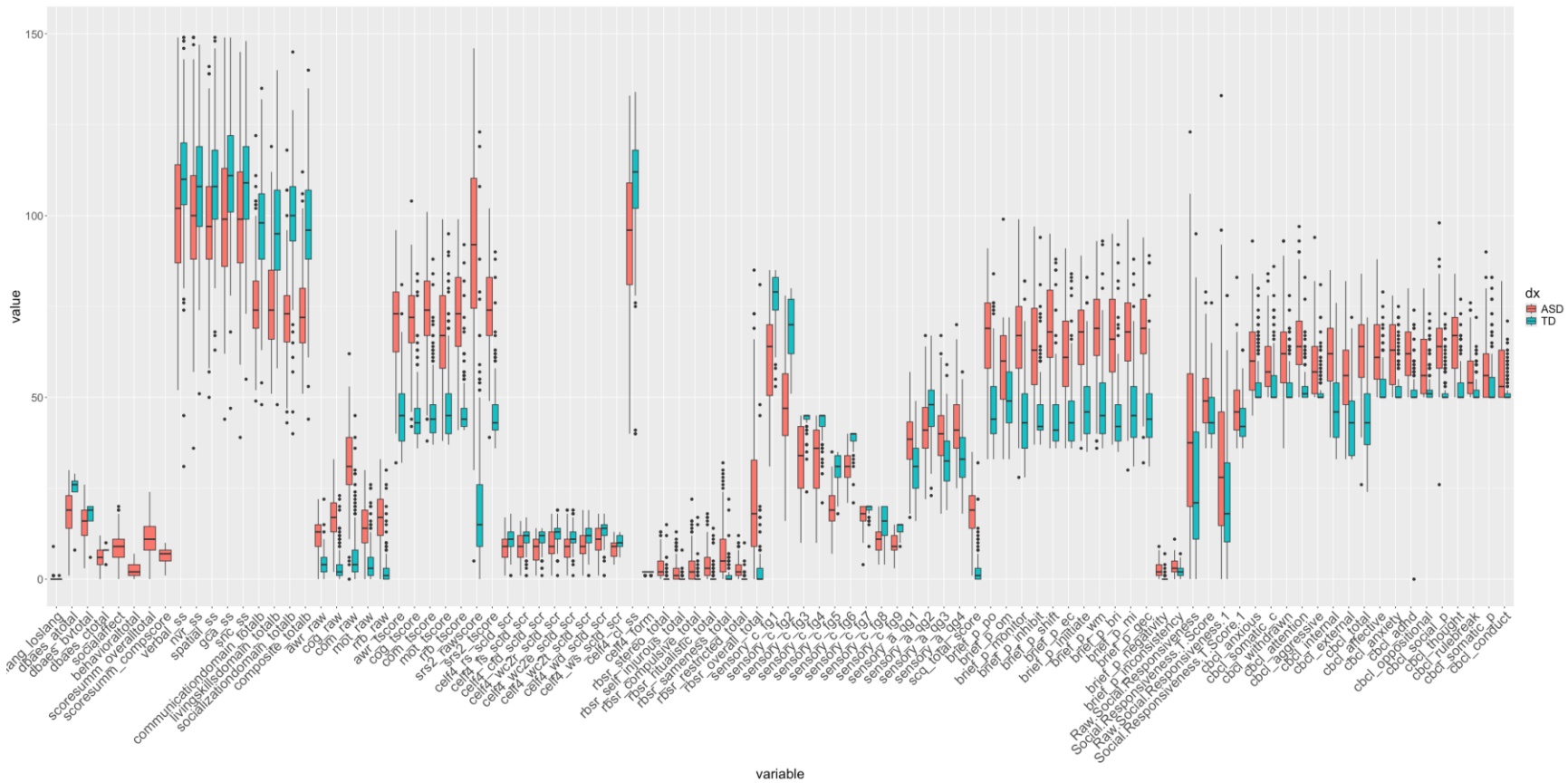


Figure 1: *Boxplots with median, upper and lower quartile (75th & 25th percentile scores, respectively), lines denoting 1.5x the inter-quartile range, and individual scores beyond that range, for each subscale metric and each test administered in this study. ASD participants typically had a higher median score than non-ASD participants with the exception of the CELF and Vineland-II tests. Most tests reliably scored either ASD or non-ASD subjects higher or lower at the group level, with the exception of the AASP*

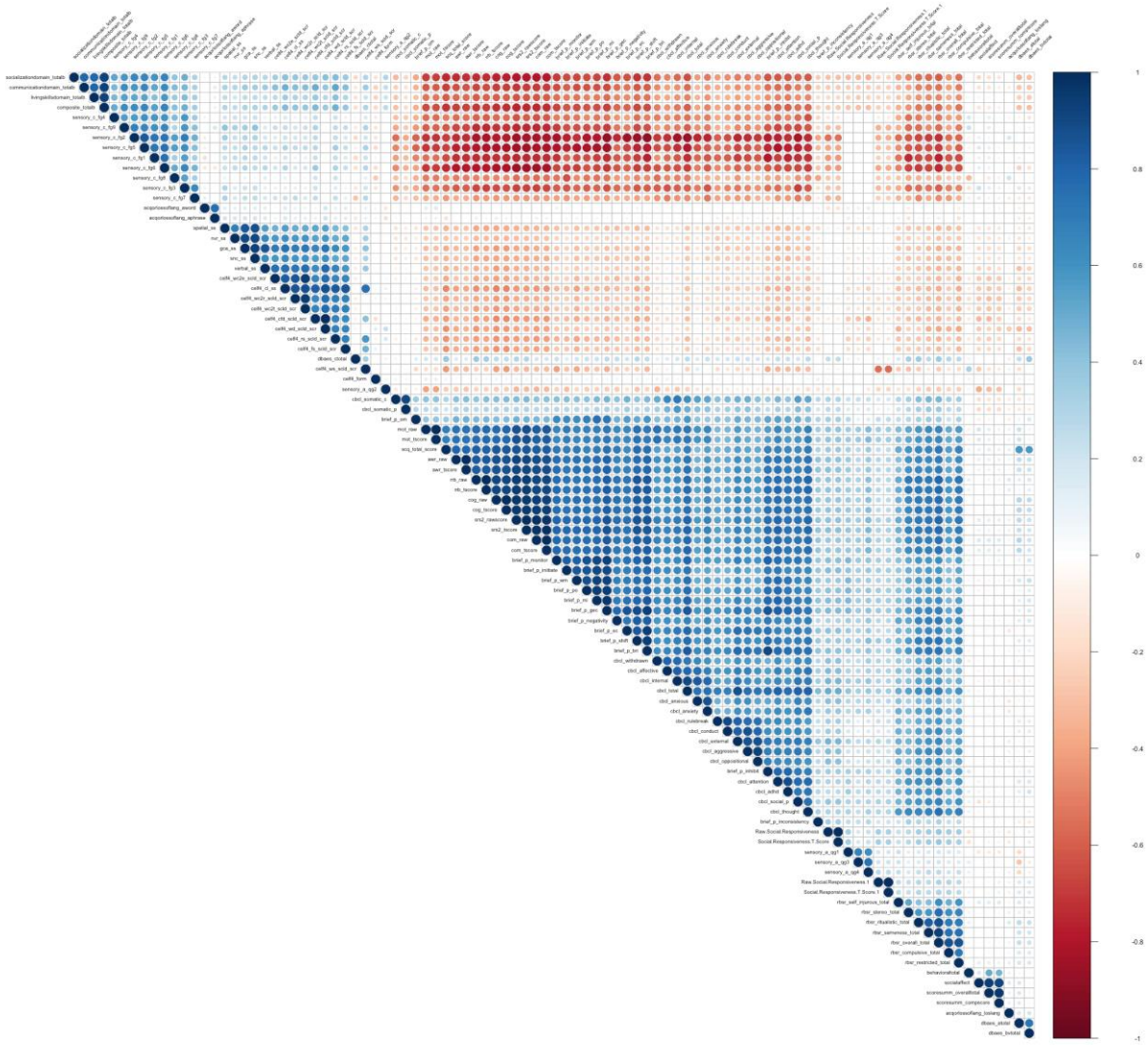


Figure 2: *Correlation plot showing the Pearson correlation coefficient between each of the cognitive and behavioral test subdomains as well as composite/total scores. Subdomains of tests were generally highly positively correlated within tests, but there was significant negative correlations between tests, particularly the AASP/CELF and behavioral metrics. Tests measuring behavioral and social functioning tended to be highly positively correlated. The underlying data composing column labels, p-values, and correlation coefficients are listed in Supplementary Table 1.*

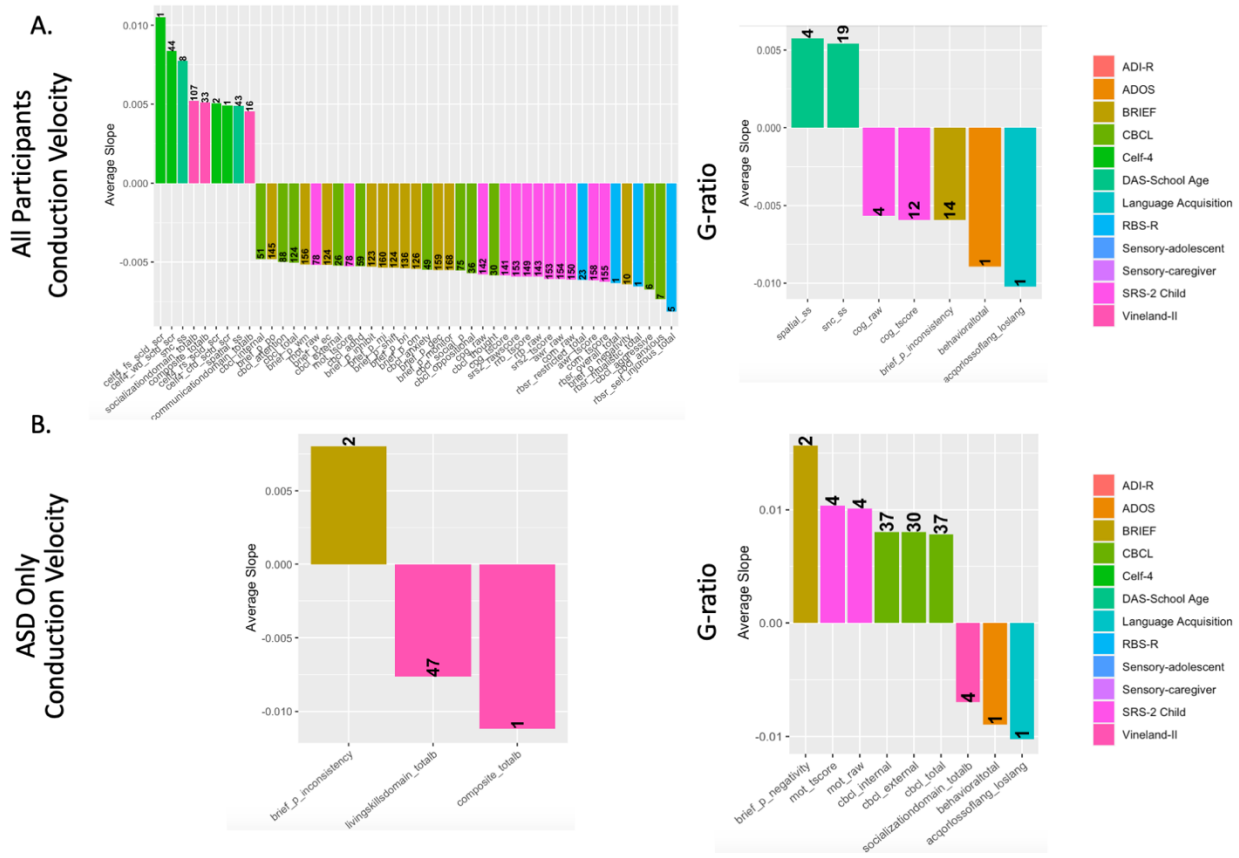


Figure 3: *Bar charts showing the mean slope of ROIs significantly associated with each subscale metric, colored by parent metric and separated by conduction velocity or g-ratio. The number on the bar specifies the number of ROIs significantly associated with each metric after multiple comparison corrections. Charts show relationships between brain cellular microstructure from a sample that includes all autistic and non-autistic participants (A) or exclusively participants diagnosed with ASD (B).*

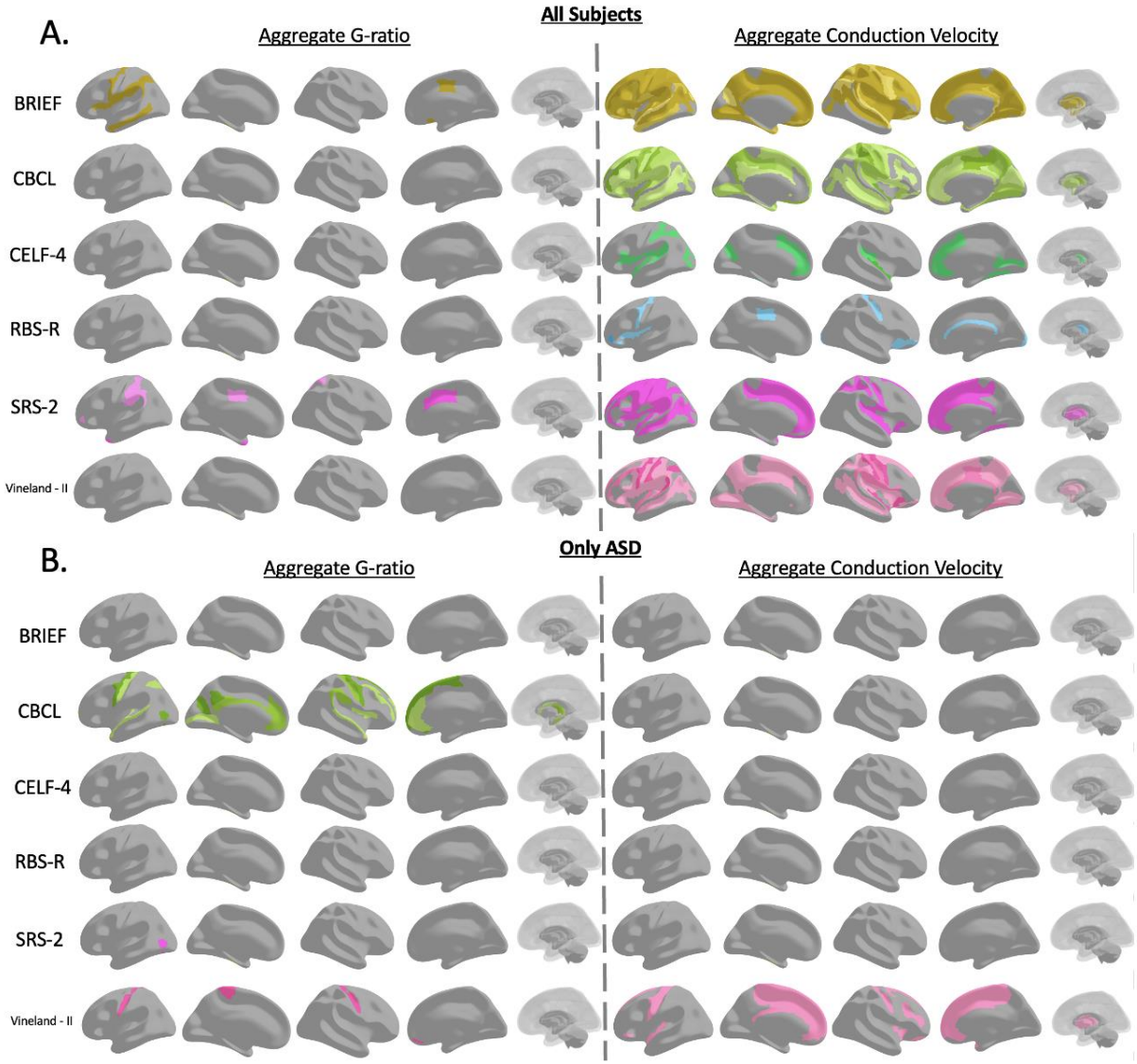


Figure 4: *Illustrations showing the location of ROIs significantly associated with each significant parent evaluation. Color is consistent with the bar charts in Fig. 3 and is darker if the region is associated with more subscales within the parent evaluation. Illustrations show relationships between brain cellular microstructure from a sample that includes all autistic and non-autistic participants (A) or exclusively participants diagnosed with ASD (B).*

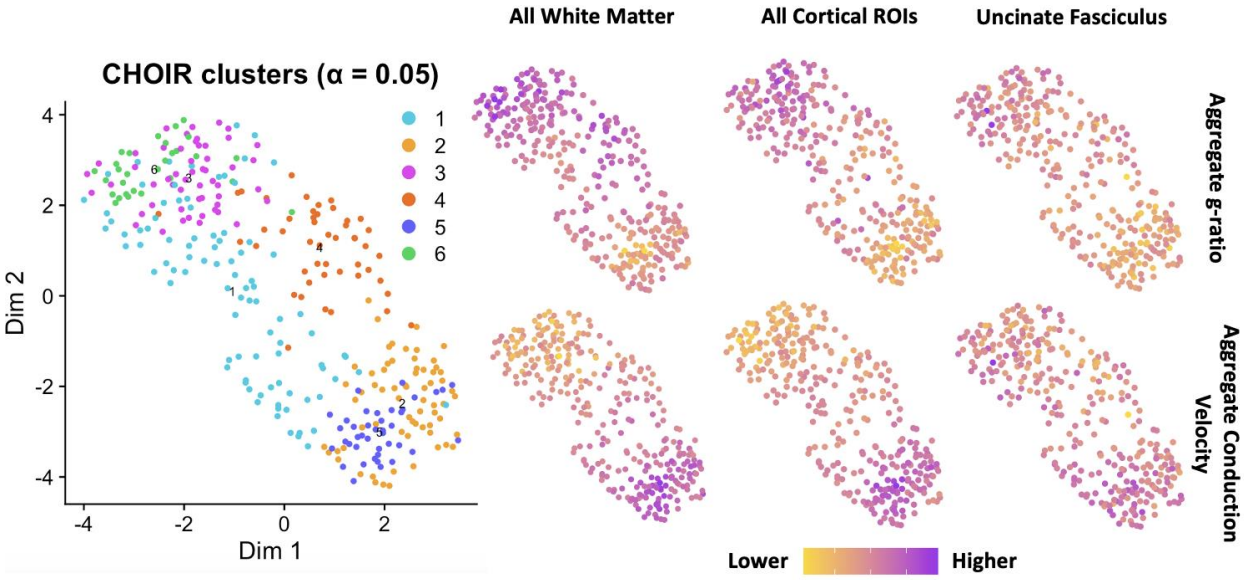


Figure 5: *Results of the CHOIR cluster analysis. CHOIR identified six total clusters from the neuroimaging data alone. These tended to be aligned along a gradient with a high aggregate conduction velocity and low aggregate g-ratio pole at clusters 2 & 5 and, in the inverse, a high aggregate g-ratio and low aggregate conduction velocity pole at clusters 4 & 6.*

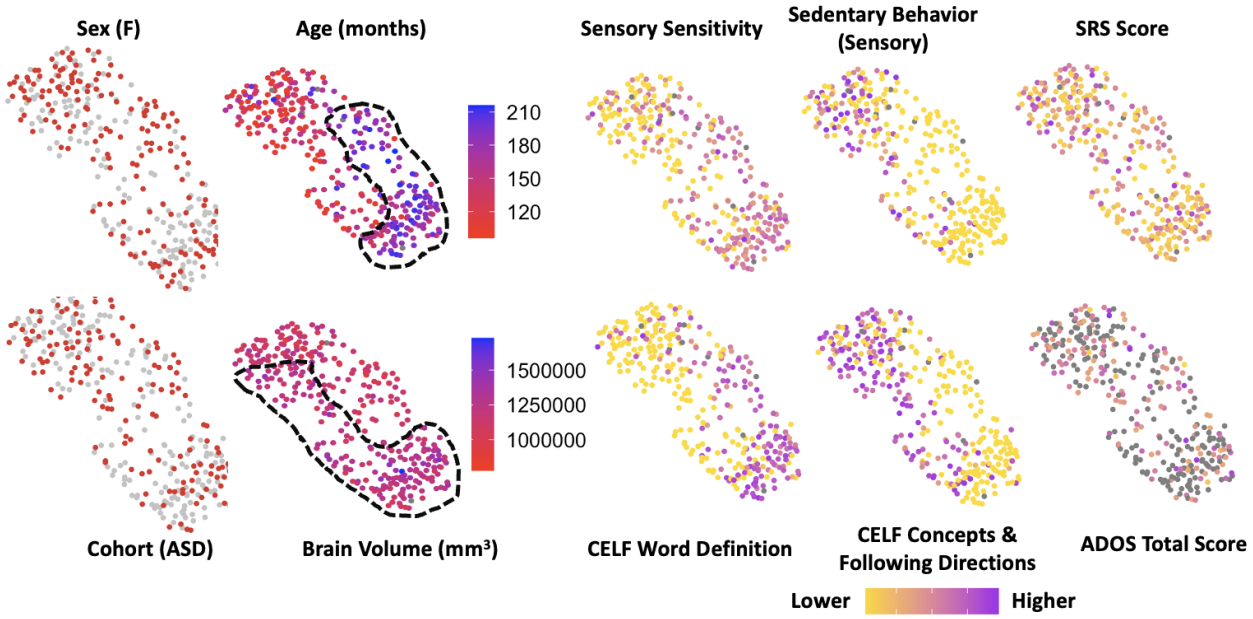


Figure 6: *Demographic, behavioral, and cognitive metrics displayed over CHOIR clusters presented previously in Figure 5. None of these metrics were used in the computation of the clusters. Sex and ASD diagnosis were not strongly related to any cluster, but age and brain volume were starkly divided. Several behavioral measures also showed stark divides, mainly following the age-related clustering but differing directions between different subscales. Some metrics, like the SRS and ADOS (ASD only), did not display mapping onto the defined clusters.*

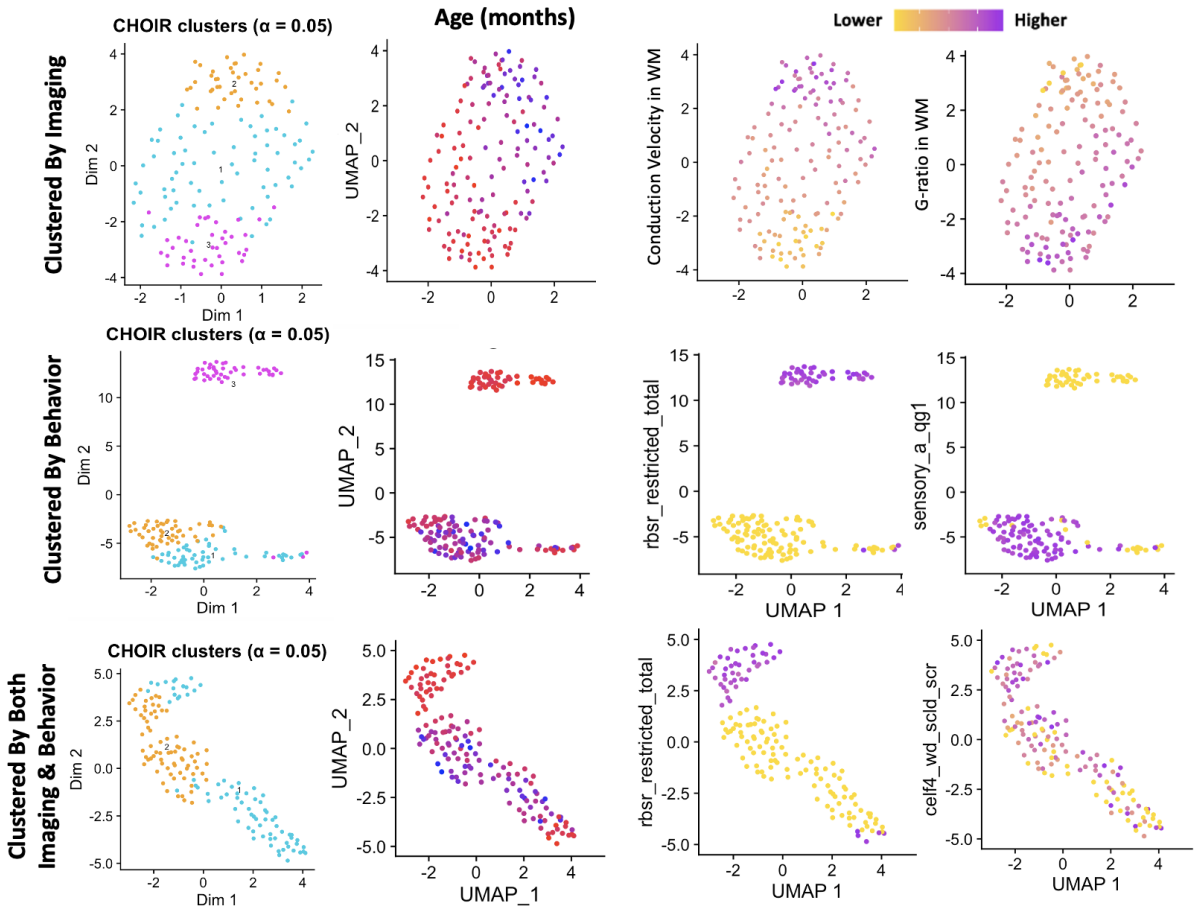
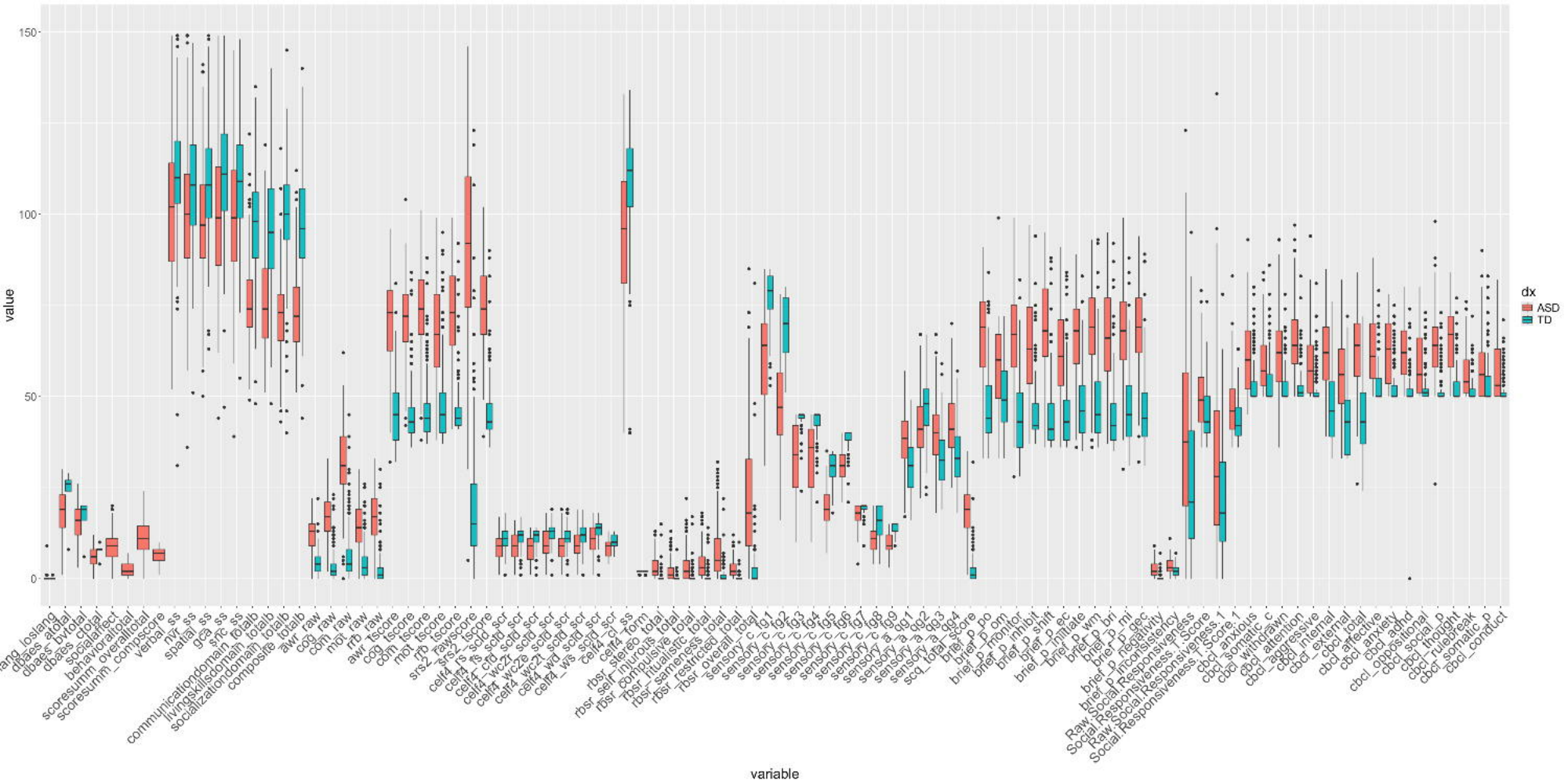
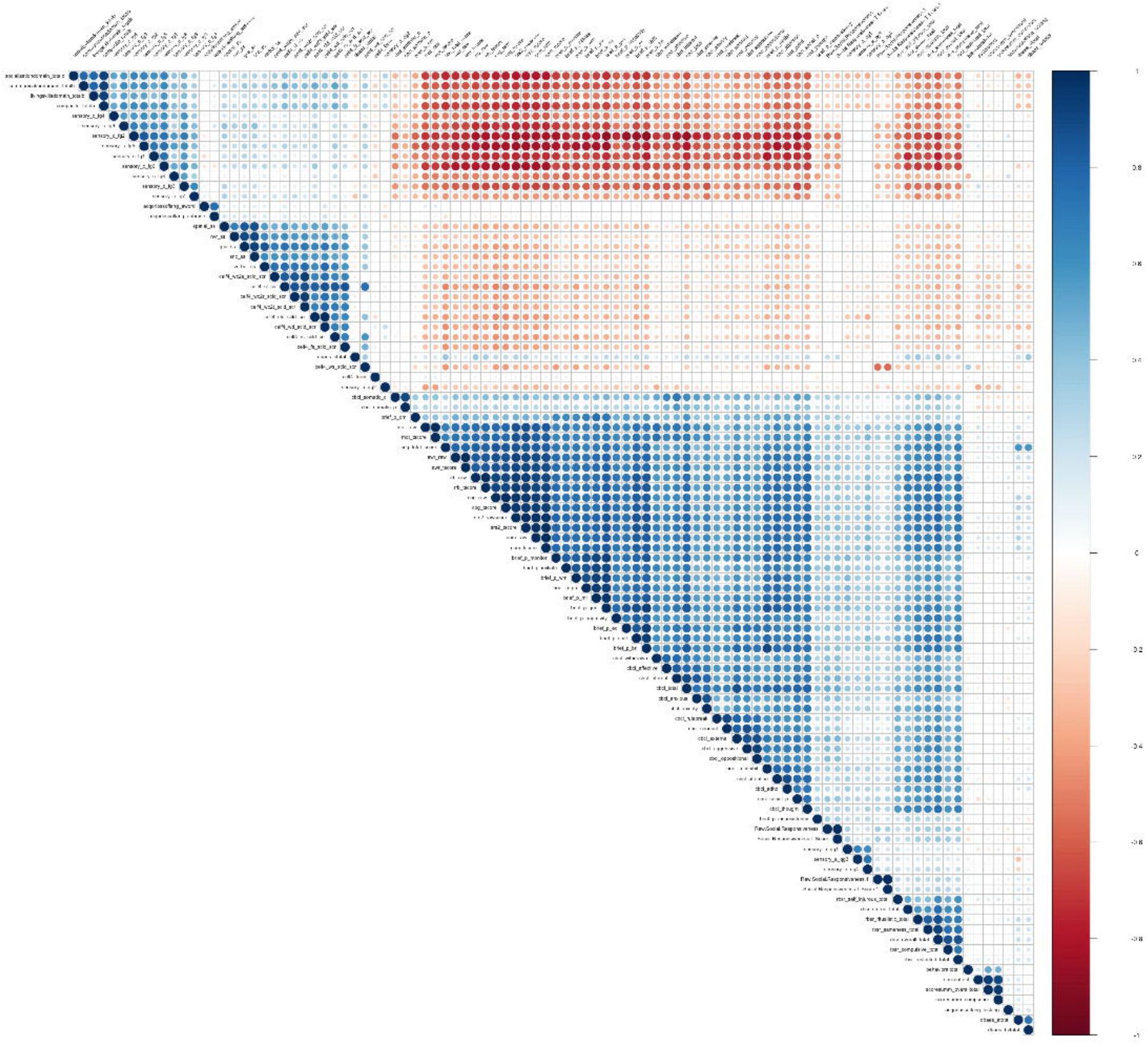
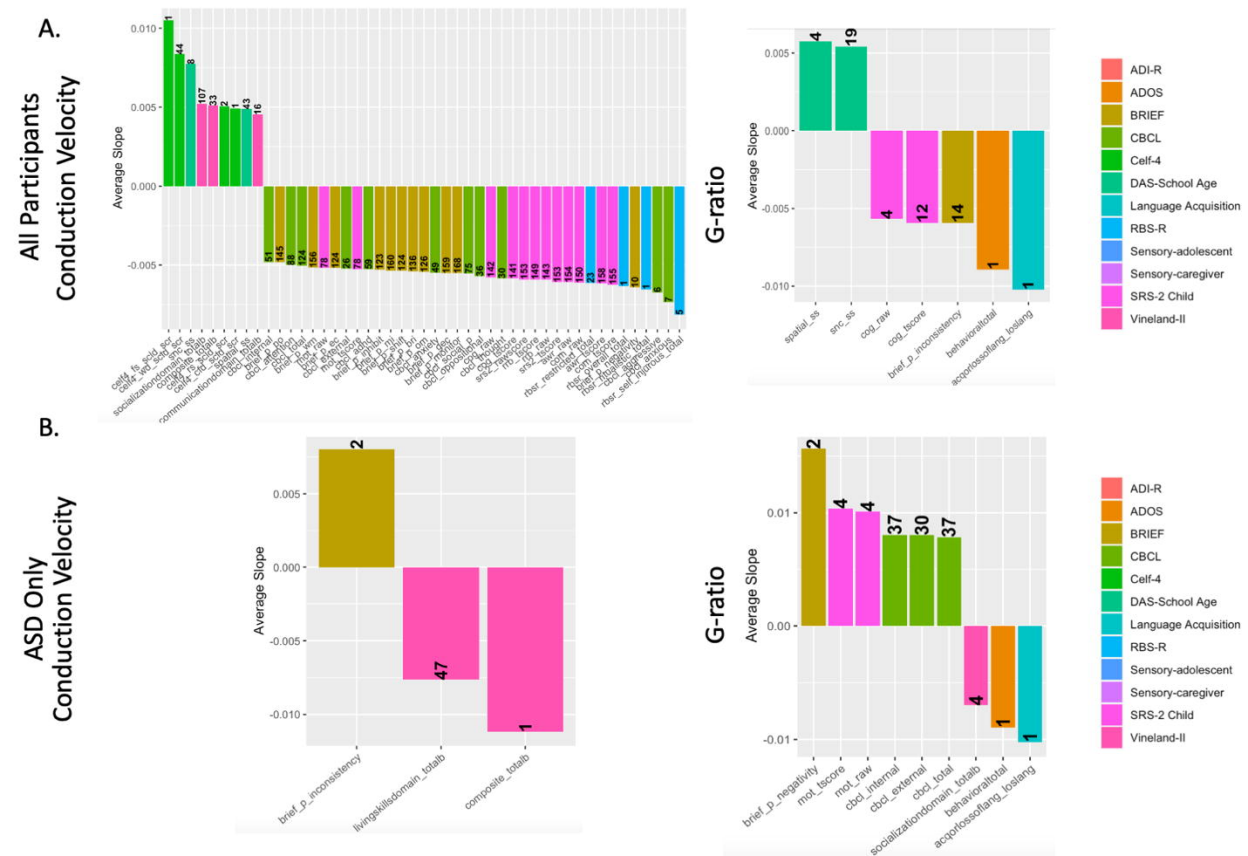


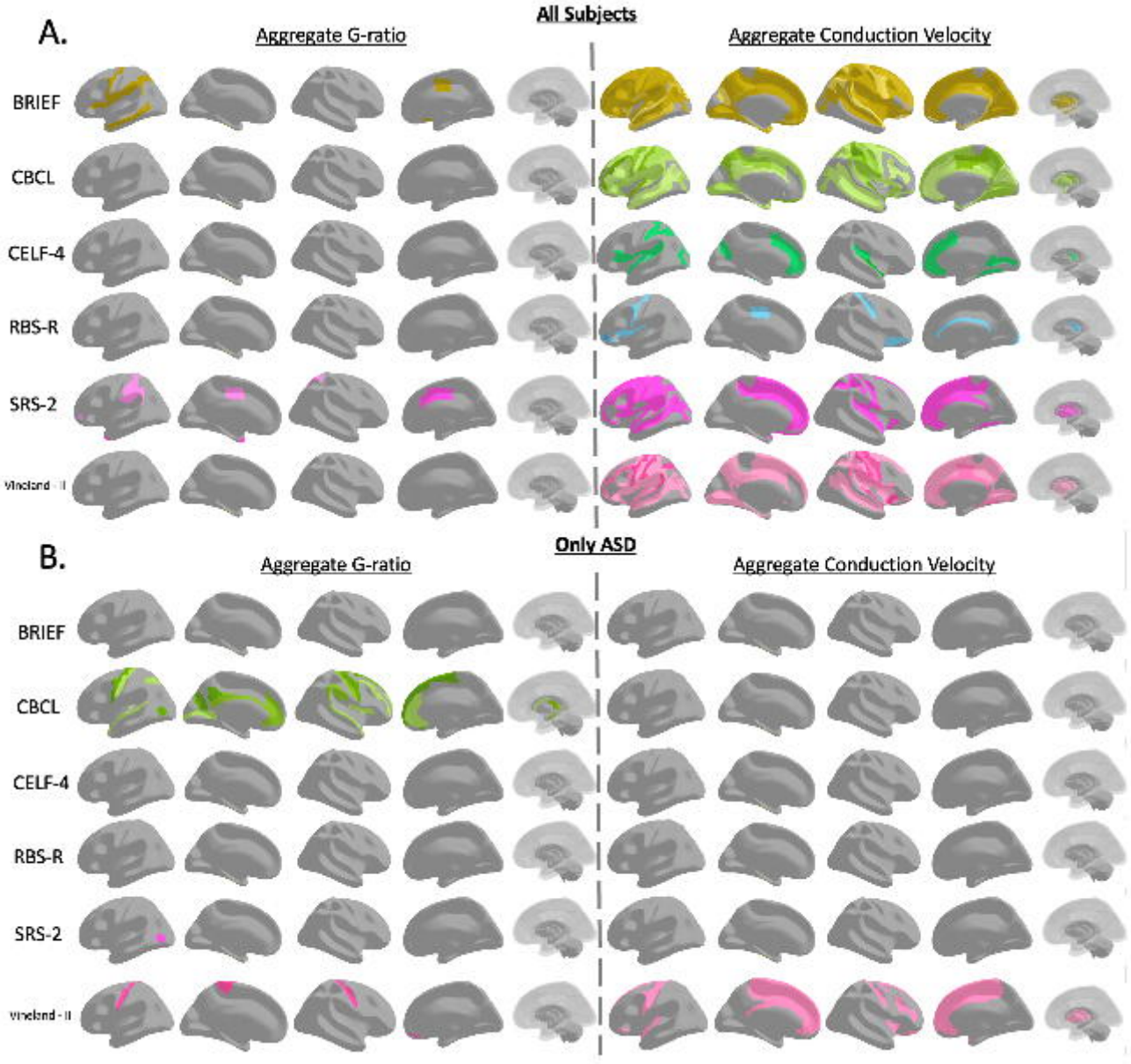
Figure 7: *CHOIR clusters of ASD participants only using just neuroimaging data (top row), behavioral metrics (middle row), or all metrics from both behavior and neuroimaging (bottom row). All cluster results were polarized by age (second column), with a particular cluster of the youngest participants in the study clearly observable (red). Age was not used to cluster any subjects, indicating that an age gradient was able to be reconstructed using exclusively the neuroimaging and behavioral metrics. Behaviors showed clear differences between the age-delineated clusters. With repetitive behaviors and language & sensory sensitivity metrics showing an inverse relationship indicating two distinct behavioral profiles.*

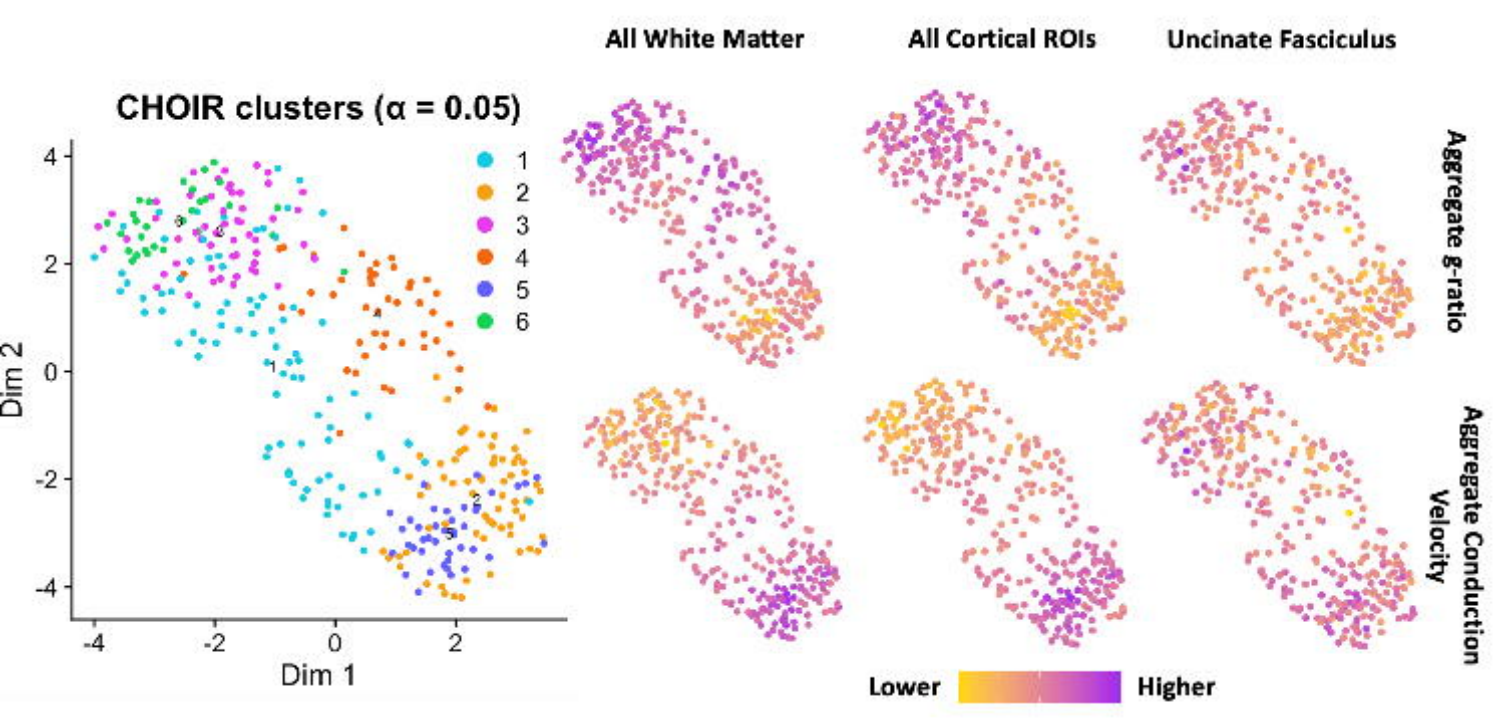


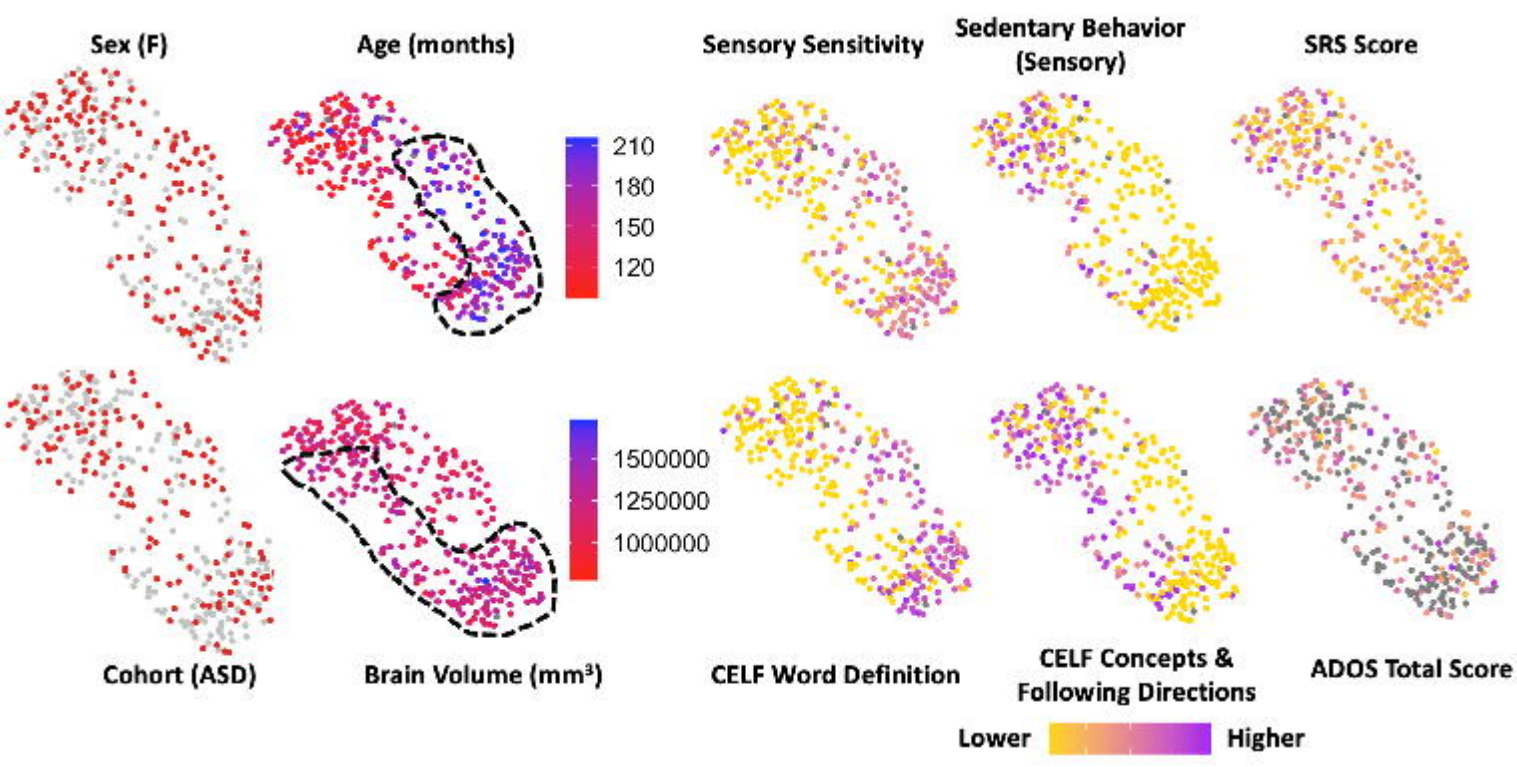
value
150
100
50
0
variable
dx
ASD
TD
lang_loslang
dbaes_atotal
dbaes_bvtotal
dbaes_ctotal
socialaffect
behavioralfotal
scoresumm_overallfotal
scoresumm_compscore
verbal_ss
nvr_ss
spatial_ss
gca_ss
snc_ss
communicationdomain_totalb
livingskillsdomain_totalb
socializationdomain_totalb
composite_totalb
awr_raw
cog_raw
com_raw
mot_raw
rrb_raw
awr_tscore
cog_tscore
com_tscore
mot_tscore
rrb_tscore
srs2_rawscore
srs2_tscore
celf4_rs_sold_scr
celf4_rs_scld_scr
celf4_cfd_scld_scr
celf4_wc2r_scld_scr
celf4_wc2e_scld_scr
celf4_wc2i_scld_scr
celf4_wd_scld_scr
celf4_ws_scld_scr
celf4_cl_ss
celf4_form
rbsr_stereo_total
rbsr_self_injurious_total
rbsr_compulsive_total
rbsr_ritualistic_total
rbsr_sameness_total
rbsr_restricted_total
rbsr_overall_total
sensory_c_tg1
sensory_c_tg2
sensory_c_tg3
sensory_c_tg4
sensory_c_tg5
sensory_c_tg6
sensory_c_tg7
sensory_c_tg8
sensory_c_tg9
sensory_a_qg1
sensory_a_qg2
sensory_a_qg3
sensory_a_qg4
scq_total_score
brief_p_po
brief_p_om
brief_p_monitor
brief_p_inhibit
brief_p_shift
brief_p_ec
brief_p_initiate
brief_p_wm
brief_p_bri
brief_p_mi
brief_p_gec
brief_p_negativity
brief_p_inconsistency
Raw.Social.Responsiveness
Social.Responsiveness.1_Score
Raw.Social.Responsiveness.1
Social.Responsiveness.1_Score.1
cbcl_anxious
cbcl_somatic_c
cbcl_withdrawn
cbcl_attention
cbcl_aggressive
cbcl_internal
cbcl_external
cbcl_total
cbcl_affective
cbcl_anxiety
cbcl_adhd
cbcl_oppositional
cbcl_social_p
cbcl_thought
cbcl_rulebreak
cbcl_somatic_p
cbcl_conduct

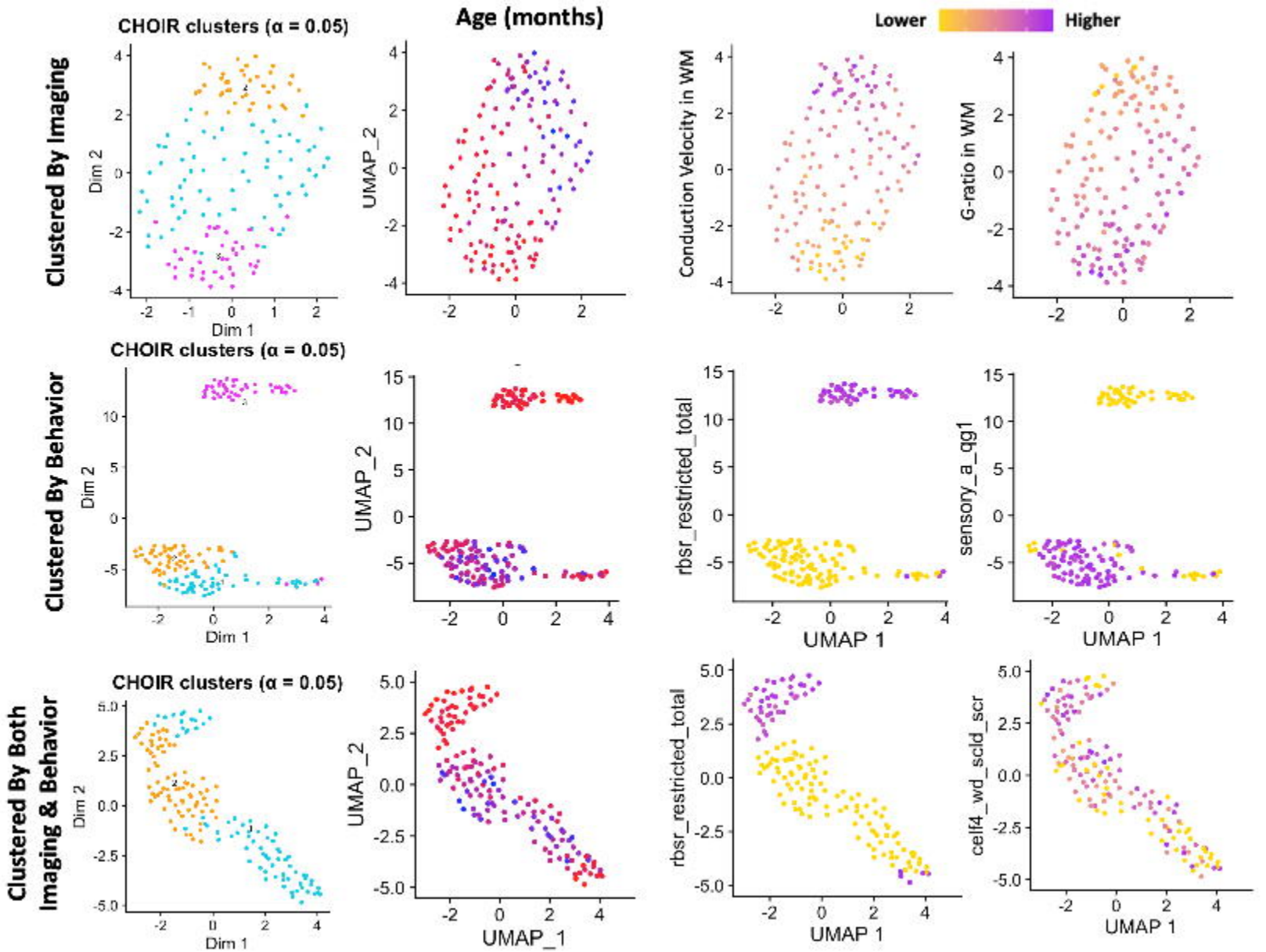



A.
All Participants
Conduction Velocity
Average Slope
G-ratio
Average Slope
ADI-R
ADOS
BRIEF
CBCL
Celf-4
DAS-School Age
Language Acquisition
RBS-R
Sensory-adolescent
Sensory-caregiver
SRS-2 Child
Vineland-II
B.
ASD Only
Conduction Velocity
Average Slope
G-ratio
Average Slope
ADI-R
ADOS
BRIEF
CBCL
Celf-4
DAS-School Age
Language Acquisition
RBS-R
Sensory-adolescent
Sensory-caregiver
SRS-2 Child
Vineland-II


All Subjects
A.
Aggregate G-ratio
Aggregate Conduction Velocity
BRIEF
CBCL
CELF-4
RBS-R
SRS-2
Only ASD
B.
Aggregate G-ratio
Aggregate Conduction Velocity
BRIEF
CBCL
CELF-4
RBS-R
SRS-2


CHOIR clusters (α = 0.05)
1
2
3
4
5
6
Dim 1
Dim 2
All White Matter
All Cortical ROIs
Uncinate Fasciculus
Aggregate g-ratio
Aggregate Conduction Velocity
Lower
Higher


Sex (F)
Age (months)
210
180
150
120
Sensory Sensitivity
Sedentary Behavior (Sensory)
SRS Score
Cohort (ASD)
Brain Volume (mm³)
1500000
1250000
1000000
CELF Word Definition
CELF Concepts & Following Directions
ADOS Total Score
Lower
Higher


Age (months)
Lower
Higher
Clustered By Imaging
CHOIR clusters (α = 0.05)
Dim 1
Dim 2
UMAP_2
Conduction Velocity in WM
G-ratio in WM
Clustered By Behavior
CHOIR clusters (α = 0.05)
Dim 1
Dim 2
UMAP_2
rbsr_restricted_total
sensory_a_qg1
UMAP 1
Clustered By Both Imaging & Behavior
CHOIR clusters (α = 0.05)
Dim 1
Dim 2
UMAP_2
UMAP_1
rbsr_restricted_total
celf4_wd_scld_scr
UMAP 1