Resting-state functional connectivity predicts longitudinal change in autistic traits and adaptive functioning in autism

Mark Plitt[a,1], Kelly Anne Barnes[a], Gregory L. Wallace[a,b], Lauren Kenworthy[a,c], and Alex Martin[a,1]

[a]Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892; [b]Department of Speech and Hearing Sciences, George Washington University, Washington, DC 20052; and [c]Center for Autism Spectrum Disorders, Children's National Health System, Washington, DC 20010

Although typically identified in early childhood, the social communication symptoms and adaptive behavior deficits that are characteristic of autism spectrum disorder (ASD) persist throughout the lifespan. Despite this persistence, even individuals without cooccurring intellectual disability show substantial heterogeneity in outcomes. Previous studies have found various behavioral assessments [such as intelligence quotient (IQ), early language ability, and baseline autistic traits and adaptive behavior scores] to be predictive of outcome, but most of the variance in functioning remains unexplained by such factors. In this study, we investigated to what extent functional brain connectivity measures obtained from resting-state functional connectivity MRI (rs-fcMRI) could predict the variance left unexplained by age and behavior (follow-up latency and baseline autistic traits and adaptive behavior scores) in two measures of outcome—adaptive behaviors and autistic traits at least 1 y postscan (mean follow-up latency = 2 y, 10 mo). We found that connectivity involving the so-called salience network (SN), default-mode network (DMN), and frontoparietal task control network (FPTCN) was highly predictive of future autistic traits and the change in autistic traits and adaptive behavior over the same time period. Furthermore, functional connectivity involving the SN, which is predominantly composed of the anterior insula and the dorsal anterior cingulate, predicted reliable improvement in adaptive behaviors with 100% sensitivity and 70.59% precision. From rs-fcMRI data, our study successfully predicted heterogeneity in outcomes for individuals with ASD that was unaccounted for by simple behavioral metrics and provides unique evidence for networks underlying long-term symptom abatement.

rs-fMRI | autism | machine learning | ASD | adaptive behavior

Although typically identified in childhood, the social communication symptoms that are characteristic of autism spectrum disorder (ASD) persist throughout the lifespan (1, 2). On average, individuals with ASD show smaller age-related improvements in adaptive behaviors, including daily living skills critical for independent living, than do typically developing (TD) peers (2–4). The burden of prolonged clinical symptom expression, coupled with limited adaptive behaviors, leads to a relatively poor prognosis for a majority of adults with ASD. For example, only 12% of adults with ASD achieve "very good" outcomes, defined by a high level of independence (5). Adolescence and young adulthood are poorly understood in ASD. Although there seems to be a great deal of change during this time, the nature of this change varies across studies, with a handful of studies reporting a decline in functioning (2, 3, 6), others reporting general improvement (7, 8), and still others reporting a quadratic course of autistic symptoms and adaptive functioning where the trajectory peaks in late adolescence (9) or the late 20s (10) and begins to fall subsequently.

Predictors of positive outcomes in ASD include higher intelligence quotient (IQ) (11–13), language ability (9, 14), less severe ASD symptoms (15), and stronger adaptive behaviors (16, 17). However, there is substantial variability in outcome even among

individuals with ASD without cooccurring intellectual disability (7, 13, 17, 18). Age, IQ, and language ability accounted for as much as 45% of the variance in outcome measures in a sample composed of predominantly individuals with both ASD and intellectual disability (11). Others reported more modest numbers for these predictors, with IQ predicting 3% of variance in outcome and language ability predicting 32% of variance in outcome (14). A study that included only individuals with ASD without cooccurring intellectual disability found age and IQ as weaker predictors, predicting 6–28% of various adaptive behavior subscales (6). Although these previous studies have been successful in predicting these outcomes using behavioral measures, in most cases, the majority of the variance in outcomes remains unexplained. Thus, it remains difficult to identify individuals with ASD who may struggle to achieve independence during adulthood and who may benefit from additional intervention.

In the present study, we explored whether a functional neuroimaging-based measure of brain connectivity, termed resting-state functional connectivity MRI (rs-fcMRI), can predict variance in behavioral outcomes in young adults with ASD beyond that explained by cognitive or behavioral measures. Functional connectivity strength in individuals with ASD has been found to predict an ASD diagnosis (19–21) and to correlate with many aspects of cognition and behavior that also predict outcome, including IQ (21, 22), and ASD symptomatology using the Autism

---

Significance

Few individuals diagnosed with autism spectrum disorder (ASD) go on to achieve high levels of independence or what are considered "very good" outcomes. As such, there is a need to identify predictors of outcomes to improve treatment and services for these individuals. Although behavioral and cognitive variables can predict substantial variance in outcomes, the majority of the variance remains unexplained. In this study, we investigated whether a measure of intrinsic functional brain connectivity [resting-state functional connectivity MRI (rs-fcMRI)] could add meaningful predictive power. Indeed, we found that connectivity involving several brain networks previously implicated in ASD could predict improvements in adaptive behaviors several years after the scan with a high degree of sensitivity.

Diagnostic Observation Schedule (21–23), the Autism Diagnostic Interview-Revised (20), and the Social Responsiveness Scale (SRS) (19, 22, 24). As such, brain measures may explain additional variance in behavioral outcomes. One previous study has shown that combining functional MRI (fMRI) data with behavioral data increased predictive power for categorical language outcomes in early developing ASD (25). Brain-derived data has also added explanatory power to predictive models of depression (26), dyslexia (27), alcoholism (28), and reading and math ability (29, 30).

We tested whether rs-fcMRI data acquired in late adolescence and early adulthood [time 1 (T1)] could predict behavioral outcomes at least 1 y after the imaging data were acquired [time 2 (T2)]. We defined behavioral outcomes with a measure of predominantly social autistic traits (SRS) and a measure of adaptive functioning [Adaptive Behavior Assessment System-Second Edition (ABAS-II)]. Using an approach that controlled for nuisance variables (e.g., variable duration between time 1 and time 2) and variables known to strongly predict outcome (e.g., age and baseline score on outcome measure), we performed regressions to investigate whether and to what extent the remaining variance in outcome could be predicted by baseline functional connectivity in networks known to be involved in ASD.

Results

The mean follow-up latency between the behavioral measures was 2 y, 10 mo and 2 y, 11 mo for the ABAS ($n = 27$) and SRS ($n = 29$), respectively. ABAS and SRS scores both generally improved from time 1 (ABAS mean = 71.96, SRS mean = 75.89) to time 2 (ABAS mean = 81.93, SRS mean = 71.97) (Fig. 1A). An increase in ABAS score indicates an improvement in adaptive behaviors; a decrease in SRS score indicates an improvement in autistic symptoms. The difference in ABAS scores was significant ($P < 0.001$), but the difference in SRS scores was not ($P = 0.14$). Age and follow-up latency accounted for 23.01% of the variance

in T2 ABAS and 43.15% of variance in T2 SRS. When controlling for these effects using an analysis of covariance (ANCOVA), T2 ABAS scores were still significantly different from T1 ABAS scores ($P < 0.001$), and the improvement in SRS scores remained nonsignificant ($P = 0.14$). Age, follow-up latency, and T1 score accounted for an initial 26.05% of variance in ΔABAS and 40.63% of variance in ΔSRS (Fig. 1 B and C). Although large amounts of variance in outcome were explained by these variables, the majority of variance in outcome remained unexplained. ΔSRS and ΔABAS scores shown in Fig. 1 have been converted to reliable change indices (RCIs), allowing us to estimate a confidence interval of true change. The RCI takes into account the SE of the measurement obtained from test–retest reliability and yields a pseudo–z-statistic. Thus, if |RCI| ≥ 1.96 for a particular measure, we can say that the individual showed a reliable change on that measure beyond the two-tailed 95% confidence interval. For ΔABAS, 12 out of 27 individuals (44.44%) showed RCIs beyond the 95% confidence interval, and, for ΔSRS, 12 out of 29 individuals (41.38%) had RCIs beyond the 95% confidence interval. Given the prevalence of significant RCIs in our sample, we next sought to investigate the predictive power of the rs-fcMRI data.

Rs-fcMRI scans were acquired at time 1, and, for each subject, we obtained a region X region correlation matrix (considered a measure of "intrinsic functional connectivity") using 264 regions of interest (ROIs) (31). These regions were previously characterized as belonging to 1 of 13 functional networks, and we have previously shown that 3 of the 13 [default-mode network (DMN), frontal-parietal task control network (FPTCN), and salience network (SN)] were particularly informative for distinguishing individuals with ASD from TD individuals (19) (see *Materials and Methods* for a full description). To increase statistical power and reduce overfitting, we restricted our analysis to these three networks. We examined, separately, whether baseline (time 1) functional connectivity within a network or across a pair of networks
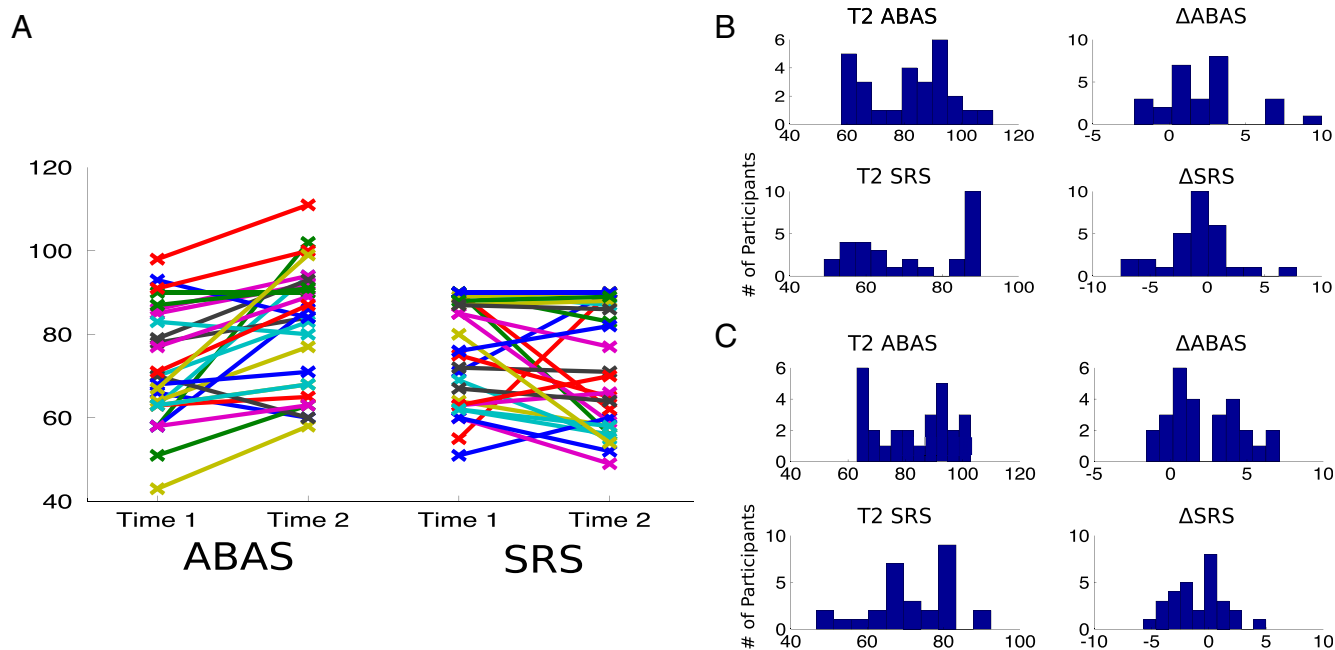
Fig. 1. Outcome measures show change over time. (*A*) Adaptive Behavior Assessment System (ABAS) General Adaptive Composite standard scores (*Left*; $n = 27$) and Social Responsiveness Scale (SRS) sum t-scores (*Right*; $n = 29$) are shown for each participant at each time point. Higher ABAS scores indicate an improvement in adaptive behaviors whereas lower SRS scores indicate improvement in social functioning. (*B*) Histograms representing the distribution of scores for each of the four outcome measures. ΔSRS and ΔABAS (the change between time 1 and time 2 SRS and ABAS, respectively) have been converted to reliable change indices. (*C*) The same outcome measures are shown after linear effects of baseline score (ΔSRS and ΔABAS only), age, and follow-up latency have been removed. Note that these outcome measures retain considerable variability after this initial regression. These final measures are those that are predicted by the brain-based statistical models (Table 1 and Fig. 2).

could predict any of the remaining variance in behavioral outcomes at time 2. It is important to note that the amount of variance explained above by the behavioral models was determined by an r^2 value obtained from typical multiple linear regression. Due to concerns about overfitting, the brain-based regression results will be presented as empirically determined r^2 values from leave-one-out cross-validation (LOOCV) with ridge regression. The amount of variance explained using these two methods (multiple linear regression vs. ridge regression with LOOCV) is not directly comparable. In addition, the r^2 values presented for the brain-based regressions will be presented as the amount of variance explained in the residual outcome measures and not the amount of total variance explained in the uncorrected outcome measures.

Using LOOCV with ridge regression, rs-fcMRI data from time 1 significantly predicted additional variance in three of the four behavioral outcome measures (time 2 SRS, ΔSRS, and ΔABAS, but not time 2 ABAS). The full results from the brain-based regressions are reported in Table 1. Time 2 SRS scores were best predicted by connectivity within the frontoparietal task control network (FPTCN–FPTCN connectivity, $r^2 = 0.1439$, $P < 0.03$). ΔSRS was best explained by intra-FPTCN connectivity ($r^2 = 0.3351$, $P < 0.001$). Finally, ΔABAS was best explained by intra-SN connectivity ($r^2 = 0.3181$, $P < 0.001$). ΔABAS was also significantly predicted by DMN–SN connectivity ($r^2 = 0.3062$, $P < 0.001$) and DMN–FPTCN connectivity ($r^2 = 0.0679$, $P < 0.003$). Scatter plots showing the predicted vs. actual outcome scores are shown in Fig. 2. None of the rs-fcMRI regressions improved when other nuisance variables, such as IQ or the number of time points censored due to motion or global signal intensity, were included as additional features. Measures of comorbid anxiety, depression, and attention deficit hyperactivity disorder (ADHD) as assessed by the Child Behavior Check List (32) or the Adult Behavior Check List (33) before the scan were available for all but one of the participants. Including the presence of any comorbidities as additional features in the regression did not improve predictive accuracy.

An important follow-up question to the above analysis is whether these models are accurately predicting the reliable changes in ABAS

and SRS scores present in >40% of individuals or whether they are predicting only the fluctuations within the SE of the measurement. To answer this question, we turn to the regressions for the ΔSRS and ΔABAS scores.

As stated above, 12 out of 27 individuals showed RCIs beyond the 95% confidence interval (|RCI| ≥ 1.96) for the ABAS. In all cases, this change was an improvement (increase) in ABAS score. Green dashed lines in Fig. 2 denote the threshold for significant improvement, and green data points represent participants that passed this threshold. For several regressions, the brain data were successful in predicting this reliable improvement. Intra-SN connections and SN–DMN connections predicted a reliable improvement in ABAS with 100% sensitivity (70.59%/66.67% and 63.16%/53.33% precision/specificity, respectively). Sensitivity can be understood as the ability of the statistical model to correctly identify all of the true positives (significant RCI values). Precision, also called positive predictive value, is a complementary measure defined as the proportion of positive estimates by the model that are true positives and can be understood as the ability of the model to not label a scan as positive when the true label is negative (nonsignificant RCI value). Lastly, specificity can be understood as the ability of the model to correctly identify all of the true negatives. DMN–FPTCN connections predicted ABAS improvement with 83.33% sensitivity (62.50% precision; 60.00% specificity). For ΔSRS, 12 out of 29 individuals had RCIs beyond the 95% confidence interval. Nine of the 12 individuals showed a reliable decrease in SRS scores (i.e., decrease in autistic traits; green data points in Fig. 2). Intra-FPTCN predicted this change with only 44.44% sensitivity (80.00% precision; 95% specificity). Three of the 12 individuals showed a reliable increase in SRS scores (i.e., increase in autistic traits; red data points in Fig. 2), but intra-FPTCN connectivity failed to correctly predict decline in each case.

Discussion

Brain-based statistical models predicted significant variance in clinically relevant outcome measures for individuals with ASD. Furthermore, this variance was unique from and beyond that which is explained by age, regression to the mean over time, and duration between testing times. On the timescale of several years, we observed that a measurable number of the individuals with ASD in our sample displayed significant change in behavioral measures of predominantly social autistic traits (SRS) or adaptive functioning (ABAS). Simple measures of functional connectivity obtained at rest using rs-fcMRI at the time of the first behavioral assessment significantly predicted both SRS scores at a later date (time 2) and the change in SRS and ABAS scores over time (ΔSRS and ΔABAS). Furthermore, several of the rs-fcMRI regression models showed high sensitivity (83–100%) in predicting reliable improvement in ABAS scores.

Predicting Outcomes in ASD. In line with previous studies, we observed a general tendency for core measures of impairment in ASD to show modest improvements during adolescence and early adulthood (4, 7, 8, 34–36) [for a review, see Seltzer et al. (18)]. In our own study, we found that time 2 ABAS scores but not time 2 SRS scores were significantly improved from mean time 1 scores. In contrast to our findings, several recent studies have found that samples of individuals with ASD without intellectual disability actually show an age-related decline in adaptive behaviors as measured by the Vineland Adaptive Behavior Scales (VABS) (2, 3, 6, 37). This difference could be due to our choice of the General Adaptive Composite scores from the ABAS rather than the VABS, or that we used longitudinal data rather than performing a cross-sectional study. Szatmari et al. (9) additionally reported a quadratic shape to adaptive behavior and autism symptoms with a peak in functioning in late adolescence (9). Taylor and Seltzer (10) report related findings with the peak occurring in the late 20s (10).

Table 1. Amount of residual variance explained (r^2) for each outcome measure using functional connections between functional networks

Outcome measures	DMN	FPTCN	SN
T2 SRS			
DMN	0	0	0
FPTCN	—	0.1439*	0
SN	—	—	0
T2 ABAS			
DMN	0	0	0.017
FPTCN	—	0	0
SN	—	—	0
ΔSRS			
DMN	0.0228	0	0
FPTCN	—	0.3351*	0
SN	—	—	0
ΔABAS			
DMN	0	0.0679*	0.3062*
FPTCN	—	0	0
SN	—	—	0.3181*

Rows and columns designate the networks from which features were derived (e.g., DMN–DMN designates all intra-DMN connections, and DMN–FPTCN designates all pairwise across-network connections between DMN and FPTCN). *$P < 0.005$. Only the upper triangle of the symmetric result matrix is shown.
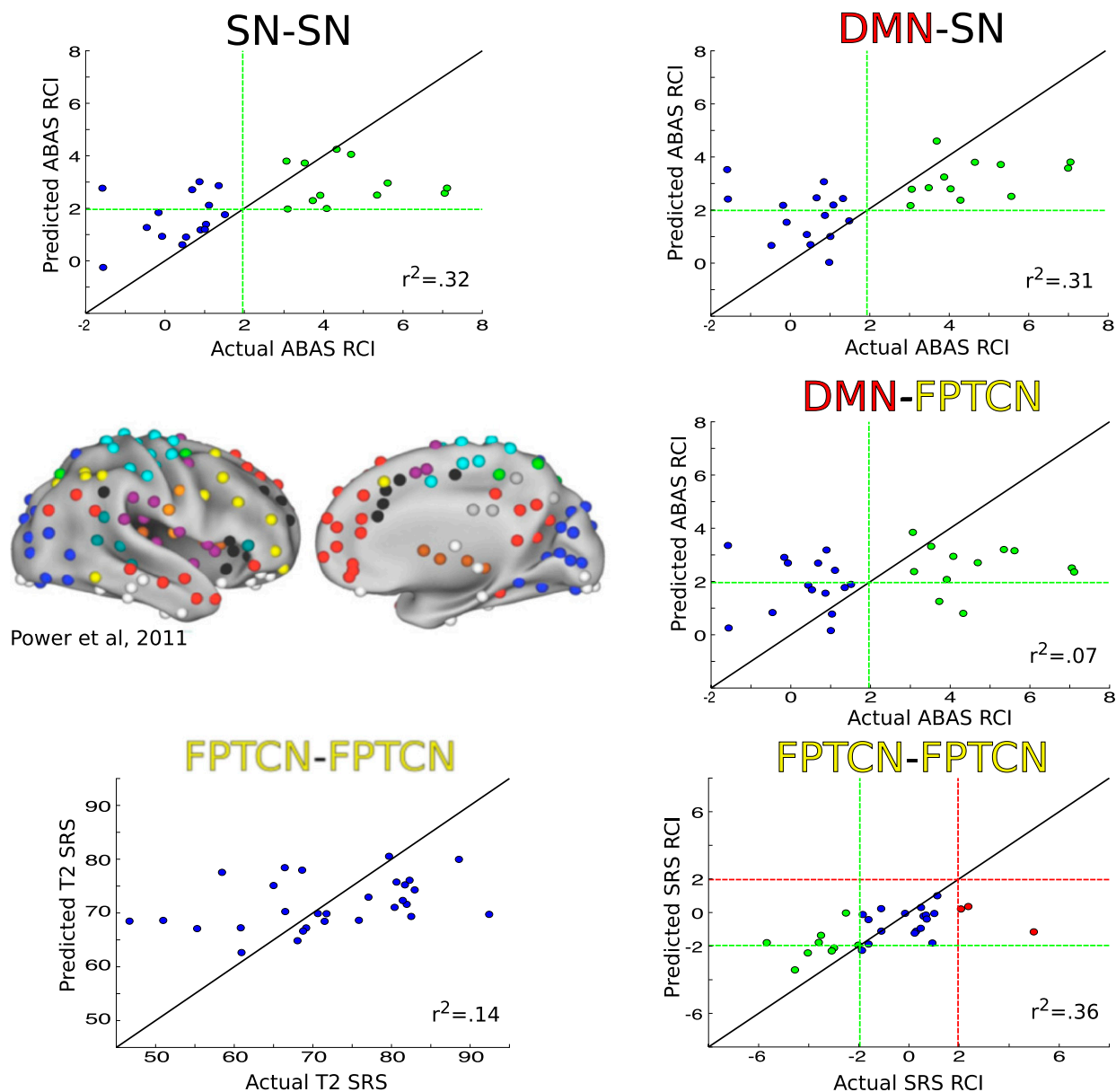
Fig. 2. Functional connectivity involving three networks previously defined in Power et al. (31) (*Center Left*) including the salience network (SN, black), default-mode network (DMN, red), and the frontoparietal task control network (FPTCN, yellow), predict significant variance in ΔABAS (*Top Left*, *Top Right*, and *Center Right*; n = 27), ΔSRS (*Bottom Right*; n = 29), and T2SRS (*Bottom Left*; n = 29). In all scatter plots, the true outcome value is on the horizontal axis and the predicted outcome value is on the vertical axis. The line of equality, representing a perfect prediction, is shown as a solid black line. For ΔABAS and ΔSRS plots, green dashed lines represent the thresholds for a reliable change index (RCI) beyond the two-sided 95% confidence interval of the measure in the direction of symptom improvement. Green data points are those that have an actual RCI beyond this threshold. For ΔSRS, red dashed lines represent the two-sided 95% confidence interval threshold for worsening of symptoms, and red data points are those that have an actual RCI beyond the threshold. The r^2 values represent residual variance explained after accounting for behavioral measures (see *Results*). (*Center Left*) Reproduced from ref. 31, with permission from Elsevier.

Also in concordance with previous studies, we observed substantial heterogeneity in individuals' trajectories (13, 17, 18). Approximately 44% of participants in our study showed significant change in ABAS scores, and 41% showed significant change in SRS scores.

Previous attempts to predict outcomes for adolescents with ASD from measurements of behavior have found several factors to be important. IQ and early language ability are perhaps the most consistent predictors of outcome in studies that included individuals with and without cooccurring intellectual disability (1, 2, 5, 6, 11, 12, 14–16, 18, 37, 38). However, the range of variance explained by IQ and language ability varies widely. Billstedt

et al. (11), for example, found that IQ, sex, and attainment of speech before age 5 y of age predicted 45% of variance in a measure of social interaction (11). Howlin et al. (14), on the other hand, found that IQ predicted only 3% of variance in a composite outcome score whereas language ability as measured by the Peabody Picture Vocabulary Test predicted 32% of variance in outcome (14). Studies of individuals with ASD without intellectual disability found that IQ may be a weak predictor of adaptive behavior as well (3, 37, 39). In line with the latter studies, we found that including IQ as an additional feature in our brain-based regression models did not improve outcome prediction. Our study included exclusively males without an intellectual disability. If we had had a

more heterogeneous sample (e.g., including females and individuals with intellectual disability), we might have found IQ and gender to be strong predictors as well.

Our strongest results were not from predicting future scores but rather from predicting the change in scores over time. We found that baseline SRS, age, and follow-up latency explained almost 41% of the variance in ΔSRS. After accounting for those variables, rs-fcMRI data were able to predict nearly 34% of the remaining variance or roughly 20% of the total variance in ΔSRS. We know of only one other study that has tried to predict the magnitude of change in SRS scores longitudinally. Constantino et al. (36) collected SRS scores at two time points and found that, whereas the two scores had a high intraclass correlation, a linear regression model using baseline SRS, age, and familial loading of Pervasive Developmental Disorder predicted only 11% of variance in the change in SRS scores over time. The lower r^2 value in their study is likely due to the much more heterogeneous sample in their analysis, including children. We also found that baseline scores, age, and follow-up latency predicted a significant portion (26%) of the variance in change in adaptive behaviors, ΔABAS. Functional connectivity data then predicted about 32% of the remaining variance, equating to roughly 23.5% of the total variance. To our knowledge, no others have attempted to directly predict the magnitude of change of adaptive behaviors. Several studies have modeled the trajectories of adaptive behaviors over time using various numbers of groups and functions (13, 40), and one study predicted the rate of change in vocational and educational activities in adults with ASD (41). The aim of these studies was often to identify predictors of trajectory group membership or identify the best model among many alternatives. Consequently, the performance of the regressions is rarely put in terms of variance explained. A related study, however, did directly predict change in maladaptive behaviors over time. Shattuck et al. (42) found that baseline measures predicted 36% of variance in the change in maladaptive behaviors and that age, intellectual disability, language ability, and sex accounted for an additional 7% of the total variance (42). Together, our results indicate that functional brain imaging data provide a marked increase in predictive power over using behavioral data alone.

A recent study of ASD outcomes by Lombardo et al. (25) corroborates this claim. The authors showed that fMRI data collected during an auditory language task on infants and toddlers, who later developed ASD, predicted a categorical language outcome 1 y postscan [68% accuracy, area under the curve (AUC), 69%]. The predictive power of the fMRI measures was comparable with that of time 1 behavioral measures (68% accuracy, 69% AUC), and combining the two modalities provided the best predictive power (80% accuracy, 81% AUC). Another study also reported significantly predicting ASD symptom severity using rs-fcMRI, but they did not report the amount of variance explained by their models (43). Previous studies with other populations, including individuals with depression (26), dyslexia (27), alcoholism (28), and academic difficulties (29, 30), have further shown that brain-based features can add additional predictive power to behavioral measures. For example, using a similar cross-validation procedure, Supekar et al. (29) were able to predict individuals' responses to math tutoring above and beyond the variance explained by behavior using structural and functional MRI data collected before tutoring. Hippocampal gray matter volume predicted 20.25% of the variance in individuals' math improvement, and connectivity between the basal ganglia and hippocampus predicted 56.25% of the variance. Myers et al. (30) found that white matter morphometry, in conjunction with baseline behavioral measures, predicted 56% of the variance in reading impairment at a later date.

Functional Networks That Predicted Outcome. To aid in feature selection, the networks that we investigated (SN, DMN, and FPTCN) were selected a priori as those that distinguish individuals with ASD from TD individuals (19). We focused foremost on creating restricted but highly predictive models rather than conducting an unbiased search for the most informative features. Such an exploratory analysis is an important future direction for this line of work. It is possible that other networks or another combination of networks could have performed similarly or better. Nonetheless, our approach yielded valuable information regarding the predictive power of specific networks.

The SN, in particular the anterior insula and anterior cingulate, has been repeatedly shown to be involved in ASD (19, 20, 24, 44). This network is thought to integrate bottom-up attention and top-down executive control to mediate interactions between other large-scale networks, such as the FPTCN and DMN, and to assign value to stimuli for later processing (45). The nodes of the SN have been implicated in disparate tasks ranging from social processing (46–48) to purely cognitive tasks (49). Intra-SN connections and SN–DMN connections were highly predictive of change in ABAS general adaptive composite (GAC) score ($r^2 > 0.3$)—a domain-general measure of adaptive behavior.

The DMN, which includes the anterior and posterior temporal lobes, the temporoparietal junction, the medial prefrontal cortex, and the precuneus, has been thought of as a "social brain" network involved in theory of mind and self-referential cognition (24, 48, 50). It is perhaps surprising that this network was not predictive of SRS scores, a measure of predominantly social traits. Rather, the DMN's connections with the SN and the FPTCN were predictive of more general improvement measured by the ABAS (51).

Given the broad nature of the ABAS, it is difficult to ascribe specific functions of these networks to their predictive power. Hearteningly, however, there is evidence that some of these regions may be predictive of other outcomes in ASD populations. The study by Lombardo et al. (25) of prospective language outcomes in ASD also showed that activity in the anterior insula and anterior cingulate of the SN, as well as medial prefrontal cortex and portions of the superior temporal cortex included in the DMN, correlated with future language ability using partial least squares correlation.

We did additionally assess outcome with a more specific measure of social autistic traits in ASD: the SRS. This measure does also include an "autistic mannerisms" subscale that taps into the restricted and repetitive behavior symptoms characteristic of ASD. T2 SRS and ΔSRS scores were significantly predicted by the FPTCN—a network that is most often associated with nonsocial reasoning, as well as cognitive and task control (48, 52, 53). However, these connections had poor sensitivity in predicting reliable improvements in SRS (44.44% sensitivity) and failed to accurately predict any of the reliable declines of SRS. In Fig. 2, one can see that intra-FPTCN connectivity underestimated ΔSRS scores and accurately estimated only individuals who had a nonsignificant RCI ($|RCI| < 1.96$).

Limitations and Future Studies. Our study has four key limitations that ought to be addressed in future research. First, although we have a comparably sized cohort ($n = 31$) to prior longitudinal studies of outcome in ASD, we have a relatively small sample size for the machine learning methods we used. Having fewer participants limits us to performing LOOCV rather than a more conservative and less variable model evaluation procedure, such as 3- or 10-fold cross-validation or separate training and test sets. Additionally, the ratio of features to participants makes a whole-brain connectivity regression model difficult. Data-driven feature selection in a small sample can be quite unstable and unreliable; thus, we were limited to choosing networks previously implicated in ASD. Future studies using a larger sample size could leverage more statistical power to make more generalizable claims about outcomes for individuals with ASD without intellectual disability

and the features that are most informative. Second, we have only two time points in our design. Behavioral studies with multiple data-collection time points (7, 13, 42) found that some participants show nonmonotonic functions of improvement or decline over time and cyclical change in impairment. Having more time points in a future study would allow us to also investigate the stability of change in symptoms. Third, the data were not originally collected with a longitudinal analysis in mind so behavioral assessments did not occur in fixed intervals. This variable follow-up latency could have affected the magnitude of change in outcome measures in a manner that was not accounted for by our linear correction. Fourth, we have a relatively homogeneous sample of adolescent and young adult males with ASD without cooccurring intellectual disability. It is unclear how our results would generalize to cohorts that include comorbid intellectual disability or females. Despite these limitations, this study highlights the potential clinical value of following individuals longitudinally for both behavioral and neuroimaging studies.

Additionally, we were not able to collect information regarding treatments and interventions received by the participants in this study between the follow-up periods. One important line of future work would be to harness such information to identify "brain types" that are most responsive to treatments.

Conclusions

Resting-state functional connectivity predicted significant variance in longitudinal measures of ASD social symptomatology and adaptive behavior even after effects of age, baseline score, and follow-up latency were removed. The performance of these models reflects an ability to predict, with high sensitivity, a meaningful behavioral change beyond the SE of the measurement. These findings provide evidence for networks that underlie changes in social and adaptive behavior and suggest that including brain data may improve the prognostic power of clinical assessments.

Materials and Methods

Ethics approval for this study was granted by the NIH Combined Neuroscience Institutional Review Board under protocol number 10-M-0027. The clinical trial number (clinicaltrials.gov) for this protocol is NCT01031407.

Thirty-one previously described male adolescents and young adults with ASD without intellectual disability (mean age = 17.93 y, SD = 3.39 y) took part in the study. Participants were recruited from the Washington, DC metropolitan area and met Diagnostic and Statistical Manual-IV and -5 diagnostic criteria as assessed by an experienced clinician. The participants received the Autism Diagnostic Interview (ADI or ADI-R) (54, 55), the Autism Diagnostic Observation Schedule (ADOS, modules 3 or 4) (56), or both, administered by a trained, research-reliable clinician. All scores from participants with ASD met cutoff for the category designated as "broad autism spectrum disorders" according to criteria established by the National Institute of Child Health and Human Development/National Institute on Deafness and Other Communication Disorders Collaborative Programs for Excellence in Autism (57). IQ scores were obtained for all participants. All full-scale IQ scores were >80 (Wechsler Abbreviated Scale of Intelligence, 29 ASD participants; Wechsler Adult Intelligence Scale-III, 2 ASD participants).

Scores on the SRS (58), an informant-based rating scale used to assess social and communication symptoms quantitatively, and the ABAS-II (59), an informant-based rating scale used to assess adaptive functioning, were obtained from parents for all ASD participants at the time of scanning [time 1 (T1)]. At least 1 y postscan, participants were contacted again to obtain follow-up scores on these two scales [time 2 (T2)], with an average of 2 y, 11 mo between SRS assessments and 2 y, 10 mo between ABAS measures. T2 behavioral scores were successfully obtained for 27 participants on the ABAS and 29 participants on the SRS (Table 2 and Fig. 1). Informed assent and consent were obtained from all participants and/or their parent/guardian in accordance with a National Institutes of Health Institutional Review Board approved protocol.

Outcome Measures. Regressions were performed to predict four separate outcome measures: T2 SRS sum t-scores (T2SRS), T2 ABAS general adaptive composite (GAC) standard scores (T2ABAS), the difference between T2 and T1 SRS sum t-scores (ΔSRS), and the difference between T2 and T1 ABAS GAC

Table 2. Demographic and behavioral outcome variables

Variable	Mean	SD
Age at scan	17.93	3.39
IQ	112.61	15.74
ADOS: soc + comm	11.61	4.62
T1 scores		
SRS	75.9	12.77
ABAS-II	71.96	13.56
T2 scores		
SRS	71.97	14.59
ABAS-II	81.93	14.53
Follow-up latency, y		
SRS	2.96	1.28
ABAS-II	2.85	1.11

soc + comm, social and communication symptoms.

standard scores (ΔABAS). Effects of age, variable duration between test times (hereby referred to as follow-up latency), and baseline score were removed from ΔSRS and ΔABAS by multiple linear regression. These residual ΔSRS and ΔABAS were then converted to reliability of change indices (RCIs) by dividing the score by each measure's SE of measurement derived from test–retest reliability (51, 58). This procedure yields pseudo–z-statistics for determining the significance of change. Effects of age and follow-up latency were also removed by the same method from T2 SRS and T2 ABAS (Fig. 1C). Regressions using the rs-fMRI data were performed on the residual outcome scores. This procedure, similar to a hierarchical regression, allowed us to determine how much unique variance is explained by the rs-fcMRI data after accounting for age, follow-up latency, and simple regression to the mean.

fMRI Acquisition. Functional MRI data were collected using a GE Signa 3T whole-body MRI scanner at the NIH Clinical Center NMR Research Facility. For each participant, a high-resolution T1-weighted anatomical image (MPRAGE) was obtained (124 axial slices, 1.2-mm slice thickness, field of view = 24 cm, 224 × 224 acquisition matrix). Spontaneous brain activity was measured during functional MRI using a gradient-echo echo-planar series with whole-brain coverage while participants maintained fixation on a central cross and were instructed to lie still and rest quietly (repetition time = 3,500 ms, echo time = 27 ms, flip angle = 90°, 42 axial interleaved slices per volume, 3.0-mm slice thickness, field of view = 22 cm, 128 × 128 acquisition matrix, single-voxel volume = 1.7 mm × 1.7 mm × 3.0 mm). Each resting scan lasted 8 min, 10 s for a total of 140 consecutive whole-brain volumes. A GE eight-channel send–receive head coil was used for all scans, with a sensitivity encoding (SENSE) factor of 2 used to reduce gradient coil heating during the session.

fMRI Preprocessing. fMRI data were preprocessed in accordance with pipelines recommended by Jo et al. (60) using the AFNI software package (61). The first four echo-planar image volumes were removed from the resting scan, and the remaining volumes were subsequently despiked to remove large transient fluctuations. Retroicor (62) and respiration volume per time (RVT) (63) regressors were created from cardiac and respiration measures. Volumes were then slice time-corrected, coregistered to the anatomical scan, resampled to 2.0-mm isotropic voxels, smoothed with an isometric 6-mm full-width half-maximum Gaussian kernel, normalized to reflect percent signal change, and transformed into the standardized Talairach and Tournoux (64) volume for the purposes of group analyses. We applied the basic ANATICOR procedure (65) for removing nuisance artifacts from the echo-planar image data as follows: The anatomical scan was segmented into tissue compartments using Freesurfer (66). Ventricle and white-matter masks were created and eroded. Masks were then applied to the volume-registered echo-planar image data before smoothing to yield pure nuisance time series for the ventricles, as well as local estimates of the blood oxygen level-dependent signal in white matter that were averaged within a 15-mm radius sphere. All nuisance time series were detrended with fourth-order polynomials before least-squares model fitting to each voxel's time series. Nuisance variables for each voxel included the following: an average ventricle time series, a local average white-matter time series, twelve parameter estimates for head motion and the first derivatives of head motion, RVT, and Retroicor. The predicted time course of these nuisance variables was then subtracted from the full voxel time series to yield a residual time series to be used in correlation analyses. The residuals were bandpass filtered (0.01–0.1 Hz) and censored simultaneously. This latter procedure removed time points

in which gross head motion exceeded 0.3 mm or in which >10% of brain voxels were determined to be outliers. Such global signal intensity outliers were determined using AFNI's 3dToutcount, which removes trend terms from the time series before setting a threshold based on the median absolute deviation. Spatial alignment and Talairach transformation were visually inspected to ensure proper registration to the template brain. All participants had <10% of the total time frames removed due to motion or global intensity changes.

Connectivity Measures and Feature Matrices. A set of 264 spherical regions of interest (ROIs) (5-mm radius) described in Power et al. (31) were used to create fMRI time course correlation matrices for subjects' processed echo-planar image (EPI) time series. Time courses were extracted and averaged within each region. Linear correlations were computed between the average time courses of each region in an ROI set and Fisher transformed. Fisher transformed r values were then z-transformed within subject to remove the effects of global levels of correlation.

In the Power et al. ROI set, each region is associated with 1 of 13 functional networks (31). In our previous study (19), we found that connections involving regions from several of these networks were highly informative in classifying scans from ASD and typically developing individuals. The four networks that were the most common among the top 25 regions were the default-mode network (DMN), frontoparietal task control network (FPTCN), the salience network (SN), and the cingulo-opercular network (CON). Because the scans used in the present analysis constituted a subset of the scans used in our previous study, we sought to ensure that our choice of networks was robust and not biased by previous analyses. We repeated the region ranking procedure described in Plitt et al. (19) using only scans not presented in the current study. DMN, FPTCN, SN, and the sensory/somatomotor-hand network were the top networks in this repeated analysis. CON failed to survive replication. Because the DMN, FPTCN, and SN replicated across both procedures, we will focus only on these three networks.

To further aid in feature selection, we performed separate regressions using correlation values either within each network (e.g., all DMN–DMN correlations) or across a pair of networks (e.g., all DMN–FPTCN correlations).

In each case, this process yielded an $N_s \times N_f$ feature matrix, F, where $N_s =$ the number of subjects and $N_f =$ the number of features (Fisher-transformed correlation values). For each regression, F has an associated label vector, L, containing the clinical scores for each subject. Subjects' time 1 behavioral scores were used to create an additional feature matrix.

Multivariate Brain-Based Regression to Predict Outcome Measures. To determine whether functional connectivity can predict future clinical scores, we implemented a simple large-scale regression algorithm, ridge regression, using Scikit-learn (67). Due to our clinical sample size and to obtain the highest possible estimates of regression accuracy, we chose to use leave-one-out cross-validation (LOOCV). In LOOCV, data from all but one participant (the "training set") is used to predict the label for the left-out participant (the "test set"), repeating the process for all subjects.

To further reduce overfitting, during each fold of cross-validation, we applied univariate feature selection to choose the 100 best features to keep in the model before model selection and fitting [see Dosenbach et al. (68) for a similar approach]. After univariate feature selection, hyperparameters were chosen based on the training set by a grid-search method using stratified-fivefold cross-validation. The final model chosen by grid-search was fit to the training set and subsequently tested on the "test" sample from the first split of the data. This procedure was repeated for all folds of the first-level cross-validation procedure to determine the overall performance of the machine-learning regression. This method resulted in a two-level cross-validation procedure similar to that described in Plitt et al. (19). This nested cross-validation is recommended to achieve unbiased estimates of hyperparameters (69).

For each algorithm, we report an empirical r^2 value derived from LOO cross-validation. The significance of this r^2 value was determined using permutations tests. In this test, the label vector, L, that contains each subject's actual clinical scores is randomly permuted many times, and the same cross-validation described above is performed after each permutation. We performed 1,000 permutations to build a null distribution for each cross-validation metric. If the cross-validation metrics achieved with the true data are greater than 95% ($P < 0.05$) of those achieved during the permutations, then we accept the performance of the classifier as significantly greater than chance.

1. Eaves LC, Ho HH (2008) Young adult outcome of autism spectrum disorders. *J Autism Dev Disord* 38(4):739–747.
2. Kanne SM, et al. (2011) The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *J Autism Dev Disord* 41(8):1007–1018.
3. Klin A, et al. (2007) Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: The Vineland and the ADOS. *J Autism Dev Disord* 37(4):748–759.
4. McGovern CW, Sigman M (2005) Continuity and change from early childhood to adolescence in autism. *J Child Psychol Psychiatry* 46(4):401–408.
5. Howlin P, Goode S, Hutton J, Rutter M (2004) Adult outcome for children with autism. *J Child Psychol Psychiatry* 45(2):212–229.
6. Pugliese CE, et al. (2015) Increasing adaptive behavior skill deficits from childhood to adolescence in autism spectrum disorder: Role of executive function. *J Autism Dev Disord* 45(6):1579–1587.
7. Woodman AC, Smith LE, Greenberg JS, Mailick MR (2015) Change in autism symptoms and maladaptive behaviors in adolescence and adulthood: The role of positive family processes. *J Autism Dev Disord* 45(1):111–126.
8. Seltzer MM, et al. (2003) The symptoms of autism spectrum disorders in adolescence and adulthood. *J Autism Dev Disord* 33(6):565–581.
9. Szatmari P, et al. (2009) Similar developmental trajectories in autism and Asperger syndrome: From early childhood to adolescence. *J Child Psychol Psychiatry* 50(12):1459–1467.
10. Taylor JL, Seltzer MM (2010) Changes in the autism behavioral phenotype during the transition to adulthood. *J Autism Dev Disord* 40(12):1431–1446.
11. Billstedt E, Gillberg IC, Gillberg C (2007) Autism in adults: Symptom patterns and early childhood predictors. Use of the DISCO in a community sample followed from childhood. *J Child Psychol Psychiatry* 48(11):1102–1110.
12. Mazurek MO, Kanne SM, Miles JH (2012) Predicting improvement in social-communication symptoms of autism spectrum disorders using retrospective treatment data. *Res Autism Spectr Disord* 6(1):535–545.
13. Gotham K, Pickles A, Lord C (2012) Trajectories of autism severity in children using standardized ADOS scores. *Pediatrics* 130(5):e1278–e1284.
14. Howlin P, Mawhood L, Rutter M (2000) Autism and developmental receptive language disorder–a follow-up comparison in early adult life. II. Social, behavioural, and psychiatric outcomes. *J Child Psychol Psychiatry* 41(5):561–578.
15. Perry A, Flanagan HE, Dunn Geier J, Freeman NL (2009) Brief report: The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *J Autism Dev Disord* 39(7):1066–1078.
16. Farley MA, et al. (2009) Twenty-year outcome for individuals with autism and average or near-average cognitive abilities. *Autism Res* 2(2):109–118.
17. Szatmari P, et al.; Pathways in ASD Study Team (2015) Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. *JAMA Psychiatry* 72(3):276–283.
18. Seltzer MM, Shattuck P, Abbeduto L, Greenberg JS (2004) Trajectory of development in adolescents and adults with autism. *Ment Retard Dev Disabil Res Rev* 10(4):234–247.
19. Plitt M, Barnes KA, Martin A (2015) Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. *Neuroimage Clin* 7:359–366.
20. Uddin LQ, et al. (2013) Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry* 70(8):869–879.
21. Anderson JS, et al. (2011) Functional connectivity magnetic resonance imaging classification of autism. *Brain* 134(Pt 12):3742–3754.
22. Nielsen JA, et al. (2013) Multisite functional connectivity MRI classification of autism: ABIDE results. *Front Hum Neurosci* 7(September):599.
23. Lynch CJ, et al. (2013) Default mode network in childhood autism: Posteromedial cortex heterogeneity and relationship with social deficits. *Biol Psychiatry* 74(3):212–219.
24. Gotts SJ, et al. (2012) Fractionation of social brain circuits in autism spectrum disorders. *Brain* 135(Pt 9):2711–2725.
25. Lombardo MV, et al. (2015) Different functional neural substrates for good and poor language outcome in autism. *Neuron* 86(2):567–577.
26. Canli T, et al. (2005) Amygdala reactivity to emotional faces predicts improvement in major depression. *Neuroreport* 16(12):1267–1270.
27. Hoeft F, et al. (2011) Neural systems predicting long-term outcome in dyslexia. *Proc Natl Acad Sci USA* 108(1):361–366.
28. Norman AL, et al. (2011) Neural activation during inhibition predicts initiation of substance use in adolescence. *Drug Alcohol Depend* 119(3):216–223.
29. Supekar K, et al. (2013) Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proc Natl Acad Sci USA* 110(20):8230–8235.

30. Myers CA, et al. (2014) White matter morphometric changes uniquely predict children's reading acquisition. *Psychol Sci* 25(10):1870–1883.
31. Power JD, et al. (2011) Functional network organization of the human brain. *Neuron* 72(4):665–678.
32. Achenbach T, Rescorla LA (2001) Manual for ASEBA School-Age Form & Profiles.
33. Achenbach TM, Rescorla LA (2003) Manual for the ASEBA Adult Forms & Profiles.
34. Piven J, Harper J, Palmer P, Arndt S (1996) Course of behavioral change in autism: A retrospective study of high-IQ adolescents and adults. *J Am Acad Child Adolesc Psychiatry* 35(4):523–529.
35. Chowdhury M, Benson BA, Hillier A (2010) Changes in restricted repetitive behaviors with age: A study of high-functioning adults with autism spectrum disorders. *Res Autism Spectr Disord* 4(2):210–216.
36. Constantino JN, et al. (2009) Developmental course of autistic social impairment in males. *Dev Psychopathol* 21(1):127–138.
37. Duncan AW, Bishop SL (2015) Understanding the gap between cognitive abilities and daily living skills in adolescents with autism spectrum disorders with average intelligence. *Autism* 19(1):64–72.
38. Mawhood L, Howlin P, Rutter M (2000) Autism and developmental receptive language disorder–a comparative follow-up in early adult life. I. Cognitive and language outcomes. *J Child Psychol Psychiatry* 41(5):547–559.
39. Howlin P (2003) Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *J Autism Dev Disord* 33(1):3–13.
40. Szatmari P, Bryson SE, Boyle MH, Streiner DL, Duku E (2003) Predictors of outcome among high functioning children with autism and Asperger syndrome. *J Child Psychol Psychiatry* 44(4):520–528.
41. Taylor JL, Mailick MR (2014) A longitudinal examination of 10-year change in vocational and educational activities for adults with autism spectrum disorders. *Dev Psychol* 50(3):699–708.
42. Shattuck PT, et al. (2007) Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *J Autism Dev Disord* 37(9):1735–1747.
43. Supekar K, et al. (2013) Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Reports* 5(3):738–747.
44. Di Martino A, et al. (2009) Functional brain correlates of social and nonsocial processes in autism spectrum disorders: An activation likelihood estimation meta-analysis. *Biol Psychiatry* 65(1):63–74.
45. Menon V, Uddin LQ (2010) Saliency, switching, attention and control: A network model of insula function. *Brain Struct Funct* 214(5-6):655–667.
46. Masten CL, et al. (2011) An fMRI investigation of responses to peer rejection in adolescents with autism spectrum disorders. *Dev Cogn Neurosci* 1(3):260–270.
47. Bolling DZ, et al. (2011) Enhanced neural responses to rule violation in children with autism: A comparison to social exclusion. *Dev Cogn Neurosci* 1(3):280–294.
48. Jack AI, et al. (2013) fMRI reveals reciprocal inhibition between social and physical cognitive domains. *Neuroimage* 66:385–401.
49. Agam Y, Joseph RM, Barton JJS, Manoach DS (2010) Reduced cognitive control of response inhibition by the anterior cingulate cortex in autism spectrum disorders. *Neuroimage* 52(1):336–347.
50. Spreng RN, Grady CL (2010) Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *J Cogn Neurosci* 22(6):1112–1123.
51. Harrison PL, Oakland T (2003) Adaptive Behavior Assessment System II (PsychCorp, San Antonio, TX), 2nd Ed.
52. Cole MW, et al. (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat Neurosci* 16(9):1348–1355.
53. Dosenbach NUF, et al. (2007) Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci USA* 104(26):11073–11078.
54. Le Couteur A, et al. (1989) Autism diagnostic interview: A standardized investigator-based instrument. *J Autism Dev Disord* 19(3):363–387.
55. 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(5):659–685.
56. Lord C, et al. (2000) The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 30(3):205–223.
57. Lainhart JE, et al. (2006) Head circumference and height in autism: A study by the Collaborative Program of Excellence in Autism. *Am J Med Genet A* 140(21):2257–2274.
58. Constantino J, Gruber C (2005) *Social Responsiveness Scale (SRS): Manual* (CA West Psychol Services, Los Angeles).
59. Oakland T, Harrison PL, eds (2008) *Adaptive Behavior Assessment System-II: Clinical Use and Interpretation* (Elsevier, San Diego), 2nd Ed.
60. Jo HJ, et al. (2013) Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state FMRI. *J Appl Math* 2013:935154.
61. Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29(3):162–173.
62. Glover GH, Li TQ, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* 44(1):162–167.
63. Birn RM, Smith MA, Jones TB, Bandettini PA (2008) The respiration response function: The temporal dynamics of fMRI signal fluctuations related to changes in respiration. *Neuroimage* 40(2):644–654.
64. Talairach J, Tournoux P (1988) *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York).
65. Jo HJ, Saad ZS, Simmons WK, Milbury LA, Cox RW (2010) Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *Neuroimage* 52(2):571–582.
66. Fischl B, et al. (2002) Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* 33(3):341–355.
67. Pedregosa F, Weiss R, Brucher M (2011) Scikit-learn. *Machine Learning in Python.* 12:2825–2830.
68. Dosenbach NUF, et al. (2010) Prediction of individual brain maturity using fMRI. *Science* 329(5997):1358–1361.
69. Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (Springer, New York), 2nd Ed.