Motor stereotypies and volumetric brain alterations in children with Autistic Disorder

Sylvie Goldman[a,b,c,*], **Liam M. O'Brien**[d,e], **Pauline A. Filipek**[f], **Isabelle Rapin**[a,b,c], and **Martha R. Herbert**[g,h]

[a]Saul R. Korey Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, United States

[b]Department of Pediatrics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, United States

[c]Rose F. Kennedy Intellectual and Developmental Disabilities Research Center, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, United States

[d]Department of Mathematics and Statistics, Colby College, 5838 Mayflower Hill, Waterville, ME 04901, United States

[e]College of Graduate Programs in Public Health, University of New England, 716 Stevens Avenue, Portland, ME, United States

[f]Department of Pediatrics, Children's Learning Institute, and the Division of Child and Adolescent Neurology, University of Texas Health Science Center, 7000 Fannin, Suite 2478, Houston, TX 77030, United States

[g]Pediatric Neurology, Martinos Center for Biomedical Imaging, 149 13th Street, 10th Floor, Charlestown, MA 02129, United States

[h]TRANSCEND Research Program, Massachusetts General Hospital, 149 13th Street, 10th Floor, Charlestown, MA 02129, United States

Abstract

Motor stereotypies are defined as patterned, repetitive, purposeless movements. These stigmatizing motor behaviors represent one manifestation of the third core criterion for an Autistic Disorder (AD) diagnosis, and are becoming viewed as potential early markers of autism. Moreover, motor stereotypies might be a tangible expression of the underlying neurobiology of this neurodevelopmental disorder. In this study, we videoscored stereotypies recorded during semi-structured play sessions from school age children with AD. We examined the effect of severity and persistence over time of stereotypies on brain volumetric changes. Our findings confirmed that the brain volume of school age children with AD is, on average, larger than that of age-matched typically developing children. However, we have failed to detect any sign of volumetric differences in brain regions thought to be particularly linked to the pathophysiology of stereotypies. This negative finding may suggest that, at least with respect to motor stereotypies, functional rather than structural alterations might be the underpinning of these disruptive motor manifestations of autism.

*Corresponding author at: Albert Einstein College of Medicine, Kennedy Center, Room 807, 1300 Morris Park Avenue, Bronx, NY 10461, United States. Tel.: +1 718 430 3713.sylviegold@aol.com(S. Goldman).

Keywords

Structural imaging; Volumetric brain analysis; Videotape scoring; Motor stereotypies; Autistic Disorder

1. Introduction

Motor stereotypies are patterned repetitive purposeless movements such as hand flapping, finger twisting, pacing, or rocking. They constitute a subcategory of "restricted repetitive and stereotyped patterns of behavior, interests and activities" which is the third core criterion required for an Autistic Disorder (AD) diagnosis as defined in the Diagnostic and Statistical Manual of Mental Disorders ([DSM IV-TR]; American Psychiatric Association, 2000). These stigmatizing repetitive motor manifestations are receiving increased attention as potential early markers suggestive of an autism diagnosis (Watt, Wetherby, Barber, & Morgan, 2008). Furthermore, given the limited advances made toward unraveling the neurobiology of autism through the study of the cardinal non-motor manifestations such as language disorder and social impairments, a growing number of researchers have turned their attention to motor impairments and stereotypies as a potential source of additional insights into the pathophysiology of autism (Lewis & Kim, 2009).

To pursue these investigations, a variety of functional and structural imaging modalities have been used in children with autism (Amaral, Schumann, & Nordahl, 2008; Schumann, Bauman, & Amaral, 2010; Stigler, McDonald, Anand, Saykin, & McDougle, 2011) with few motor studies focusing on stereotypies (Estes et al., 2011; Hardan, Kilpatrick, Keshavan, & Minshew, 2003; Thakkar et al., 2008). None of these studies has revealed a consistent pattern of neuroanatomical characteristics in children with autism and stereotypies. Some investigators (Hollander et al., 2005; Rojas et al., 2006; Sears et al., 1999) report an increase in caudate nucleus volume, when total brain volume is taken into account. Other magnetic resonance imaging (MRI)-based brain volumetric investigations yielded no consensus between repetitive behaviors and the size of orbitofrontal cortex, anterior cingulate cortex, the basal ganglia, or thalamus, the brain structures most commonly thought to be implicated in repetitive behaviors (Langen, Durston, Staal, Palmen, & van Engeland, 2007; Langen et al., 2009).

There are at least four possible reasons for this lack of correspondence. First, size or structure may not correlate systematically with function or, if there are volumetric alterations, they may be too slight to be resolved with current low resolution morphometric technology. Second, methodologic differences in image acquisition or analysis may mask subtle disparities, given the challenges of recording optimal images from individuals with autism, especially from younger children in whom stereotypies may be particularly prominent. Third, heterogeneity in age, cognitive abilities, and autism severity, together with small sample sizes (Kates, Lanham, & Singer, 2005), jeopardize the likelihood of gathering optimally representative data (Toal et al., 2010). Finally, a lack of comparable measures of stereotypies likely contributes to between-study inconsistencies. There are several validated instruments such as questionnaires and observational scales like the Repetitive Behavior Scale-Revised ([RBS-R]; Lam & Aman, 2007) to measures stereotypies. However, none of these provides standardized measures for the videocoding of stereotypies and thus far no systematic attempt has been made to examine consistency of results across different types of measures such as observations, interviews, and questionnaires (Honey, Rodgers, & McConachie, 2012). Some investigations emphasize behavioral features whereas others focus on motor components (DiGennaro Reed, Hirst, & Hyman, 2012). Rating scales and questionnaires yield less objective and accurate data than scoring of videos which provide

the opportunity for repeated detailed examination by several observers (Goldman et al., 2009). A major limitation of video studies is limited sampling time and the rarity of records obtained under standardized circumstances ideally suited to enhancing the expression and characterization of stereotypies in all or most subjects. These considerations have limited the utility to date of video recordings.

Here we report an analysis of motor stereotypies videoscored using a rigorous approach and correlated with volumetric measures of selected basal ganglia and other relevant brain regions. Subjects were a subset of a larger population of children with a uniformly documented preschool Autistic Disorder (AD) diagnosis ([DSM III-R]; American Psychiatric Association, 1987; Rapin, 1996) in whom MRI data were obtained at school age. In this cohort we had carried out video data analysis of stereotypies in a standardized play setting at preschool (Goldman et al., 2009) and again at school age. Previous studies from school age MRIs in an overlapping subset of children eligible for the present study had yielded brain volumetric data showing significant differences between AD and control subjects in language regions (Herbert et al., 2002), total and regional brain volumes (Herbert et al., 2003), localization of white matter enlargement (Herbert et al., 2004), and brain asymmetries (Herbert et al., 2002, 2005). These studies demonstrated sufficient quality of the morphometric data to justify our exploring the relationship between brain volumes and the presence and severity of motor stereotypies.

Herein, we compared brain volumetric measures in the AD school age sample group with available videocoded stereotypy data to corresponding volumes in a control group of typically developing (TD) children free of stereotypies. We hypothesized that in children with AD, stereotypies scored from videotapes would be associated with volumetric alterations in basal ganglia circuitry, most likely the striatum, as compared with controls.

2. Methods

2.1. Participants

Participants included 61 school age children (M= 9.12 years, SD = 1.27); 31 children with a preschool diagnosis of AD who were matched on chronological age with 30 TD (Table 1). All of the TD subjects were recruited specifically for imaging purposes; they all had normal birth weights, normal developmental histories without seizures or significant head injury. Their screening neurological examination was normal and school performance had never required special help. IQ was not measured. English was the primary language of each child's family. Exclusionary criteria are described in Herbert et al. (2005) and included hearing or gross sensory-motor deficits; clinical evidence of progressive encephalopathy; or frequent seizures. Children were also excluded from imaging if they had any potentially paramagnetic metal, a clinically evident focal brain lesion, brain atrophy, or ventriculomegaly detected by a neuroradiologist who examined each scan and assessed its normalcy. None of the children had been or was medicated. No sedation was used for the scanning. Parents of all the children signed permission for their child's videotaping and participation in the study and the Institutional Review Boards of each of the four participating institutions approved the study (Boston, MA, the Bronx, NY, Cleveland, OH, and Trenton, NJ).

The children with AD were part of a nosological longitudinal study conducted between 1985 and 1992 involving comprehensive behavioral, neurologic, neuropsychologic, and psychiatric evaluations that parsed them into four diagnostic groups based on AD diagnosis and performance cognitive level with preschool subgroups split at nonverbal IQ of 80 (Rapin, 1996).

2.2. Procedure

Videotaping of semi-structured play sessions was an integral part of the comprehensive assessment protocol at each visit of the longitudinal study; MRI scans were only added to the protocol at school age and performed on a randomly selected group of children. Thus, AD subjects in the present study included the subset of the larger cohort who had undergone both MRI scans and videotaping of a school age play session.

2.3. Measures

2.3.1. Diagnostic measures—The children were diagnosed with AD at preschool age according to DSM III-R (American Psychiatric Association, 1987) on the basis of trained psychiatrists' interviews and validated with the parent's responses to the Wing Autistic Disorder Interview Checklist ([WADIC]; Wing, 1996). These children had been recruited from four sites: Boston, MA, the Bronx, NY, Cleveland, OH, and Trenton, NJ and followed up at age 7 and 9 years. Details of recruitment, group assignment, testing procedures, and behavioral measures at preschool have been published (Rapin, 1996).

At school age, AD children's non-verbal performance intellectual quotient (PIQ) was assessed using the Stanford-Binet Fourth Edition ([SB-IV]; Thorndike, Hagel, & Satler, 1986). The SB-IV has strong external validity and high correlations with intelligence scales such as the Wechsler Intelligence Scale for Children-Revised ([WISC-R]; Wechsler, 1974). At the time of imaging, children with AD included in this study had a mean performance intelligence quotient (PIQ) of 85.6 (±28.1).

2.3.2. Intellectual quotient—As indicated above, all 30 TD children fulfilled the criteria for typical controls but no actual IQ measurement for these children was available. The use of standard deviation (SD) cutoffs to categorize cognitive impairment has been previously used (DeRosier, Swick, Davis, McMillen, & Matthews, 2011; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998). In light of these previous studies, our strategy for overcoming the lack of IQ scores when comparing AD subjects to TD has been to construct "IQ status" categories based on the Stanford-Binet average IQ of 100 to which was added (or subtracted) multiples of its SD, that is 15. Accordingly, all TD control subjects were given a score of 0, since they were all typically developing children and thus their IQ scores had to fall within 100 ± 15. Using the IQ scores available for the AD subjects, the children were given an "IQ status" with a value of +1 for IQ = 115–129 (mean + 1 × SD to <mean + 2 × SD); +2 for IQ = 130–144 (mean + 2 × SD to <mean + 3 × SD), etc. Similarly, they were given a value of −1 for IQ = 85–71 (mean − 1 × SD to >mean − 2 × SD), −2 for IQ of 70–56 (mean − 2 × SD and >mean − 3 × SD), etc.

The scale shown in Table 2 allowed us to use our available IQ data in making direct comparisons between TD and AD subjects. This strategy was only necessary when comparing AD subjects to TD directly since we had the actual IQ scores for the within-AD analyses.

2.3.3. Measures of motor stereotypies—Stereotypies were coded from 30-min videotaped semi-structured play sessions. All play sessions started with a 5-min problem solving game where the child was instructed by the same trained examiner at each site to utilize different tools to get prizes through a large Plexiglas maze. After that short period the examiner engaged in interactive play using a similar script and sets of toys adapted to school age. The amount of time the child and experimenter played with each toy varied somewhat, depending on the child's social abilities, cognitive functioning, and play skills. All sessions were videotaped from an adjacent observation room through a one-way mirror.

Each stereotypy defined as repetitive purposeless patterned motor behavior (Sanger et al., 2010) seen at least twice (to document its repetitiveness) during the first 15 min of the 30-min semi-structured videotaped play session, was characterized and counted. For all stereotypies, the parts of the body involved as well as the characteristics were scored, so that each repetitive movement was described and assigned to one of eight discrete mutually exclusive subtypes as described in Goldman et al. (2009). Three trained observers achieved inter-rater coding reliability (Kappa 0.8).

2.3.4. Stereotypies grouping—Two types of group assignments were tested based on the videoscoring of the stereotypies (Goldman et al., 2009). AD subjects were divided into subgroups twice, each time by a different set of criteria. The first subgrouping was by number of recorded stereotypies: (a) 0, (b) 1–10, and (c) greater than 10. The second subgrouping was by persistence versus loss of stereotypies over time, since all AD children had documented stereotypies at preschool: (a) stereotypies present at preschool but no longer at school age, and (b) stereotypies still present at school age. None of the TD children had stereotypies, so their brain volumes were not included in these particular analyses. Table 3 presents the number of subjects included in each of these analyses.

To mitigate the impact of missing data on stereotypies due to two missing videotapes, we substituted the videoscoring by the information from the neurologic examination and the responses from the WADIC parent questionnaire. We have previously shown (Goldman et al., 2009) that videoscoring, neurologic observation, and parent questionnaire data on stereotypies are highly correlated.

2.4. MRI data

2.4.1. Image acquisition—MRI was performed on either General Electric 1.5 T Signa (Milwaukee, WI, USA) or Siemens 1.5 T Magnetom (Iselin, NJ, USA) systems. Scanning was performed between 1989 and 1992. On the GE system, volumetric acquisition parameters were: pulse sequence = 3D-SPGR or 3D-CAPRY, repetition time (TR) = 34–50 ms; echo time (TE) = 5–9 ms, flip angle = 45–50, field of view (FOV) = 24–26 cm, slice thickness = 3.0–3.1 mm, number of slices = 60 contiguous, matrix = 256 3 256, number of excitations = 1. On Siemens systems, volumetric acquisition parameters were: pulse sequence = 3D-FLASH, TR = 40 ms, TE = 10 ms, flip angle = 40, FOV = 30 cm, slice thickness = 3.1 mm, number of slices = 60 contiguous, matrix = 256 3 256, number of excitations = 1. Images on the two systems were found to be comparable for quantitative segmentation analysis (Filipek, Kennedy, & Caviness, 1991).

2.4.2. Image analysis—Imaging data were analyzed on Sun Microsystems (Mountain View, CA, USA) workstations. Using methods previously described (Kennedy et al., 1994), the initial image data set was normalized with respect to Talairach stereotactic space (Talairach & Tournoux, 1988), minimizing the need for precise uniformity of head position at the time of imaging. Neuroanatomical segmentation, dividing the brain into gray matter and white matter subdivisions, was performed using semi-automated algorithms based upon intensity contour mapping and differential intensity contour algorithms that have been described previously (Filipek et al., 1989; Filipek, Richelme, Kennedy, & Caviness, 1994; Kennedy et al., 1994). Cerebral cortex–white matter distinctions were segmented in a semi-automated fashion, while deep gray nuclei were delineated manually.

Segmentation was performed between 1990 and 1993 by five raters, who met laboratory standards of reliability (Filipek et al., 1994; Seidman et al., 1999). The neocortical ribbon was then parcellated into 48 primarily gyral-based parcellation units (PUs) per hemisphere, according to a procedure described previously (Caviness, Kennedy, Richelme, Rademacher,

& Filipek, 1996; Kennedy et al., 1998; Rademacher, Galaburda, Kennedy, Filipek, & Caviness, 1992).

Briefly, sulcal patterns were identified and labeled by a neuroanatomically trained rater, on multiplanar orthogonal views allowing the sulcal markers to be tracked three-dimensionally. In addition, anatomical markers for other mainly anterior–posterior divisions of larger gyral units were identified. A canonical set of PUs was then identified and labeled with a colorcode system. Cortical parcellation was performed between 1995 and 1998 by four raters who met the previously reported laboratory standards for inter-rater reliability for this method (Caviness et al., 1996). Volumes for each neuroanatomical unit were derived by summing the voxels within that segmented brain region or cortical PU.

2.4.3. Choice of brain regions pertinent to neuroanatomy of stereotypies—
Although the quantitative volumetric analyses were performed on the entire brain, including segmentation of gray and white matter and parcellation of cortex and white matter, and while prior publications (Herbert et al., 2003, 2004, 2005) presented analyses that worked with all these units of analysis, here we worked with a subset of brain regions of interest (ROIs) chosen because of their pertinence to the neuroanatomical substrate of stereotypies. These ROIs included four cortical regions (supplementary motor cortex, fronto-orbital cortex, dorsolateral prefrontal cortex [DLPFC], and anterior cingulate gyrus) and four subcortical areas (caudate, putamen, globus pallidus, and diencephalon [thalamus plus hypothalamus, which were not divided from each other in this dataset]). Volumes for each ROI in this analysis are total volumes represented by the sum of right and left hemispheres.

2.4.4. Use of raw volumes for inter-group and gender-specific standardized volumes for intra-group comparisons

a. TD versus AD: As described for Analysis 1 in Table 4, raw volumes were used in comparisons between TD controls (none of whom had stereotypies) and AD subjects (all of whom had stereotypies at preschool) since they allow the most direct volumetric comparison between these two groups. These comparisons were performed by relating each ROI volume to "IQ status" and a stereotypy grouping variable, described above, via a regression modeling approach. Total brain volume was also included as a covariate to control for differences in overall brain size.

b. Comparisons within AD: For Analyses 2 and 3 in Table 4, which involved comparisons of AD stereotypy groupings, we used standardized volumes. Standardization of brain volumes is done by calculating a z-score expressing the difference between an AD subject's brain volume and the mean control volume. The z-score is calculated as follows: (i) subtract the mean ROI volume for controls from the ROI volume of each individual in the study group, and (ii) divide the resulting number by the standard deviation of the control group. The z-score number is unique for each individual and within that individual for each of their ROIs. It represents the number of standard deviations by which that individual's ROI differs from the control mean for that ROI. This standardization into z-scores thus allows each individual to be compared with the control group. An additional benefit to using standardized z-score volumes is that it frees us to use the actual PIQ scores available for this group as covariate in the regression model, since we are comparing brain volumes of study group with controls without the need to covary for IQ within the comparison. Standardization thus allows us to perform within-AD group assessments of whether AD subjects' differences from control volumes are different by stereotypy grouping.

Since our AD group included both genders, we constructed the standardized volumes separately for males and females because of the well-known brain volume differences between the genders. We did this so that in our analyses we would be deriving AD group differences relative to the difference of each AD subject from a comparable control subject. We performed this gender-specific standardization by calculating z-scores within gender. That is, mean of controls was calculated by gender rather than for the entire mixed-gender control group—so that for each of the male subjects, the mean ROI volume calculated from all control boys was subtracted from each individual AD boy's ROI volume, and then this difference was divided by the standard deviation of the control boys' ROI volumes. The same procedure was applied within the female group. Once calculated separately by gender, these standardized volumes could then be grouped to include both genders and this group of standardized volumes could then be related to stereotypy grouping variables, along with the actual performance IQ at school age, using a linear regression modeling strategy as it was in the AD versus TD comparisons. Total brain volume was also used as a covariate in within-AD comparisons because standardization was only done with respect to the average control volumes for each ROI, not for the whole brain, which leaves open the possibility that there may still be effects in the standardized volumes due to differences in overall brain size (O'Brien et al., 2011).

2.5. Statistical analysis

We performed three sets of regression analyses, one between TD and AD groups and two within-AD groups; these analyses are summarized in Table 4. The TD-AD comparison (Analysis 1, details in Table 5) compared raw brain volumes. The first set of within-AD analyses (Analysis 2, details in Table 6) addressed the number of stereotypies by using the subgrouping based on number of stereotypies. The second set of within-AD analyses (Analysis 3, details in Table 7) addressed the temporal changes in stereotypies by using the subgrouping that distinguished the children who had stereotypies at preschool that went away at school age and those children who retained stereotypies at the later age.

Each of these regression analyses generated coefficients indicating the expected change in standardized brain volumes relative to a unit change in another variable (i.e., a stereotypy or an IQ variable), holding all other covariates constant. For both Analyses 1 and 3 in Table 4, only one coefficient was generated (see analyses results detailed in Tables 5 and 6), since each involved only two subgroups and therefore only one pairwise comparison. However for Analysis 2 in Table 4, which involved not two but three stereotypies subgroupings (number of stereotypies) two group coefficients were generated (see Table 6), to allow an omnibus test of volumetric differences among all groups under consideration. Further details of these coefficients are explained in the table captions for Tables 5–7 respectively. The p-value for each comparison indicates whether or not there was a significant difference in regional brain volumes between the subgroups. All statistical analyses were performed with the statistical software Statistical Analysis System 9.2 ([SAS]; SAS Institute Inc., Cary, NC), and regression modeling techniques were primarily used.

3. Results

3.1. Analysis of demographic variables

To investigate possible demographic differences when making group comparisons, we performed a series of a priori analyses. First, in regard to ethnicity, based on previous reports published about this cohort the two AD and TD groups did not differ in ethnicity distribution. The available data about the ethnicity of the AD group shows that 70% of the participants were Caucasian (Rapin, 1996). Since so far no published data on ethnicity has suggested difference in brain volume we did not covary for ethnicity. Second, in regard to

age, we found no significant difference in age between the TD controls and AD subjects (t-test, $t(59) = 047$, $p = .963$). The age differences between the AD subjects who had stereotypies at school age versus those who did not also did not achieve significance at the .05 level (t-test, $t(29) = 1.33$, $p = .195$); nor were there significant differences among the three stereotypy severity groupings (one-way ANOVA, $F(2, 58) = 2.53$, $p = .088$).

Third, as for gender, since TD controls were evenly split by gender but the AD group had the expected much higher proportion of boys, there was a significant gender difference between TD and AD subjects (Pearson chi-square test, $x^2(1) = 7.94$, $p = .005$). Thus, we examined the possibility of a gender effect on the TD versus AD analyses. In these analyses, a gender covariate was included in addition to the other covariates of interest. We also examined the gender effect for the within-AD analyses. We found no significant gender association with groups defined by the severity of stereotypies (Fisher's exact test, $p = 1.00$) or for groups defined by stereotypy persistence (Fisher's exact test, $p = .116$).

3.2. Between-group comparisons

3.2.1. Correlations of autism and stereotypies with total brain volume and PIQ
—Prior to comparing the volume of the selected cortical and subcortical neuroanatomical units of interest between children with AD and TD controls, we examined the total brain volume in these two groups. This analysis revealed that children with AD had a significantly larger mean total brain volume (1400 ± 144 cm^3) compared to controls (1308 ± 110 cm^3, Student's t-test, $t(59) = -2.76$, $p = .008$), as previously reported for a partially overlapping subset of this population comprised only of children with PIQ > 80 (Herbert et al., 2003).

3.2.2. Analysis 1: volumetric comparisons of anatomical units between AD and TD children—The general linear regression model used to compare regional mean volumes of the anatomical units did not detect any significant difference in the total (left plus right) volume of any stereotypy-pertinent ROI in children with AD compared to TD controls when controlling for PIQ status (see Analysis 1, Table 5). We found no gender effect on these results (data not shown). The TD versus AD analyses included the "IQ status" covariate, while the within-AD analyses included the actual performance IQ values for each subject as covariates. All analyses included total brain volume to adjust for differences in overall brain sizes.

3.3. Within-group comparisons

3.3.1. PIQ and stereotypy—To assess the relationship between PIQ and stereotypy, we utilized a Pearson correlation. Consistent with our previous observation (Goldman et al., 2009), in the AD group we found a marginally significant correlation between lower PIQ and greater number of stereotypies ($r = .34$, $p = .054$).

3.3.2. Analysis 2: lack of relationship between stereotypy number and volumetric alterations—In this analysis we examined whether the severity of stereotypies present at school age had any impact on the anatomical unit volume differences. We found that in the school age AD group 17 children had no stereotypy, six had one to ten stereotypies, and eight had more than ten stereotypies during the first 15 min of the videotaped play sessions. Statistical analysis revealed no difference among any of the brain ROIs with respect to severity (i.e., number) of stereotypy (Table 6).

3.3.3. Analysis 3: lack of volumetric differences in children with transient as compared to persistent stereotypies—All of the children with AD enrolled in this study had stereotypies at preschool, but of these 31 children, only 45% (14 of 31) still exhibited stereotypies at school age, meaning that 17 of these children no longer had

stereotypies that we observed or were reported in the WADIC parent questionnaire. As described in Goldman et al. (2009) the stereotypies observed at preschool involved mostly both arms, including hands and fingers (e.g., flapping, clapping) and often involved handling objects. Among the 14 school age children in whom stereotypies persisted, 10 children showed a pattern of motor repetitive behaviors that was unchanged from the pattern each such individual had shown in preschool. In four of the 14 children (28%) the pattern did change; such that rocking and pacing persisted but the arm and finger movements decreased or disappeared.

While it is of interest to determine whether brain volumes were impacted over time by whether the stereotypies did or did not persist in these children, MRI scanning only became available for the school age phase of the study (1989–1992); consequently we have no preschool imaging data to compare with our imaging of the children when they reached school age. We addressed this question instead by examining whether the group of children with persistent stereotypies showed volumetric difference compared to those whose stereotypies did not persist into school age. This comparison yielded no difference between these subgroups (Table 7).

4. Discussion

This study examines the relationship of neuroanatomical volumes to stereotypies in autism. We found that the total brain volumes of children with AD were larger than those of typically developing children with neither autism nor stereotypies. We also found that in AD, lower PIQ was marginally associated with more stereotypies. However, in contrast with previous brain imaging studies of subjects from the same cohort which had reported significant morphometric differences in brain regions relevant to language (Herbert et al., 2003, 2002, 2005), we failed to detect any volumetric differences among regions of the brain thought to be relevant to stereotypies, such as the basal ganglia. Moreover, we did not detect brain volumetric differences that correlated with loss versus persistence of stereotypies between preschool and school age assessments.

How should our negative results be interpreted in the context of other morphometric studies (Estes et al., 2011; Hardan et al., 2003; Sears et al., 1999; Thakkar et al., 2008) that report a variety of alterations that are inconsistent with each other and often based upon small sample sizes? According to the limitations of our study discussed below, we propose that the results should be characterized as "failure to detect" rather than "negative findings" or "absence of difference." Our MRI anatomic data were acquired early in the MRI era when image resolution was lower and before functional imaging and tractography became available. We were limited to structural data as a proxy or indirect measure relevant to function. Our failure to detect any anatomical difference associated with the presence or absence of stereotypies at school age does not exclude the possibility that we might find such anatomical differences in a larger sample or with more up-to-date volumetric imaging techniques. It also does not exclude the possibility of underlying functional differences that simply do not produce a macroanatomically detectable impact. On the other hand, if our failure to detect an anatomical difference reflects an actual lack of such a difference, perhaps the influence of the underlying biology on factors affecting cell size or density was too slight in this cohort to cause aggregate volume changes measurable via MRI volumetrics. The underlying biology might include alterations in neurotransmitters/neuromodulators or in synaptic networks not associated with changes in actual neuronal or glial size or number, but rather in the neural circuitry.

There are reports of neurologic conditions, for example streptococcal infection, in which stereotypies fluctuate and may be associated with changes in volumes of the basal ganglia

(Wolf & Singer, 2008). Group analyses might wash out statistically significantly increased volumes if children at different phases of such an illness were included. But reports of such phenomena are anecdotal (Giedd, Rapoport, Garvey, Perlmutter, & Swedo, 2000) and there is no convincing evidence that the stereotypies of autism have an encephalopathic basis. Moreover, the decrease in prevalence of stereotypies with age in our cohort suggests not so much an intermittent process as a dynamic one that tends to resolve over time. It remains to be seen whether maturation, adaptation to genetic or environmental causes of stereotypies, or development of capacities for behavioral self-control or inhibition explain the waning. Reduction in the older children of arm and finger movements but not rocking and pacing is tantalizing as a possible clue to pertinent biology, but explanation awaits availability of more refined anatomical and other neurobiological data. At least with respect to abnormal repetitive behaviors, functional rather than structural alteration may be the underpinning of these disrupting motor manifestations of autism.

Failure to detect volumetric differences in this carefully designed and implemented study brings up some critical issues inherent to autism research, some of which are responsible for the divergent and inconclusive reports in the literature. Frequent shortcomings in autism research include small sample sizes, heterogeneity in AD phenotypes, lack of means to parse people with autism according to biological etiologies, and lack of standardized tools to quantify motor manifestations such as stereotypies. Such flaws contribute to preventing investigators from interpreting positive or negative results, including those of this study, in a definitive manner.

4.1. Limitations

Several limitations to this study need to be considered. First, given our sample size, a real consideration is insufficient statistical power, at least for some analyses, to detect significant volumetric differences among subgroups. Although the comparisons between the AD and TD subjects would have had adequate power to detect a difference (78.5% probability of detecting a difference of 91 cm^3 in 61 children), the analyses involving only the 31 AD subjects were significantly less powered. For instance, the probability of detecting a mean volume difference of 90 cm^3 between AD subjects with one to ten stereotypies versus those with more than ten was only 26.8% given our sample sizes. Similarly, the correlation of 0.34 between number of stereotypies and PIQ, assuming it is the population correlation, had only 47.8% chance of being detected. This makes the marginally significant $p = .054$ value obtained encouraging but not robust.

Second, another possible explanation of failure to detect volume abnormality might relate various methodological issues related to performing volumetric analyses. The anatomic parcellation units selected might not have conformed to the functionally relevant anatomical structures. Affected brain areas could have had overlapped regional boundaries or might have occupied only a portion of a morphometrically defined anatomical unit. In this case, volumetric differences could have gotten washed out by the lack of difference in that unit's other sub-parts, or could have had different shapes but not different volumes among subjects. Brain differences could also have involved differences in tissue or metabolic composition that might not affect volumes significantly. More recent imaging methods such as shape analysis (Casanova et al., 2011; Qiu, Adler, Crocetti, Miller, & Mostofsky, 2010), Diffusion Tensor Imaging (DTI), or spectroscopy (Langen et al., 2012; Sivaswamy et al., 2010) offer alternate windows into the anatomical substrates of these movements. A third limitation to this study is the use of the "IQ status" variable. Since no IQ information was available for control subjects, they all received the presumed score of 0. However, it is possible that some TD controls had IQ scores outside the range of 85–115 and thus should have received different status scores. Although this is a legitimate concern, given the small

sample size of this cohort, it is quite unlikely that having these scores would have changed our results.

A fourth limitation pertains to the fact that, while detailed quantitative videocoding of stereotypies is unique to this study, it does not guarantee a representative sample of the stereotypies of each child. Fifteen minutes is a small time sample. Finally, our choice to split the continuous variable "number of stereotypies" into a categorical variable with three levels (i.e., subjects with (1) 0, (2) 1–10 and (3) greater than 10 stereotypies during the 15 min of videocoding) is defensible; it is a reasonable way to render the great variability in the number of stereotypies more tractable. However, the arbitrary nature of this grouping may have watered down any potential meaningful relationship between stereotypy severity and brain volume changes.

It is also possible that other studies which relied on questionnaires rather than standardized scoring might have yielded less detail and precision but perhaps more reliable information on prevalence. Automated technologies for investigating movement abnormalities, such as wearable accelerometers, could overcome this concern in future studies.

4.2. Future research

Going forward and in view of these limitations, we propose that future research make every effort to include a larger cohort with equal numbers of girls and boys matched on ethnicity and IQ. Given that no significant stereotypy-related volumetric alteration at school age was found in the present cohort, we also suggest that imaging modalities aimed at assessing function such as Diffusion Tensor Imaging (DTI), magnetic resonance spectroscopy (MRS), or EEG coherence and not just or not primarily anatomical size of brain regions may be valuable to probe the potential link between stereotypies and changes in specific regions, pathways or networks. Furthermore, it is possible that the relationship between repetitive behaviors and regional brain differences might be detectable by comparing imaging analyses over time. Thus future study may benefit from longitudinal rather than cross-sectional design.

As for the behavioral assessment of stereotypies, it is expected that computerized pattern recognition of videocoded stereotypies through body sensors may yield more densely quantifiable data. Therefore future studies combining functional imaging and emerging quantitative behavioral technologies should be considered.

Acknowledgments

References

Amaral DG, Schumann CM, Nordahl CW. Neuroanatomy of autism. Trends in Neurosciences. 2008; 31(3):137–145. [PubMed: 18258309]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders, DSM III-R. 3rd Rev. ed.. Washington, DC: American Psychiatric Association; 1987.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders, DSM IV-TR. 4th Rev. ed.. Washington, DC: American Psychiatric Association; 2000.

Casanova MF, El-Baz A, Elnakib A, Switala AE, Williams EL, Williams DL, et al. Quantitative analysis of the shape of the corpus callosum in patients with autism and comparison individuals. Autism. 2011; 15(2):223–238. [PubMed: 21363871]

Caviness VS Jr, Kennedy DN, Richelme C, Rademacher J, Filipek PA. The human brain age 7–11 years: A volumetric analysis based on magnetic resonance images. Cerebral Cortex. 1996; 6(5): 726–736. [PubMed: 8921207]

DeRosier ME, Swick DC, Davis NO, McMillen JS, Matthews R. The efficacy of a Social Skills Group Intervention for improving social behaviors in children with High Functioning Autism Spectrum disorders. Journal of Autism and Developmental Disorders. 2011; 41(8):1033–1043. [PubMed: 21042870]

DiGennaro Reed FD, Hirst JM, Hyman SR. Assessment and treatment of stereotypic behavior in children with autism and other developmental disabilities: A thirty year review. Research in Autism Spectrum Disorders. 2012; 6(1):422–430.

Estes A, Shaw DW, Sparks BF, Friedman S, Giedd JN, Dawson G, et al. Basal ganglia morphometry and repetitive behavior in young children with autism spectrum disorder. Autism Research. 2011; 4(3):212–220. [PubMed: 21480545]

Filipek PA, Kennedy DN, Caviness VS Jr. Volumetric analyses of central nervous system neoplasm based on MRI. Pediatric Neurology. 1991; 7(5):347–351. [PubMed: 1764136]

Filipek PA, Kennedy DN, Caviness VS Jr, Rossnick SL, Spraggins TA, Starewicz PM. Magnetic resonance imaging-based brain morphometry: Development and application to normal subjects. Annals of Neurology. 1989; 25(1):61–67. [PubMed: 2643919]

Filipek PA, Richelme C, Kennedy DN, Caviness VS Jr. The young adult human brain: An MRI-based morphometric analysis. Cerebral Cortex. 1994; 4(4):344–360. [PubMed: 7950308]

Giedd JN, Rapoport JL, Garvey MA, Perlmutter S, Swedo SE. MRI assessment of children with obsessive–compulsive disorder or tics associated with streptococcal infection. American Journal of Psychiatry. 2000; 157(2):281–283. [PubMed: 10671403]

Goldman S, Wang C, Salgado MW, Greene PE, Kim M, Rapin I. Motor stereotypies in children with autism and other developmental disorders. Developmental Medicine and Child Neurology. 2009; 51(1):30–38. [PubMed: 19087102]

Hardan AY, Kilpatrick M, Keshavan MS, Minshew NJ. Motor performance and anatomic magnetic resonance imaging (MRI) of the basal ganglia in autism. Journal of Child Neurology. 2003; 18(5): 317–324. [PubMed: 12822815]

Herbert MR, Harris GJ, Adrien KT, Ziegler DA, Makris N, Kennedy DN, et al. Abnormal asymmetry in language association cortex in autism. Annals of Neurology. 2002; 52(5):588–596. [PubMed: 12402256]

Herbert MR, Ziegler DA, Deutsch CK, O'Brien LM, Kennedy DN, Filipek PA, et al. Brain asymmetries in autism and developmental language disorder: A nested whole-brain analysis. Brain. 2005; 128(Pt 1):213–226. [PubMed: 15563515]

Herbert MR, Ziegler DA, Deutsch CK, O'Brien LM, Lange N, Bakardjiev A, et al. Dissociations of cerebral cortex, subcortical and cerebral white matter volumes in autistic boys. Brain. 2003; 126(Pt 5):1182–1192. [PubMed: 12690057]

Herbert MR, Ziegler DA, Makris N, Filipek PA, Kemper TL, Normandin JJ, et al. Localization of white matter volume increase in autism and developmental language disorder. Annals of Neurology. 2004; 55(4):530–540. [PubMed: 15048892]

Hollander E, Anagnostou E, Chaplin W, Esposito K, Haznedar MM, Licalzi E, et al. Striatal volume on magnetic resonance imaging and repetitive behaviors in autism. Biological Psychiatry. 2005; 58(3):226–232. [PubMed: 15939406]

Honey E, Rodgers J, McConachie H. Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: Selecting a questionnaire or interview. Research in Autism Spectrum Disorders. 2012; 6(1):757–776.

Kates WR, Lanham DC, Singer HS. Frontal white matter reductions in healthy males with complex stereotypies. Pediatric Neurology. 2005; 32(2):109–112. [PubMed: 15664770]

Kennedy DN, Lange N, Makris N, Bates J, Meyer J, Caviness VS Jr. Gyri of the human neocortex: An MRI-based analysis of volume and variance. Cerebral Cortex. 1998; 8(4):372–384. [PubMed: 9651132]

Kennedy, DN.; Meyer, KA.; Filipek, PA.; Caviness, VS.; John, ER.; Huerta, M. Functional neuroimaging: Technical foundations. San Diego, CA: Academic Press; 1994. MRI-based topographic segmentation; p. 201-208.

Lam KS, Aman MG. The Repetitive Behavior Scale-Revised: Independent validation in individuals with autism spectrum disorders. Journal of Autism and Developmental Disorders. 2007; 37(5): 855–866. [PubMed: 17048092]

Langen M, Durston S, Staal WG, Palmen SJ, van Engeland H. Caudate nucleus is enlarged in high-functioning medication-naive subjects with autism. Biological Psychiatry. 2007; 62(3):262–266. [PubMed: 17224135]

Langen M, Leemans A, Johnston P, Ecker C, Daly E, Murphy CM, et al. Fronto-striatal circuitry and inhibitory control in autism: Findings from diffusion tensor imaging tractography. Cortex. 2012; 48(2):183–193. [PubMed: 21718979]

Langen M, Schnack HG, Nederveen H, Bos D, Lahuis BE, de Jonge MV, et al. Changes in the developmental trajectories of striatum in autism. Biological Psychiatry. 2009; 66(4):327–333. [PubMed: 19423078]

Lewis MH, Kim SJ. The pathophysiology of restricted repetitive behavior. Journal of Neurodevelopmental Disorders. 2009; 1(1):114–132. [PubMed: 21547711]

O'Brien LM, Ziegler DA, Deutsch CK, Frazier JA, Herbert MR, Locascio JJ. Statistical adjustments for brain size in volumetric neuroimaging studies: Some practical implications in methods. Psychiatry Research. 2011; 193(2):113–122. [PubMed: 21684724]

Qiu A, Adler M, Crocetti D, Miller MI, Mostofsky SH. Basal ganglia shapes predict social, communication, and motor dysfunctions in boys with autism spectrum disorder. Journal of the American Academy of Child and Adolescent Psychiatry. 2010; 49(6):539–551. [PubMed: 20494264]

Rademacher J, Galaburda AM, Kennedy DN, Filipek P, Caviness VS. Human cerebral cortex: Localization, parcellation, and morphometry with magnetic resonance imaging. Journal of Cognitive Neurosciences. 1992; 4:352–374.

Rapin, I. Developmental language disorder, autism, low IQ. Clinics in developmental medicine ed.. London: Mac Keith; 1996. Preschool children with inadequate communication.

Rojas DC, Peterson E, Winterrowd E, Reite ML, Rogers SJ, Tregellas JR. Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. BMC Psychiatry. 2006; 6:56. [PubMed: 17166273]

Sanger TD, Chen D, Fehlings DL, Hallett M, Lang AE, Mink JW, et al. Definition and classification of hyperkinetic movements in childhood. Movement Disorders. 2010; 25(11):1538–1549. [PubMed: 20589866]

Schumann CM, Bauman MD, Amaral DG. Abnormal structure or function of the amygdala is a common component of neurodevelopmental disorders. Neuropsychologia. 2010; 49(4):745–759. [PubMed: 20950634]

Sears LL, Vest C, Mohamed S, Bailey J, Ranson BJ, Piven J. An MRI study of the basal ganglia in autism. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 1999; 23(4):613–624. [PubMed: 10390720]

Seidman LJ, Faraone SV, Goldstein JM, Goodman JM, Kremen WS, Toomey R, et al. Thalamic and amygdala-hippocampal volume reductions in first-degree relatives of patients with schizophrenia:

An MRI-based morphometric analysis. Biological Psychiatry. 1999; 46(7):941–954. [PubMed: 10509177]

Sivaswamy L, Kumar A, Rajan D, Behen M, Muzik O, Chugani D, et al. A diffusion tensor imaging study of the cerebellar pathways in children with autism spectrum disorder. Journal of Child Neurology. 2010; 25(10):1223–1231. [PubMed: 20179000]

Stigler KA, McDonald BC, Anand A, Saykin AJ, McDougle CJ. Structural and functional magnetic resonance imaging of autism spectrum disorders. Brain Research. 2011; 1380:146–161. [PubMed: 21130750]

Stothard SE, Snowling MJ, Bishop DV, Chipchase BB, Kaplan CA. Language-impaired preschoolers: A follow-up into adolescence. Journal of Speech Language and Hearing Research. 1998; 41(2): 407–418.

Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme; 1988.

Thakkar KN, Polli FE, Joseph RM, Tuch DS, Hadjikhani N, Barton JJ, et al. Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). Brain. 2008; 131(Pt 9):2464–2478. [PubMed: 18550622]

Thorndike, RL.; Hagel, EP.; Satler, JM. Stanford-Binet intelligence scale. 4th ed.. Chicago, IL: Riverside Publishing Co; 1986.

Toal F, Daly EM, Page L, Deeley Q, Hallahan B, Bloemen O, et al. Clinical and anatomical heterogeneity in autistic spectrum disorder: A structural MRI study. Psychological Medicine. 2010; 40(7):1171–1181. [PubMed: 19891805]

Watt N, Wetherby AM, Barber A, Morgan L. Repetitive and stereotyped behaviors in children with autism spectrum disorders in the second year of life. Journal of Autism and Developmental Disorders. 2008; 38(8):1518–1533. [PubMed: 18266099]

Wechsler, D. Wechsler intelligence scale for children-revised. 4th ed.. New York: Psychological Corporation; 1974.

Wing, L. Wing autistic diagnostic interview checklist. In: Rapin, I., editor. Clinics in developmental medicine No. 139: Preschool children with inadequate communication. London: Mac Keith; 1996. p. 247-251.

Wolf DS, Singer HS. Pediatric Movement Disorders: an update. Current Opinion in Neurology. 2008; 21(4):491–496. [PubMed: 18607212]

Table 1

Demographic information.

	Group	
	Autistic Disorder	**Typically developing**
Number of subjects	31	30
Gender (boys/girls)	26/5	15/15
Ethnicity	70% Caucasian	NA
Mean age at MRI (years) (SD)	9.13 (1.36)	9.11 (1.19)

NA, not available.

Table 2

Transformation of IQ scores into IQ status for Autistic Disorder (AD) and typically developing (TD) groups.

Group	IQ scores	IQ status	n
AD	25–39	−4	5
AD	40–54	−3	1
AD	55–69	−2	2
AD	70–84	−1	4
TD	85–115	0	30
AD	116–130	1	1
AD	131–145	2	0
AD	146–160	3	17

Table 3

Subject numbers by analysis.

	Total AD	Total TD	AD with no stereotypies at school age	AD with some stereotypies at school age	AD with 1–10 stereotypies at school age	AD with >10 stereotypies at school age	Total in each analysis
Analysis 1	31	30					61
Typically developing versus AD							
Analysis 2			17		6	8	31
AD with no stereotypies at school age versus AD with 1–10 stereotypies at school age versus AD with more than 10 stereotypies at school age							
Analysis 3			17	14			31
AD with no stereotypies at school age versus AD with stereotypies at school age							

Table 4

Table of analyses performed.

Group comparison of primary interest	Type of outcome measure	Type of statistical model	Factors controlled for in model	Results
(1) Typically developing versus AD	Raw volumes (cm^3)	Multiple linear regression	Performance IQ (measured as number of SDs from normal)	Total brain volume (cm^3)	Table 5
(2) AD with no stereotypies at school age versus AD with 1–10 stereotypies at school age versus AD with more than 10 stereotypies at school age	Standardized volumes representing each subject's deviation from a TD subject of the same gender, measured in number of SDs	Multiple linear regression	Actual performance IQ	Total brain volume (cm^3)	Table 6
(3) AD with no stereotypies at school age versus AD with stereotypies at school age	Standardized volumes representing each subject's deviation from a TD subject of the same gender, measured in number of SDs	Multiple linear regression	Actual performance IQ	Total brain volume (cm^3)	Table 7

Table 5

Analysis 1: volumetric comparisons between school age children with Autistic Disorder ($n = 31$) and typically developing children controlling for IQ status ($n = 30$).

	Total raw regional volume controlled for total brain volume	
	Group effect coefficient	Group effect p value
Cortical areas		
Supplemental motor cortex	2.102	0.352
Fronto-orbital cortex	−6.366	0.101
DLPFC	−8.785	0.320
Anterior cingulate gyrus	−0.468	0.874
Subcortical areas		
Caudate	−3.150	0.026
Putamen	−0.328	0.739
Globus pallidus	−0.925	0.103
Diencephalon	0.548	0.735

Anatomical unit volumes were compared between groups by a multiple regression linear model.

Group effect coefficient indicates the expected change in volume (in cubic centimeters, cm^3) for an autistic subject relative to a control for a given IQ status and total brain volume. A positive coefficient means a larger volume in AD subjects and a negative mean a larger volume in controls.

DLPFC, dorsolateral prefrontal cortex; diencephalon, thalamus + hypothalamus.

Table 6

Analysis 2: comparisons of standardized brain volumes in relation to number of stereotypies among school age children with Autistic Disorder. No stereotypies ($n = 17$), 1–10 stereotypies ($n = 6$) or >10 stereotypies ($n = 8$).

	Total standardized volume controlled for total brain volume					
	Group 1 coefficient	Group 1 p-value	Group 2 coefficient	Group 2 p-value	IQ coefficient	IQ coefficient p-value
Cortical areas						
Supplemental motor cortex	0.988	0.185	1.280	0.159	−0.019	0.096
Fronto-orbital cortex	0.253	0.622	0.479	0.614	0.001	0.917
DLPFC	0.646	0.206	0.707	0.253	−0.008	0.313
Anterior cingulate gyrus	0.163	0.771	−0.327	0.632	−0.005	0.598
Subcortical areas						
Caudate	−0.360	0.633	−0.108	0.907	0.012	0.291
Putamen	−0.779	0.151	−0.811	0.222	0.014	0.097
Globus pallidus	0.706	0.374	1.038	0.291	0.017	0.151
Diencephalon	0.556	0.499	−0.906	0.375	−0.006	0.632

Anatomical unit volumes were compared between groups by a multiple regression linear model.

Group 1 effect coefficient indicates the expected change in standardized volume (cm^3) for an AD subject with no stereotypies at school age relative to an AD subject with more than 10 stereotypies at school age. A positive coefficient indicates larger volumes in AD subjects with no stereotypies and a negative coefficient indicates larger volumes in AD subjects with more than 10 stereotypies at school age.

Group 2 effect coefficient indicates the expected change in standardized volume (cm^3) for an AD subject with 1–10 stereotypies at school age relative to an AD subject with more than 10 stereotypies at school age. A positive coefficient indicates larger volumes in AD subjects with 1–10 stereotypies and a negative coefficient indicates larger volumes in AD subjects with more than 10 stereotypies at school age.

IQ *coefficient* indicates the expected increase in standardized volume for a 1-unit increase in IQ score within a stereotypy grouping.

DLPFC, dorsolateral prefrontal cortex; diencephalon, thalamus + hypothalamus.

Table 7

Analysis 3: comparisons of standardized brain volumes in relation to temporal changes: loss versus persistence of stereotypies in children with Autistic Disorder. Stereotypies at both preschool age and at school age ($n =$ 17) versus stereotypies at preschool but no longer at school age ($n = 14$).

	Total standardized volume controlled for total brain volume			
	Group effect coefficient	Group effect *p*-value	IQ coefficient	IQ *p*-value
Cortical areas				
Supplemental motor cortex	−0.478	0.443	0.019	0.126
Fronto-orbital cortex	0.443	0.249	0.004	0.544
DLPFC	0.171	0.644	−0.003	0.702
Anterior cingulate gyrus	−0.412	0.433	−0.004	0.637
Subcortical areas				
Caudate	0.194	0.763	0.012	0.291
Putamen	0.217	0.658	0.013	0.141
Globus pallidus	−0.359	0.602	0.019	0.126
Diencephalon	−0.992	0.151	−0.004	0.738

Anatomical unit volumes were compared between groups by a multiple regression linear model.

Group effect coefficient indicates the expected change in standardized volume (cm^3) for an AD subject with stereotypies at school age relative to an AD subject without stereotypies at school age. A positive coefficient indicates larger volumes in AD subjects with stereotypies and a negative coefficient indicates larger volumes in AD subjects without stereotypies at school age.

IQ coefficient indicates the expected increase in standardized volume for a 1-unit increase in IQ score within a stereotypy grouping.

DLPFC, dorsolateral prefrontal cortex; diencephalon, thalamus + hypothalamus.