Subcortical Brain and Behavior Phenotypes Differentiate Infants With Autism Versus Language Delay

Supplemental Information

Supplemental Methods and Materials

Participants. This study includes data from the Infant Brain Imaging Study (IBIS), an ongoing longitudinal study of infants at high and low familial-risk for autism spectrum disorder (ASD). The National Institutes of Health funded IBIS, through an Autism Center for Excellence Network award. The parent network includes four clinical sites: University of North Carolina, Chapel Hill; University of Washington, Seattle; The Children's Hospital of Philadelphia; and Washington University, St. Louis. Montreal Neurological Institute at McGill University managed data coordination. The University of North Carolina and the Scientific Computing and Imaging Institute at the University of Utah performed neuroimaging data processing. Parents provided written informed consent prior to participating in this study. The Institutional Review Boards at each clinical data collection site approved the procedures for this study.

High-risk (HR) infants, *n*= 382, had a sibling who met ASD criteria on the Social Communication Questionnaire (SCQ , 1), Autism Diagnostic Interview-Revised (2), and diagnosis was confirmed by medical records. Low-risk (LR) infants, *n*= 143, had typically developing older siblings who did not meet ASD screening criteria on the Family Interview for Genetic Studies (3). LR infants did not have first-degree relatives with ASD or intellectual disability. Further exclusionary criteria included: significant medical conditions known to affect brain development, sensory impairment, low birth weight (< 2,200 g) or prematurity (<36 weeks gestation), perinatal brain injury secondary to birth complications or exposure to specific medication or neurotoxins

during gestation, non-English speaking immediate family, contraindication for MRI, adoption, and first degree relative with psychosis, schizophrenia, or bipolar disorder. Two low-risk infants were excluded for outlying Mullen Early Learning Composite standard scores (-2SD from the *M*). Two low-risk infants met criteria for ASD and ten met criteria for language delay. These infants were excluded from all analyses, as the groups were too small to be analyzed separately. Table 1 includes participant demographics.

MRI Acquisition. Pediatric imaging was completed during natural sleep at each clinical site using identical 3-T Siemens TIM Trio scanners (Siemens Medical Solutions, Malvern, Pa.) equipped with 12-channel head coils. The imaging protocol included 1) a localizer scan, 2) 3D T1 MPRAGE: TR=2400ms, TE=3.16ms, 160 sagittal slices, FOV=256, voxel size = $1mm^3$, 3) 3D T2 FSE TR=3200ms, TE=499ms, 160 sagittal slices, FOV=256, voxel size = $1mm^3$, and 4) a 25 direction DTI: TR=12800ms, TE=102ms, slice thickness = 2mm isotropic, variable b value = maximum of $1000s/mm^2$, FOV=190.

A number of quality control procedures were employed to assess scanner stability and reliability across sites, time, and procedures. Geometry phantoms were scanned monthly and human phantoms (two adult subjects) were scanned annually to monitor scanner stability at each site across the study period. Details on the stability procedures for IBIS and scanner quality control checks are described elsewhere (4).

Image Preprocessing. All T1- and T2-weighted images were corrected for geometric distortions (5) and intensity non-uniformity (6). T2-weighted images underwent linear, rigid registration to the corresponding T1-weighted images via mutual information registration (7). Subsequently, both T1- and T2-weighted images were transformed to stereotactic space based on the registration of the T1 scan. The skull was extracted using a “majority voting approach” between the T1 atlas

mask, T2 atlas mask, and the T1 and T2 images jointly via FSL Brain Extraction Tool (8). The resulting brain masks were manually corrected if necessary. All corrected and skull-stripped T1 and T2 images were used as input for an expectation, maximization-based, tissue segmentation tool (AutoSeg pipeline (9)) to obtain white matter, gray matter, and CSF (10).

Segmentation of Subcortical Brain Structures. A graph-based, multi-atlas method developed by our laboratory was employed to segment the subcortical structures (9). A brief summary is presented here, see Wang et al (9) for complete details. First, all atlases and participant MR images were paired and co-registered via symmetric diffeomorphic registration using the ANTS (Advanced Normalization ToolS) registration tool (11). Second, a directed graph with edge weights based on intensity and shape similarity was constructed between all atlases and the participant MR image (9). Third, the shortest path from each atlas to the participant image was computed (with atlases sharing the same shortest paths combined into the same cluster), and the atlas closest to the participant for each cluster was selected as the neighboring template (9). Finally, the final segmentation was produced by fusing the propagated label files of the neighboring templates via weighted majority voting (9). For the 12 and 24 months MRI data, atlas templates were derived from 8 cases at the 12-month time point as well as 8 cases at the 24-month time point, which were manually segmented by a single experimenter. In order to eliminate any template induced asymmetric laterality biases, we employed left-right flipped versions of all atlas images in the multi atlas segmentation, resulting in a total of 32 atlas templates. These atlas templates were employed in the multi atlas segmentation as it was applied to all 12-month data sets in this study. The multi-atlas segmentation method was validated in a leave-one-out validation analysis that achieved high Dice coefficients for all structures (mean=91.5%, SD=.03, range=87.2-96.0%). All segmentations underwent visual quality inspection by two experimenters (blind to diagnosis,

risk status, sex, and scan site). Ninety-eight percent of scans met quality inspection criteria for inclusion in the final analysis (N=368 scans). Figure S1 shows the results of the segmentation of left and right caudate, amygdala, and thalamus. The Autoseg software pipeline for multi-atlas-based segmentation is publicly available on the NIH NITRC website (Neuroimaging Informatics Tools and Resources Clearinghouse) at http://www.nitrc.org/projects/autoseg.

sMRI Substructures of Interest (Figure S1). The amygdalae are a set of nuclei in the limbic system located deep medially within the temporal lobes (12). The amygdala contains major bidirectional pathways to the cortex via the ventral amygdalofugal pathway, to subcortical structures via the stria terminalis, as well as direct connections to the hippocampus, dorsomedial nucleus of the thalamus, entorhinal cortex, and brainstem. The thalamus is a midline structure composed of two halves that are each bulb-shaped. The thalamus is connected to cortex, spinal cord, and hippocampus via white matter fiber tracts. The caudate nucleus is the lateral part of the dorsal striatum and is located next to the thalamus in each of the two hemispheres of the brain. The caudate receives afferent projections from nearby cortical regions and the thalamus, and primarily projects efferent connections to the substantia nigra. Brain-behavior associations were examined for only those substructures described above, and not on all substructures generated from our segmentation pipeline.

Given that laterality is commonly demonstrated with language function, we examined structural laterality for each subcortical structure to determine if left and right subcortical volumes should be treated separately, or summed together. To access potential hemispheric differences a Laterality Index (LI) was created according to the following formula $LI = 1 * \frac{(Left-Right)}{(Left+Right)}$ (13). This formula results in values ranging from -1 to +1, and a common threshold for hemispheric dominance is 0.2 (14, 15). The sign of LI indicates the direction of asymmetry, with positive values

indicating larger left side volumes, and negative values indicating larger right side volumes. As indicated in Table S3, LI values for each subcortical structure were all virtually 0, and no individual in the study exceeded 0.08, far less than the threshold for hemispheric dominance of 0.2 (14, 15). Given the lack of laterality, the left and right substructure volumes were summed to create a total volume.

Statistical Analysis. As of December 1st, 2016, data were available for 525 infants who completed at least two behavioral visits and had a completed DSM-IV-TR checklist at the 24-month diagnostic evaluation. Of these infants, 368 also had a sMRI acquisition at 12-months that passed quality control procedures (HR-ASD *n*= 46, HR-LD *n*= 29; HR-Neg *n*= 189, LR-Neg *n* = 104). Longitudinal language analyses utilized data from the full sample of 525 infants. The raw data that support the findings from this study will be publicly available from the NIH National Database for Autism Research (NDAR).

A two-step process was utilized for model fitting. First, potential covariates (mother's education and clinical data site) were examined for inclusion in the final model to determine if they were significantly associated with the dependent variable. Mother's education has been shown to be associated with emerging language skills (16). Clinical data collection site was investigated as a potential covariate to account for potential differences in clinical data collection and MRI across the four sites. Second, *a priori* covariates were examined and included in the final model regardless of their model contribution. *A priori* covariates include MSEL NVDQ, sex of the infant, chronological age at assessment or scan, and total cerebral volume (TCV; for brain-behavior analyses only). MSEL NVDQ was chosen as an *a priori* covariate to ensure that language findings were not reflecting general cognitive ability. Sex of the infant was included as an *a priori* covariate due to known sex differences in language acquisition and sexual dimorphism in subcortical

structure size. Age at assessment or scan was included to account for variations in age allowed by the age window. Total cerebral volume was included in all brain-behavior analyses so analyses would reflect differences in subcortical size irrespective of overall brain size. This model fitting procedure was completed separately for longitudinal language analyses and cross-sectional brain-behavior analyses.

Longitudinal language skills were examined across 6, 12, and 24-month visits using general linear mixed models (GLMM). For each model, the language measure was the dependent variable and fixed effects for the model included visit, group, and group x visit interaction. The intercept term is treated as random with the objective to reduce the subject-to-subject variation. Model fitting procedures resulted in a final model, which included maternal education and clinical site as potential covariates, and MSEL NVDQ and sex of the infant as *a priori* covariates. The primary hypothesis was for language trajectories to diverge by group. This hypothesis was assessed by the group x visit interaction in each model. Significant group x visit interactions were followed up with planned cross-sectional analyses at each time point. Post hoc group comparisons were adjusted for multiple comparisons using an adaptive FDR procedure (17). Adjusted *q*-values are presented with the significance threshold at $q < 0.05$.

GLMM were applied to examine brain-behavior associations. In an *a priori* selection of variables, all brain-behavior analyses utilized brain substructure volume at the 12-month time point and language skills at the 24-month time point. We focused on 24-month language scores because there was the greatest variability in scores at this time point, and because this time point is closer to when we would expect delays in language to be most apparent. We focused on 12-month subcortical volumes because our main aim was to examine brain development prior to the measured behavior, following previous research that brain changes precede behavioral changes

(18–20). For each model, the language variable was the dependent variable and the fixed effects for the model included group, brain subcortical structure volume, and a group x brain interaction. A similar model fitting process as described above was conducted for the brain-behavior analyses. Model fitting procedures resulted in a final model that included clinical data collection site as a potential covariates, and MSEL NVDQ, sex of the infant, and total cerebral volume as *a priori* covariates. For completeness all four diagnostic groups were included in the GLMM and the group x brain interaction term was reported. Our main aim for these analyses was to test the effect of one specific contrast, HR-ASD vs. HR-LD, to determine if these two groups differed in their brain-behavior association. This planned contrast was designed to indicate if the HR-ASD and HR-LD group have similar or dissimilar brain-behavior phenotypes. Since this contrast was planned, results were analyzed regardless of the results of the omnibus test (21). Lastly, as a post-hoc test to a significant omnibus test, we estimated the simple slopes of the subcortical volume within each group and tested the slopes to see if they were different from zero. An FDR procedure was used to correct for multiple comparisons. Adjusted *p*-values were presented as *q*-values.

Lastly, GLMM was used to conduct follow-up brain-behavior analyses. The aim of these analyses was to determine if HR-ASD infants with language delay displayed brain-behavior phenotypes that were more similar to the HR-LD infants or the HR-ASD without language delay infants. The previously described language delay criteria were applied to the HR-ASD group to create two groups, one with ASD and language delay (ASD-LD+, n= 28), and one with ASD but without language delay (ASD-LD-, n= 16). The HR-LD group was also included in this analysis. The same fixed effects described above (subcortical structure, clinical site, MSEL NVDQ, sex of the infant, and total cerebral volume) were included in this model, a group x brain interaction term was also included, and receptive advantage scores were the dependent variable. To determine if

the groups differed in their brain-behavior associations, three contrasts were tested, ASD-LD+ vs. ASD-LD-, ASD-LD+ vs. HR-LD, and ASD-LD- vs. HR-LD. We also tested the significance of the effect within each group. An FDR procedure was used to correct for multiple comparisons. Adjusted *p*-values were presented as *q*-values.

All analyses were done using SAS statistics software, version 9.3 (SAS Institute INC, Cary, NC, USA).

Supplemental Results

Participant Characteristics. As expected, there were more males than females who were classified in the HR-ASD group than the HR-Neg and LR-Neg groups (overall X^2= 15.86, p = .001; HR-ASD vs. HR-Neg, X^2= 15.17, p < .0001; HR-ASD vs. LR-Neg, X^2= 8.78, p = .003). The remaining pair-wise comparisons were not significant, p > .147. The LR-Neg group had mothers with higher levels of educational attainment when compared to all other groups (overall X^2= 31.56, p = .0002; HR-ASD vs. LR-Neg, X^2= 12.56, p < .002; HR-LD vs. LR-Neg, X^2= 15.14, p = .001; HR-Neg vs. LR-Neg, X^2= 11.95, p = .002). Additionally, the HR-Neg group had mothers with a higher level of educational attainment than the HR-LD group (X^2= 6.10, p = .047). The groups did not differ in composition of infant race (overall X^2= 8.32, p = .759). The HR-ASD group had higher ADOS Severity Scores than all other groups, while the remaining groups did not differ from one another (overall F (3, 497) = 316.87, p < .0001; HR-ASD vs. HR-LD, t= 18.55, p < .0001; HR-ASD vs. HR-Neg, t= 28.86, p < .0001; HR-ASD vs. LR-Neg, t= 27.24, p < .0001).

Table S1. Least square means for development of language skills.

	HR-ASD (a)	**HR-LD (b)**	**HR-Neg (c)**	**LR-Neg (d)**	**Post-Hoc Comparisons**
	LSM (SE)	*LSM (SE)*	*LSM (SE)*	*LSM (SE)*	
6-month visit					
MSEL VDQ	89.99 (2.19)	86.71 (2.64)	86.88 (1.12)	90.61 (1.42)	n/a
Exp. Lang. t-score	45.05 (1.04)	43.22 (1.25)	44.69 (0.53)	46.43 (0.68)	n/a
Rec. Lang. t-score	48.48 (1.28)	50.26 (1.56)	49.19 (0.65)	49.98 (0.83)	n/a
Vineland Comm. SS	95.43 (1.97)	96.31 (2.41)	96.97 (1.02)	99.40 (1.30)	n/a
Receptive Adv.	0.22 (0.23)	0.82 (0.28)	0.40 (0.12)	0.24 (0.15)	n/a
12-month visit					
MSEL VDQ	82.08 (2.08)	85.85 (2.72)	93.58 (1.07)	97.58 (1.47)	a, b < c, d, c<d
Exp. Lang. t-score	41.95 (1.42)	42.44 (1.86)	48.37 (0.72)	49.71 (1.00)	a, b < c, d
Rec. Lang. t-score	40.07 (1.04)	42.42 (1.38)	43.92 (0.53)	46.22 (0.78)	a < c, d; b, c< d
Vineland Comm. SS	91.60 (1.52)	93.98 (1.96)	99.38 (0.78)	102.71 (1.07)	a, b < c, d, c<d
Receptive Adv.	-0.28 (0.37)	0.39 (0.49)	-0.82 (0.19)	-0.56 (0.26)	n/a
24-month visit					
MSEL VDQ	82.46 (1.78)	80.34 (2.17)	103.97 (0.91)	104.79 (1.25)	a, b < c, d
Exp. Lang. t-score	42.34 (1.17)	38.86 (1.52)	50.91 (0.60)	51.03 (0.82)	a, b < c, d
Rec. Lang. t-score	40.65 (1.07)	39.00 (1.40)	54.21 (0.55)	54.73 (0.75)	a, b < c, d
Vineland Comm. SS	92.14 (1.12)	95.20 (1.45)	103.34 (0.58)	104.71 (0.81)	a, b < c, d
Receptive Adv.	-0.41 (0.59)	1.57 (0.77)	1.90 (0.30)	2.08 (0.41)	a < b*, c, d

Notes: *MSEL VDQ*, MSEL Verbal Developmental Quotient; *Exp. Lang. t-score*, Expressive Language t-score; *Rec. Lang. t-score*, Receptive Language t-score; *Vineland COM SS*, Vineland Communication Subscale Standard Score; *Receptive Adv*., Receptive Advantage Score. * HR-ASD vs HR-LD comparison did not survive multiple comparison corrections (p = .033, corrected p = .067).

Table S2. Tests of fixed effects for longitudinal language analyses.

Dependent Variable	**Sex of Infant**		**Mullen NVDQ**		**Site**		**Mother's Education**		**Age**		**Diagnostic Group[a]**		**Age x Group**	
Longitudinal	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p/q*[b]	*F*	*q*
MSEL VDQ	3.00	.083	68.74	.0001	16.14	<.0001	3.76	.024	4.49	.034	10.09	<.0001	52.38	<.0001
Exp. Lang. *t*-score	2.19	.139	33.05	<.0001	12.57	<.0001	1.93	.146	0.50	.480	5.49	.001	34.13	<.0001
Rec. Lang. *t*-score	1.84	.174	48.91	<.0001	13.09	<.0001	5.46	.004	8.42	.003	15.74	<.0001	58.07	<.0001
Vineland COM SS	10.19	.001	24.99	<.0001	2.93	.033	2.73	.066	3.48	.062	0.66	.578	14.41	<.0001
Receptive Adv.	0.00	.946	0.00	.964	5.54	.001	0.84	.432	12.48	.0005	2.86	.036	4.25	.005
6-month visit														
MSEL VDQ	0.66	.416	2.21	.138	28.22	<.0001	0.08	.927	14.95	.0001	1.56	.332		
Exp. Lang. *t*-score	1.45	.229	1.06	.303	23.00	<.0001	0.29	.751	3.12	.078	2.18	.332		
Rec. Lang. *t*-score	0.23	.631	1.13	.287	18.89	<.0001	0.35	.702	10.18	.001	0.50	.684		
Vineland Comm.	2.67	.103	0.94	.332	7.75	<.0001	0.81	.445	4.96	.026	1.14	.332		
Receptive Adv.	0.38	.540	0.01	.928	8.03	<.0001	0.46	.629	0.34	.561	1.32	.332		
12-month visit														
MSEL VDQ	4.19	.041	4.27	.039	8.49	<.0001	1.83	.161	0.76	.385	12.21	<.0001		
Exp. Lang. *t*-score	2.25	.134	2.92	.088	1.13	.337	0.61	.545	0.36	.546	8.09	<.0001		
Rec. Lang. *t*-score	4.17	.041	2.75	.097	21.21	<.0001	5.16	.006	15.37	.0001	7.24	<.0001		
Vineland Comm.	7.47	.006	3.81	.051	2.49	.059	0.12	.888	1.24	.265	11.99	<.0001		
Receptive Adv.	0.90	.767	0.13	.714	8.63	<.0001	1.05	.351	6.15	.013	1.91	.127		
24-month visit														
MSEL VDQ	3.02	.083	137.19	<.0001	1.20	.311	4.20	.015	0.04	.837	57.67	<.0001		
Exp. Lang. *t*-score	4.20	.041	78.88	<.0001	4.95	.002	3.90	.020	1.00	.317	26.68	<.0001		
Rec. Lang. *t*-score	1.35	.245	105.08	<.0001	0.79	.499	2.75	.065	1.51	.219	63.88	<.0001		
Vineland Comm.	7.07	.008	31.39	<.0001	1.80	.146	4.08	.017	0.22	.641	32.40	<.0001		
Receptive Adv.	0.78	.337	0.33	.568	7.97	<.0001	0.50	.604	4.14	.042	4.31	.005		

Notes: MSEL VDQ, MSEL Verbal Developmental Quotient; *Exp. Lang. t-score*, Expressive Language t-score; *Rec. Lang. t-score*, Receptive Language t-score; *Vineland COM SS*, Vineland Communication Subscale Standard Score; *Receptive Adv*., Receptive Advantage Score. [a]Model covariates include non-verbal developmental quotients (NVDQ) calculated from visual reception and fine motor subscales, maternal education, sex of the infant, and clinical site. [b] *P*-values are reported for longitudinal models; *q*-values are reported for cross-sectional models.

Table S3. Laterality Index (LI) by subcortical structure and diagnostic group.

	Thalamus LI		**Amygdala LI**		**Caudate Nucleus LI**	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**	**Mean (SD)**	**Range**
HR-ASD	0.0001 (0.0077)	-0.0178 - 0.0224	0.0003 (0.0305)	-0.0546 - 0.0856	-0.0059 (0.0178)	-0.0603 - 0.0414
HR-LD	0.0019 (0.0046)	-0.0054 - 0.0127	-0.0146 (0.0194)	-0.0517 - 0.03042	-0.0133 (0.0142)	-0.0433 - 0.0256
HR-Neg	0.0012 (0.0060)	-0.0159 - 0.0162	-0.0086 (0.0215)	-0.0843 - 0.0470	-0.0088 (0.0155)	-0.0488 - 0.0371
LR-Neg	0.0024 (0.0071)	-0.0183 - 0.0190	-0.0052 (0.0250)	-0.0664 - 0.0789	-0.0092 (0.0156)	-0.04189 - 0.0341

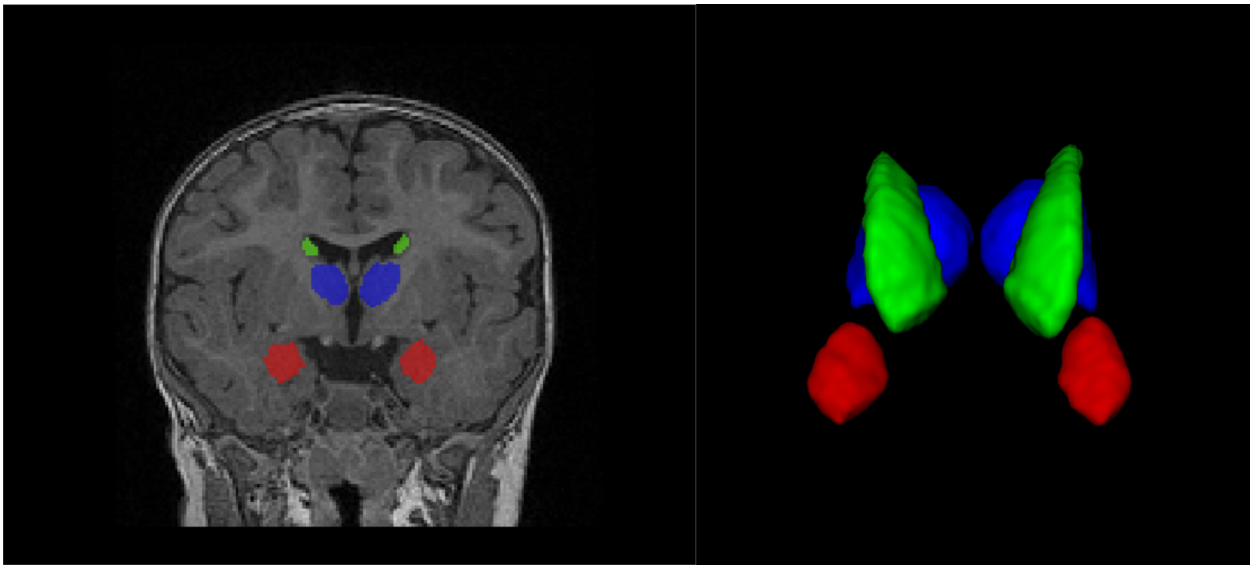

Figure S1. Example segmentation of subcortical brain structures. Coronal slice on left and 3D presentation on the right, showing amygdala in red, thalamus in blue, and caudate nucleus in green.

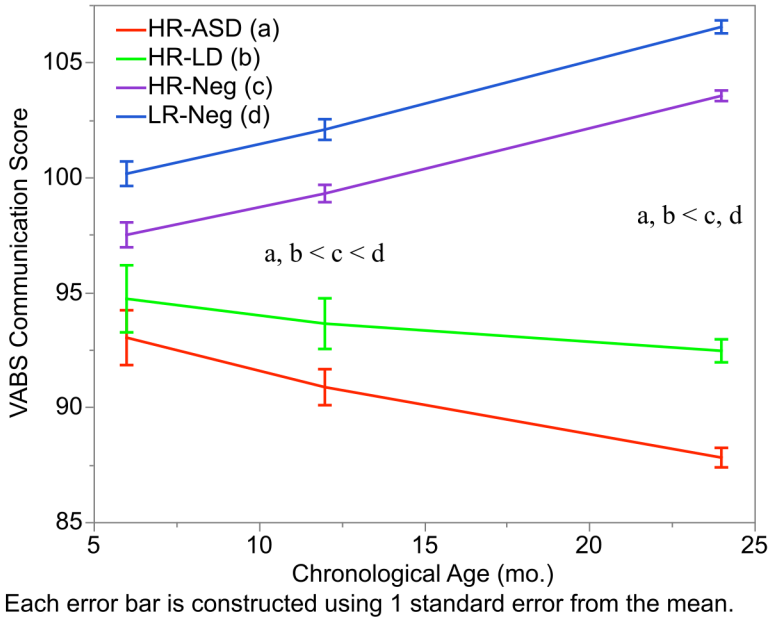

Figure S2. Adaptive language skills are delayed at 12-months in HR-ASD and HR-LD infants and delays were more evident at 24-months. VABS-II Communication standard scores from 6-24 months, *n* data points = 1330. *Note: Contrast legend is as follows: HR-ASD (a), HR-LD (b), HR-Neg (c), and LR-Neg (d). Lines represent LS means which are adjusted for covariates in model (maternal education, clinical site, MSEL NVDQ, and sex of the infant). Error bars = ±1 SEM.*

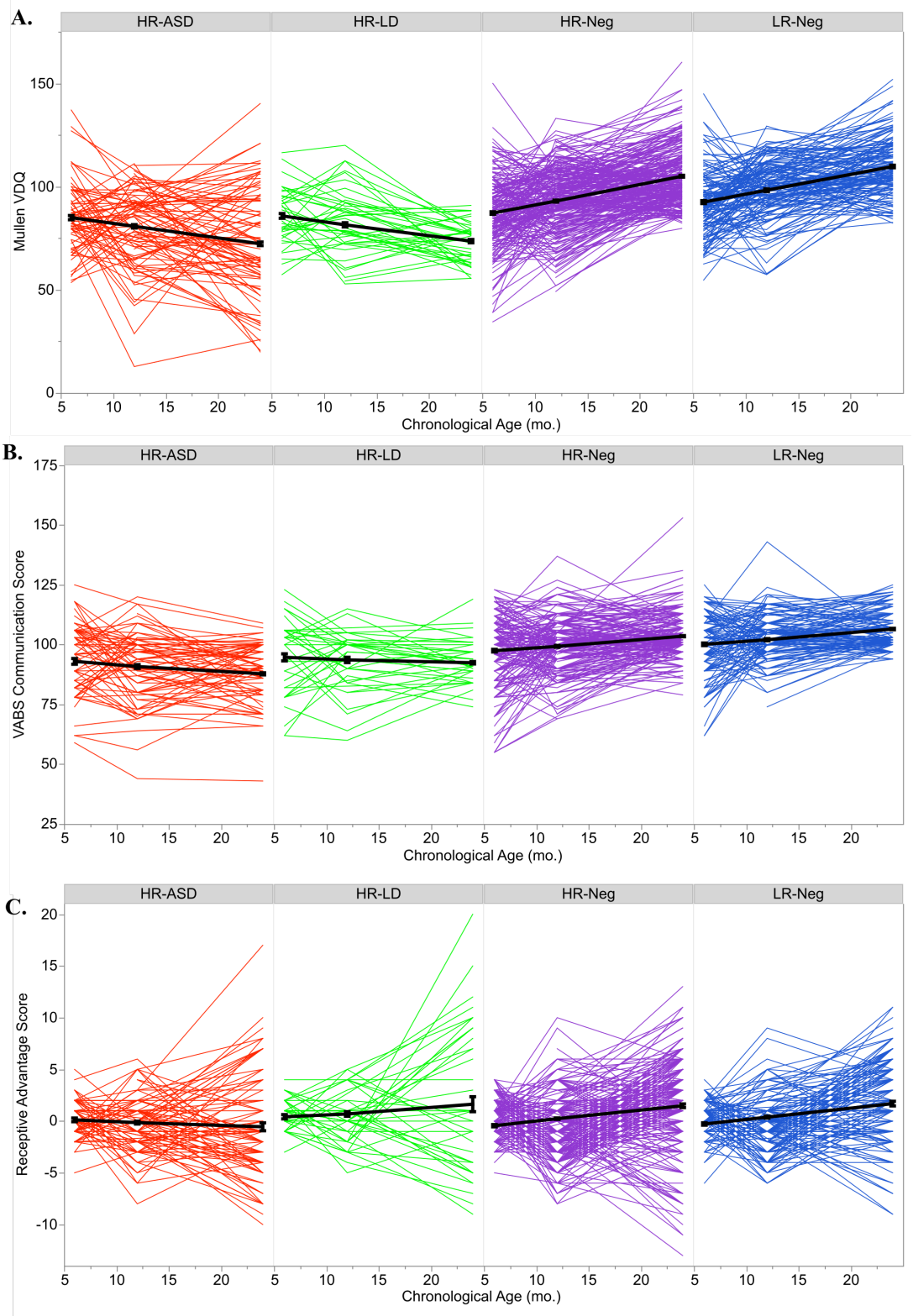

Figure S3. Spaghetti plots show individual trajectories of language development. Panel A, MSEL VDQ from 6-24 months, *n* data points = 1369. Panel B, VABS Communication scores from 6-24 months, *n* data points = 1330. Panel C, receptive advantage scores from 6-24 months, *n* data

points = 1366. *Note: Bold lines represent LS means which are adjusted for covariates in model (maternal education, clinical site, MSEL NVDQ, and sex of the infant). Error bars = ±1 SEM.*

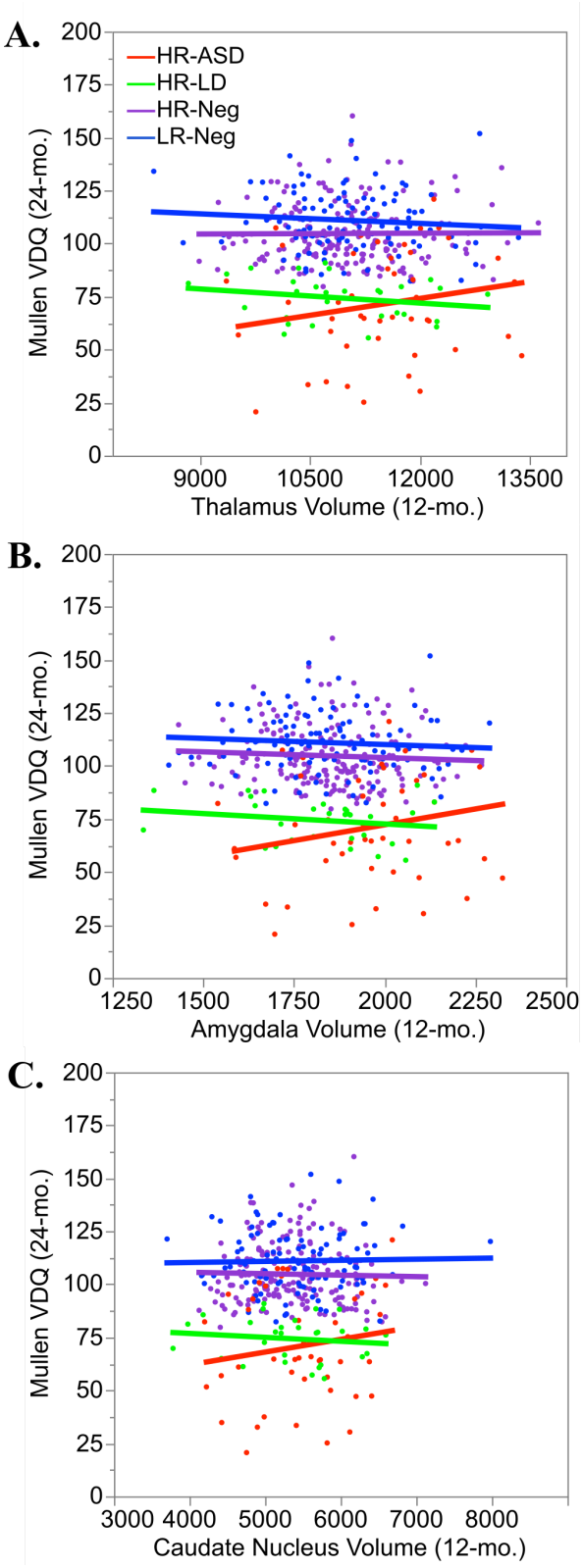

Figure S4. Associations between 12-month bi-lateral subcortical volume and 24-month MSEL VDQ show different brain-behavior associations in the HR-ASD and HR-LD groups. Panel A, association between total thalamus volume (mm^3) and MSEL VDQ (*n* data points = 365),

Panel B, association between total amygdala volume (mm^3) and MSEL VDQ (*n* data points = 365), Panel C, association between total caudate nucleus volume (mm^3) and MSEL VDQ (*n* data points = 365). *Note: Bold lines represent LS means which are adjusted for covariates in model (TCV, age at scan, clinical site, MSEL NVDQ, and sex of the infant).*

Supplemental References

1. Rutter M, Bailey A, Lord C, Cianchetti C, Fancello G (2003): *SCQ: Social Communication Questionnaire: manuale*. Los Angeles, CA: Western Psychological Services.

2. Lord C, Rutter M, Le Couteur A (1994): Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 24: 659–685.

3. Maxwell M (1992): *Family Interview for Genetic Studies (FIGS): a Manual for FIGS*. Bethesda, MD: National Institutes of Mental Health.

4. Gouttard S, Styner M, Prastawa M, Piven J, Gerig G (2008): Assessment of reliability of multi-site neuroimaging via traveling phantom study. *Med Image Comput Comput Interv 2008*. (Vol. 5242), Springer, pp 263–270.

5. Fonov VS, Janke A, Caramanos Z, Arnold DL, Narayanan S, Pike GB, Collins DL (2010): Improved Precision in the Measurement of Longitudinal Global and Regional Volumetric Changes via a Novel MRI Gradient Distortion Characterization and Correction Technique. Springer, Berlin, Heidelberg, pp 324–333.

6. Sled JG, Zijdenbos AP, Evans AC (1998): A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 17: 87–97.

7. Collins DL, Neelin P, Peters TM, Evans AC (1994): Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*. 18: 192–205.

8. Smith SM (2002): Fast robust automated brain extraction. *Hum Brain Mapp*. 17: 143–155.

9. Wang J, Vachet C, Rumple A, Gouttard S, Ouziel C, Perrot E, *et al.* (2014): Multi-atlas segmentation of subcortical brain structures via the AutoSeg software pipeline. *Front Neuroinform*. 8: 7.

10. Gouttard S, Styner M, Joshi S, Smith RG, Cody Hazlett H, Gerig G (2007): Subcortical structure segmentation using probabilistic atlas priors. In: Pluim JPW, Reinhardt JM, editors. *Med Imaging*. International Society for Optics and Photonics, p 65122J–65122J–11.

11. Avants BB, Epstein CL, Grossman M, Gee JC (2008): Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and

neurodegenerative brain. *Med Image Anal*. 12: 26–41.

12. Kandel E, Schwartz J, Jessell T (2000): *Principles of neural science*, 4th ed. New York: McGraw-Hill.

13. Seghier ML (2008): Laterality index in functional MRI: methodological issues. *Magn Reson Imaging*. 26: 594–601.

14. Springer JA, Binder JR, Hammeke TA, Swanson SJ, Frost JA, Bellgowan PSF, *et al.* (1999): Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain*. 122: 2033–2046.

15. Deblaere K, Boon P a, Vandemaele P, Tieleman A, Vonck K, Vingerhoets G, *et al.* (2004): MRI language dominance assessment in epilepsy patients at 1.0 T: region of interest analysis and comparison with intracarotid amytal testing. *Neuroradiology*. 46: 413–420.

16. Hart B, Risley TR (1995): *Meaningful differences in the everyday experience of young American children.* Baltimore: P.H. Brookes.

17. Hochberg Y, Benjamini Y (1990): More powerful procedures for multiple significance testing. *Stat Med*. 9: 811–818.

18. Hazlett HC, Gu H, Munsell BC, Kim SH, Styner M, Wolff JJ, *et al.* (2017): Early brain development in infants at high risk for autism spectrum disorder. *Nature*. 542: 348–351.

19. Shen MD, Kim SH, McKinstry RC, Gu H, Hazlett HC, Nordahl CW, *et al.* (2017): Increased Extra-axial Cerebrospinal Fluid in High-Risk Infants Who Later Develop Autism. *Biol Psychiatry*. . doi: 10.1016/j.biopsych.2017.02.1095.

20. Wolff JJ, Gerig G, Lewis JD, Soda T, Styner MA, Vachet C, *et al.* (2015): Altered corpus callosum morphology associated with autism over the first 2 years of life. *Brain*. 138: 2046–2058.

21. Ruxton GD, Beauchamp G (2008): Time for some a priori thinking about post hoc testing. *Behav Ecol*. 19: 690–693.