Identification of neural connectivity signatures of autism using machine learning

Gopikrishna Deshpande[1,2], Lauren E. Libero[3], Karthik R. Sreenivasan[1], Hrishikesh D. Deshpande[3] and Rajesh K. Kana[3]*

[1] AU MRI Research Center, Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA
[2] Department of Psychology, Auburn University, Auburn, AL, USA
[3] Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, USA

Alterations in interregional neural connectivity have been suggested as a signature of the pathobiology of autism. There have been many reports of functional and anatomical connectivity being altered while individuals with autism are engaged in complex cognitive and social tasks. Although disrupted instantaneous correlation between cortical regions observed from functional MRI is considered to be an explanatory model for autism, the causal influence of a brain area on another (effective connectivity) is a vital link missing in these studies. The current study focuses on addressing this in an fMRI study of Theory-of-Mind (ToM) in 15 high-functioning adolescents and adults with autism and 15 typically developing control participants. Participants viewed a series of comic strip vignettes in the MRI scanner and were asked to choose the most logical end to the story from three alternatives, separately for trials involving physical and intentional causality. The mean time series, extracted from 18 activated regions of interest, were processed using a multivariate autoregressive model (MVAR) to obtain the causality matrices for each of the 30 participants. These causal connectivity weights, along with assessment scores, functional connectivity values, and fractional anisotropy obtained from DTI data for each participant, were submitted to a recursive cluster elimination based support vector machine classifier to determine the accuracy with which the classifier can predict a novel participant's group membership (autism or control). We found a maximum classification accuracy of 95.9% with 19 features which had the highest discriminative ability between the groups. All of the 19 features were effective connectivity paths, indicating that causal information may be critical in discriminating between autism and control groups. These effective connectivity paths were also found to be significantly greater in controls as compared to ASD participants and consisted predominantly of outputs from the fusiform face area and middle temporal gyrus indicating impaired connectivity in ASD participants, particularly in the social brain areas. These findings collectively point toward the fact that alterations in causal connectivity in the brain in ASD could serve as a potential non-invasive neuroimaging signature for autism.

Keywords: autism, effective connectivity, fMRI, classification, machine learning, theory-of-mind

INTRODUCTION

A biological origin for autism spectrum disorders (ASD) had been proposed even in the earliest published accounts of the disorder (Kanner, 1943; Asperger, 1944). Despite several decades of research since then, a focal neurobiological marker for autism has been rather elusive. Brain imaging techniques in the last decade, particularly functional and structural MRI, have pointed to disrupted cortical connectivity as a defining neural feature of ASD (Kana et al., 2011; Just et al., 2012). Neuroimaging studies have reported functional under connectivity (weaker synchronization of activated brain areas) between frontal and posterior brain areas (Just et al., 2004, 2007; Villalobos et al., 2005; Kana et al., 2006, 2007, 2009; Koshino et al., 2008; Mason et al., 2008; Solomon et al., 2009; Damarla et al., 2010; Jones et al., 2010; Mizuno et al., 2011; Schipul et al., 2011), and intact or increased functional connectivity within relatively posterior brain areas (Villalobos et al., 2005; Kana et al., 2006; Damarla et al., 2010; Kana et al., under review). Similar findings have also been reported during task-free resting state in autism (Cherkassky et al., 2006; Assaf et al., 2010; Murdaugh et al., 2012). Furthermore, diffusion tensor imaging (DTI) studies have reported disruptions in anatomical connectivity in ASD (Barnea-Goraly et al., 2004, 2010; Alexander et al., 2007; Keller et al., 2007; Jou et al., 2011; see Travers et al., 2012 for a review). Although there is converging evidence for connection abnormalities, the neural connectivity model of ASD is based primarily on functional connectivity, with some contributing evidence from white matter integrity. While the insights gained from these models are valuable, functional connectivity is a method for assessing zero-lag correlations, and does not provide insight into the time-lagged relationships and direction of such causal influence.

Effective connectivity, on the other hand, refers to the influence one neural system exerts over another with respect to a given experimental context (Buchel and Friston, 2000), thus helping uncover more information about how brain areas communicate. Effective connectivity can provide information about the transfer of information from one node to another, and differentiate between top-down vs. bottom-up effects. Thus, effective connectivity findings have enriched models of cognitive function by emphasizing the dynamic and interactive nature of neural instantiations (McIntosh et al., 2010). Studying such interactions is important not only for understanding typical brain functioning, but also is critical in learning more about diseases. Considering relatively consistent reports of disruptions in functional connectivity in ASD, it is perhaps a logical and valuable next step to study how information transfer is accomplished in ASD brains. Of particular interest is to explore the information transfer among brain areas that are part of a team to perform higher-order cognitive and social functions, which people with ASD particularly struggle with.

Understanding the information transfer, or the lack of it, between specific nodes in the brain may help uncover the neural bases of behavioral and social problems in ASD. It should be noted that only four previous studies have examined effective connectivity between brain regions in ASD (Bird et al., 2006; Wicker et al., 2008; Shih et al., 2010; Shen et al., 2012). These studies only permit limited inferences as they used a small number of regions and made prior assumptions about the underlying connectional architecture. This is because they used confirmatory methods such as dynamic causal modeling (Friston et al., 2003) and structural equation modeling (McIntosh and Gonzales-Lima, 1994) in their studies. In contrast, the present study applies multivariate autoregressive (MVAR) modeling for obtaining Granger causality between a large number of brain regions. This is an exploratory technique which does not make any prior assumptions about the underlying connectional architecture. In addition, it is capable of obtaining condition-specific causal influences between a large number of brain regions using relatively shorter time series. According to the principle of Granger causality, the directional causal influence from time series X to time series Y can be inferred if past values of time series X help predict the present and future values of the time series Y (Granger, 1969). MVAR models have been used to characterize the predictive relationship between the time series from different brain regions in many previous studies (Roebroeck et al., 2005; Abler et al., 2006; Deshpande et al., 2008, 2009b; Sathian et al., 2011). But according to many recent studies, the spatial variability of the hemodynamic response is considered to be of vascular origin, and hence confounding the Granger causal estimates obtained from raw fMRI time series (David et al., 2008; Deshpande et al., 2010b). Removing the smoothing effect of the hemodynamic response function (HRF) will increase the effective temporal resolution of the signal in addition to accounting for the inter-subject and inter-regional variability of the HRF (Handwerker et al., 2004). This can be accomplished using blind hemodynamic deconvolution methods where in the underlying hidden neuronal variable for the fMRI time series can be estimated. We employed this approach in this study by deconvolving the hemodynamic

response from fMRI time series using a Cubature Kalman filter (CKF) (Havlicek et al., 2011). Subsequently, these hidden neuronal variables were input into the MVAR model to obtain directional connectivity measures.

Investigating the directional interactions among brain areas in ASD could supplement functional connectivity findings, and potentially may serve as a neural signature for the disorder. Thus, connection abnormalities at anatomical, functional, and causal levels may be considered for potential diagnosis of ASD and/or to supplement the behavior-based diagnosis. However, such attempts will need to test and validate the diagnostic utility of connection abnormalities in ASD. Questions pertaining to diagnostic utility may be best answered through pattern classification analyses using sophisticated machine learning algorithms (Deshpande et al., 2010a; Weygandt et al., 2011; Shinkareva et al., 2013). In this regard, earlier studies have used pattern recognition and machine learning algorithms reliably in classification. Craddock et al. (2009) showed that by using resting state functional connectivity metrics as features in SVM based machine learning classifier, Major Depressive Disorder (MDD) patients were successfully distinguished from healthy controls. In another study, the treatment type provided to patients with MDD was accurately identified using SVM classifier based on the effective connectivity measures (Deshpande et al., 2009a). A pattern recognition approach using structural networks as biomarkers was proposed (Marquand et al., 2013) for classification of Parkinson's Disorder. This method of analysis accurately predicted the diagnosis in patients with Parkinson disorders. A study by Mirowski and colleagues (2009) showed that machine learning classifiers can be successfully used in prediction of seizures in patients with epilepsy. Given the success of pattern recognition and classification methods based on machine learning techniques in other fields and contexts, they could potentially prove to be useful to correctly identify participants with ASD after replication and fine tuning. In these lines, diagnostic information (although preliminary) has been obtained from even short fMRI BOLD sequences, such as characterization of subject age (Dosenbach et al., 2010), classification of dementia (Chen et al., 2011), and autism (Anderson et al., 2011; Murdaugh et al., 2012; Wang et al., 2012). For a neurodevelopmental disorder such as ASD, which is currently diagnosed solely by behavioral observation and in-person interviews by clinicians, classification by brain imaging signatures could be applied to gain more accurate (and perhaps earlier) diagnosis of the disorder. Classification studies have utilized a wide range of data sources to differentiate participants into ASD and TD groups, including functional connectivity (Anderson et al., 2011; Murdaugh et al., 2012; Wang et al., 2012), voxel based morphometry (Uddin et al., 2011; Calderoni et al., 2012), fMRI activation patterns (Coutanche et al., 2011), EEG (Duffy and Als, 2012), and DTI (Ingalhalikar et al., 2011). Yet, none of these methods are currently employed to diagnose the disorder. Issues remain regarding generalizability, such as whether the classification techniques can still be accurately applied to younger children. When other disorders also show functional connectivity and resting state abnormalities, such as schizophrenia (Lawrie et al., 2002; Meyer-Lindenberg et al., 2005; Garrity et al., 2007) and ADHD (Tian et al., 2006; Cubillo et al., 2010),

it begs the question about the specificity of these metrics to ASD. However, notably, effective connectivity markers have not been used in classification of ASD individuals. In this regard, effective connectivity could be an additional data source utilized to add to classification of ASD participants, potentially providing sufficient information to serve as a biomarker for the disorder. In other words, effective connectivity could contribute significantly to the global connectivity-based neural characterization of ASD. Also, whereas traditional statistical analyses can uncover significant group differences in brain activation and connectivity, classification analyses can serve to identify brain imaging signatures which are not only able to separate or distinguish the groups, but also predict the group membership of a new subject.

In the current study we explored the causal influences between brain regions that may underlie the processing of theory-of-mind (ToM) in young adults with ASD and typically developing (TD) control participants. The original fMRI study on ToM was published earlier (Kana et al., 2012), reporting findings of brain activity, functional connectivity and white matter integrity. In the current study, we obtained causal connectivity between 18 brain regions activated in the ToM task in our previous publication (Kana et al., 2012). We used these causal connectivity weights along with the following metrics from our previous study—assessment scores, functional connectivity values and fractional anisotropy (FA) obtained from DTI data—as features for classification. We employed recursive cluster elimination to select important features and a support vector machine (SVM) classifier to classify participants into ASD and TD based on the entire feature set. This paper is novel in that it takes into consideration different aspects of brain connectivity, instead of a single index, to characterize the nature of brain functioning in individuals with ASD for classification purposes.

METHOD

PARTICIPANTS

Fifteen adolescents and young adults with high-functioning ASD (mean age: 21.14 years) and 15 age-and-IQ-matched individuals with typical development (TD) (mean age: 22.18 years) participated in this fMRI study. Functional connectivity, structural connectivity, behavioral data, and brain activation measures from the same participants were reported elsewhere (Kana et al., 2012). All participants were required to have an IQ of 80 or above measured by the Wechsler Abbreviated Scale of Intelligence (WASI).

The participants with ASD were recruited from the University of Alabama ASD Clinic and surrounding service providers. The study was approved by the Institutional Review Board of the University of Alabama at Birmingham, and all participants provided informed consent for their participation in the study. Participants with ASD had received a previous diagnosis of an ASD based on Autism Diagnostic Interview (ADI-R) symptoms, and Autism Diagnostic Observation Schedule (ADOS). Eight of the 15 ASD participants in this study had received a diagnosis of Asperger's Disorder. The TD participants were recruited through newspaper advertisements and through the University of Alabama at Birmingham's Psychology 101 course subject pool. They were screened through a parent-report (for participants younger than 18 years) or self-report history questionnaire to rule out neurological disorders, such as ASD, ADHD, or Tourette's Disorder, that could potentially confound the results. All participants completed the Autism Spectrum Quotient (AQ) questionnaire (Baron-Cohen et al., 2001b), and the Reading the Mind in the Eyes (RME) test (Baron-Cohen et al., 2001a). Demographic information about the participants is shown in **Table 1.**

EXPERIMENTAL PARADIGM AND IMAGING PARAMETERS

The stimuli consisted of a series of black and white comic strip vignettes (adapted from Brunet et al., 2000) depicting scenarios that demand either a physical causal attribution or an intentional causal attribution to arrive at a logical ending. The first part of the vignette was presented for 5 s and the participants' task was to choose a logical ending to the story from the three choices in the second panel presented for 6 s. The entire vignette remained on the screen for a total of 11 s. The experiment was designed in an event-related format. All data were collected using a Siemens 3.0 Tesla Allegra head-only scanner (*Siemens* Medical Inc., Erlangen, Germany). For functional imaging, a single-shot gradient-recalled echo-planar pulse sequence was used for rapid image acquisition ($TR = 1000$ ms, $TE = 30$ ms, flip angle = 60 degrees). Seventeen adjacent oblique-axial slices were acquired in an interleaved sequence with 5 mm slice thickness, 1 mm slice gap, a 24×24 cm^2 field of view (FOV), and a 64×64 matrix, resulting in an in-plane resolution of $3.75 \times 3.75 \times 5$ mm^3. More information on the experimental paradigm and imaging parameters for the 3D MPRAGE structural MRI data and diffusion weighted echo-planar imaging data can be found in Appendix A (for further details, please refer to Kana et al., 2012).

Table 1 | Demographic information of the ASD and TD control participants.

| | Autism | | | | | Control | | | | | Group difference | |
| | N = 15 | | | | | N = 15 | | | | | | |
	Mean	Range	*SD*	Median	IQR	Mean	Range	*SD*	Median	IQR	*t*-value	*p*-value
Age	21.14	16–29	0.99	20.08	5	22.28	16–34	1.08	21.83	3.3	0.77	0.44
VIQ	104.8	74–139	5.02	102	31	113.93	98–127	2.2	116	14	1.66	0.11
PIQ	107.7	73–129	4.33	106	25	107.2	89–124	2.48	108	6	0.11	0.92
FSIQ	106.93	80–140	4.84	105	35	112	96–128	2.24	113	13	0.94	0.35
MIE	19.07	15–24	0.7	20	4	21.6	18–24	0.55	22	4	2.84	0.01
AQ	26.5	9–38	2.04	29	12	14.06	4–22	1.45	14	9	5.47	<0.001

DATA ANALYSES

Head motion correction and regions of interest (ROI) definition

Within-group brain activation was examined for the whole group (ASD + TD) of participants (see Kana et al., 2012). Functional ROIs were defined on the group activation map for the whole group (ASD + TD) for the contrast (Intentional Causality + Physical Causality) vs. Fixation, so that it best represented the study. Because head motion can impact connectivity analyses (Satterthwaite et al., 2012; Van Dijk et al., 2012), a conservative threshold of 0.5 mm was set for head motion in any direction. In addition, the root mean square (RMS) values of head motion were measured in three translational directions (x, y, and z) and three rotations (pitch, roll, and yaw) for each individual participant in the study (see Appendix B **Table B1**). We examined group differences in head motion on this data using a Mann-Whitney U Test, which is a non-parametric test and may be more appropriate in case assumptions about normality of sample distributions are not met.

Eighteen ROIs were identified: supplementary motor area (SMA), left and right inferior frontal gyrus (LIFG, RIFG), left and right precentral cortex (LPRCN, RPRCN), left and right middle temporal gyrus (LMTG, RMTG), right superior temporal gyrus (RSTG), left and right inferior parietal lobule (LIPL, RIPL), left and right fusiform gyrus (LFFG, RFFG), left and right superior parietal lobule (LSPL, RSPL), left and right middle occipital gyrus (LMOG, RMOG), and left and right temporal parietal junction (LTPJ, RTPJ). A sphere was defined for each cluster (with a radius ranging from 8 to 12 mm) that best captured the cluster of activation in the contrast map for each group. The radius was selected to specifically encompass as much of the activation cluster as possible, without including surrounding (not significantly activated) areas. Selecting ROIs of the same radius or utilizing anatomically defined ROIs may entail those ROIs not encompassing the entire cluster of activation, or may include tissue that is not significantly active for the task. As a result, extracting time courses from ROIs defined in these ways may result in time series variability that does not reflect the cognitive task being processed.

The effective connectivity model

Let l fMRI time series be represented as $X(t) = [x_1(t)x_2(t)\ldots x_l(t)]$. Below, we present a model linking observed fMRI time series to underlying latent neuronal variables. A dynamic state-space model can be described as follows.

$$\tilde{h}_T^l = \begin{bmatrix} h_T^l \\ u_T^l \\ \theta_T^l \end{bmatrix} = \begin{bmatrix} f(h_{T-1}^l, u_{T-1}^l, \theta_{T-1}^l) \\ u_{T-1}^l \\ \theta_{T-1}^l \end{bmatrix} + \begin{bmatrix} P_{T-1}^l \\ Q_{T-1}^l \\ R_{T-1}^l \end{bmatrix} \quad (1)$$

Where h is the hidden neuronal state variable, u is the exogenous input and θ represents the HRF parameter variables. f is the function which links the current neuronal state to the previous neuronal states, exogenous inputs and parameters. The subscript T indicates continuous time and the superscript l indicates the number of time series in the model. P, Q, and R are the zero mean

Gaussian state noise vectors. The observation equation links the state to observation variables as given below.

$$x_l(t) = \boldsymbol{m}(\tilde{h}_t^l) + \varepsilon_{t-1} \quad (2)$$

where ε is the measurement noise, t is discrete time and $\boldsymbol{m}$ is the measurement function which links the state variables to measurement variables. The exogenous inputs u, which is the experimental boxcar function, and $x_l(t)$ are the inputs to the model. As demonstrated before, using the CKF (Havlicek et al., 2011), the hidden neuronal variables can be estimated successfully. The CKF performs very efficient joint estimation of the hidden neuronal state variables and parameters. In addition, since Eq. 1 represents a continuous time model, the neuronal variables can be estimated with a highly improved temporal resolution up to 10 times smaller than the TR. When the hidden neuronal state variables $h_l(t)$ are input into the MVAR model, we get the following equation.

$$\begin{bmatrix} h_1(t) \\ h_2(t) \\ . \\ . \\ h_l(t) \end{bmatrix} = \begin{bmatrix} 0 & a_{12}(0) & \cdots & a_{1l}(0) \\ a_{21}(0) & 0 & & a_{2l}(0) \\ . & . & 0 & . \\ . & . & & . \\ a_{l1}(0) & a_{l2}(0) & \cdots & 0 \end{bmatrix} \times \begin{bmatrix} h_1(t) \\ h_2(t) \\ . \\ . \\ h_l(t) \end{bmatrix}$$

$$+ \sum_{j=1}^{p} \begin{bmatrix} a_{11}(j) & a_{12}(j) & \ldots & a_{1l}(j) \\ a_{21}(j) & a_{22}(j) & & a_{2l}(j) \\ . & . & & . \\ . & . & & . \\ a_{l1}(j) & a_{l2}(j) & \ldots & a_{ll}(j) \end{bmatrix} \times \begin{bmatrix} h_1(t-j) \\ h_2(t-j) \\ . \\ . \\ h_l(t-j) \end{bmatrix}$$

$$+ \begin{bmatrix} e_1(t) \\ e_2(t) \\ . \\ . \\ e_l(t) \end{bmatrix} \quad (3)$$

where p is the model order estimated by the Akaike/Bayesian information criterion (Deshpande et al., 2009b), a represent the model coefficients and e represents the error of the MVAR model. From the above equation it can be observed that $a(0)$ represents the instantaneous influences between the time series, and the Granger causal influences between them is indicated by $a(j)$, $j = 1 \ldots . p$. Both terms are used in the model because including both instantaneous and causal terms in the model minimizes the "leakage" of instantaneous correlation into causality (Deshpande et al., 2010c). The multivariate model we have used is less sensitive to the effects of missing variables than the traditionally used pairwise bivariate models (Kuś et al., 2004). Also, since we included all 18 regions which were activated in the effect of interest, it guaranteed to a certain level that all regions involved in the task were indeed included in the model.

Effective connectivity analysis

Mean time series from 18 activated regions were obtained for each of the 15 participants with ASD and the 15 typical control participants. Using the boxcar function corresponding to

"intentional causality" as the exogenous input, hidden neuronal variables corresponding to normalized mean fMRI time series were obtained and input into the MVAR model. The Granger causal relationships between the 18 regions for each participant (ASD and TD) were obtained. The number of coefficients in the MVAR model is equal to k^2p (k is the number of time series and p is the model order) (Kuś et al., 2004). This must be smaller than the number of time points in each time series. We had 18 ROI time series, each of length 460. Since we used a first order model, $k^2p = 324$ which is less than 460. Therefore, we were able to estimate the model.

Classification using support vector machine

The statistical separation of neural signatures (e.g., t-test) does not guarantee generalizability or predictive power of those signatures for diagnosis. Therefore, in this study, we also used machine learning approaches for identification of metrics which can accurately classify individuals with ASD from individuals with typical development. A Recursive Cluster Elimination based Support Vector Machine (RCE-SVM) (Deshpande et al., 2010a) was used in this study to classify the participants based on granger causal path weights between the 18 ROIs, functional connectivity z-scores for all pairs of the 18 ROIs, assessment scores (AQ and RME scores) and FA values for the white matter tract extending into the temporal lobe as the input features. The functional connectivity, assessment and DTI FA values were obtained from our prior study (Kana et al., 2012).

Our choice of SVM for classification was motivated by its wide applicability as a machine learning approach (Vapnik, 1995) for classification in many different fields (Wang, 2005). Previous studies have demonstrated that using discriminatory features enhances SVM classification (Craddock et al., 2009). Therefore, to enhance the performance of the SVM classifier, filtering and wrapper methods for feature selection have been used. Filtering methods are based on extraction of features that are statistically different between classes. They can be extracted using statistical tests such as a t-test. The wrapper approach is based on iteratively eliminating features to minimize the prediction error. RCE is one of the wrapper methods that is an iterative process were the feature selection and classification steps are embedded with each other. The main steps of the RCE-SVM algorithm, shown in the flowchart in **Figure 1**, are the cluster step, the SVM scoring step and the RCE step. Initially, the features that were input into the classifier were divided into training and testing data sets. Fifty such splits were carried out in order to ensure the generalizability of the results. In the clustering step, k-means algorithm (Yang et al., 2003) was used to cluster the training data into n clusters. The number of clusters was first set to the number of features, and was progressively decreased by one until there were no empty clusters. The n obtained by this iteration served as the initial n for the RCE-SVM loop.

In the SVM scoring step, each cluster was scored based on its ability to differentiate the two categories by applying linear SVM. In order to rate the clusters, the training data was randomly partitioned into 10 non-overlapping subsets of equal sizes (10 folds). Using 9 subsets, the linear SVM was trained and performance was calculated using the remaining subset. Different
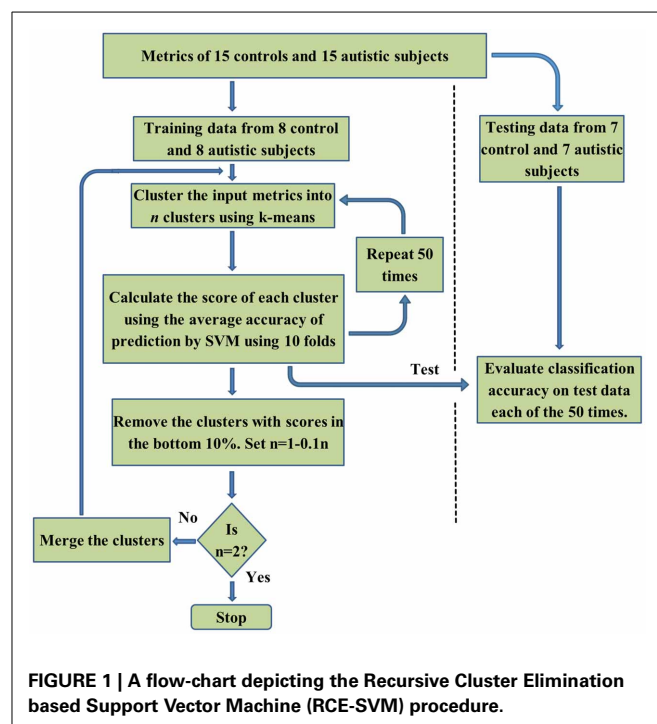

FIGURE 1 | A flow-chart depicting the Recursive Cluster Elimination based Support Vector Machine (RCE-SVM) procedure.

possible partitions were taken into account by repeating the clustering and cross validation procedure 50 times. For each of these 50 repetitions, the classification accuracy of SVM was ascertained using the test data. The average value of this accuracy, taking into account the repetitions and all the folds was assigned as the score of the cluster. The bottom 10% of low score clusters were eliminated in the RCE step. The remaining features were merged and the value of n was decreased by 10% and the cluster step, the SVM scoring step and the RCE step were repeated again in an iterative manner. After each iteration, the performance of the classifier was assessed using the testing data and lesser number of features compared to the earlier iteration. When the number of clusters was equal to two, the procedure was stopped. Complete separation of testing and training data in this algorithm eliminates bias in performance accuracy (Kriegeskorte et al., 2009). The accuracy at every RCE-SVM loop was calculated as a mean value of accuracy obtained over 50 repetitions of each loop and each train-test split, using the feature clusters of test data available at the corresponding loop and split. The statistical significance of mean accuracies was calculated by estimating the p-values of a binomial null distribution $B(\eta, \rho)$, η being the number of participants and ρ is the probability of accurate classification as in previous studies (Pereira et al., 2009). Only accuracies whose p-values were less than 0.05 after correcting for multiple comparisons using Bonferroni method were considered as statistically significant.

The causal connectivity weights obtained from the MVAR model, the behavioral assessment scores, the functional connectivity z-scores for each ROI pair, and DTI FA metrics for each of the 30 subjects (15 ASD and 15 TD) were input into the RCE-SVM classifier to determine the accuracy with which the classifier can predict a novel subject's group membership (autism or control).
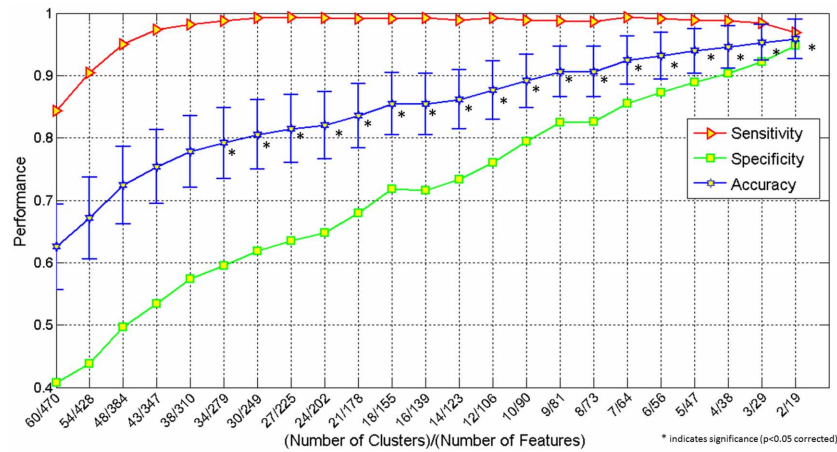
FIGURE 2 | Graph showing classification accuracy, sensitivity and specificity obtained by simultaneously using the following features: behavioral scores, functional connectivity, effective connectivity and fractional anisotropy obtained **from DTI.** The X-axis shows number of clusters/number of features and the Y-axis, the performance (classification accuracy, sensitivity and specificity). *indicates significance ($p < 0.05$ corrected).

RESULTS

The main results of this study are summarized as follows: (1) The effective connectivity path weights were able to successfully classify participants by diagnosis with 95.9% accuracy. These path weights were the most discriminative features among all the different metrics used in classification; (2) Effective connectivity paths most important for classification were significantly reduced ($p < 0.05$) in ASD participants compared to typical control participants; and (3) The paths that were among the top ranked features in the classification analysis were found to be negatively correlated with the AQ and positively correlated with the RME test scores.

The first set of results pertains to a pattern classification analysis involving several indices of connectivity (functional connectivity, effective connectivity, white matter integrity) and performance accuracy in this ToM task. In this analysis, utilizing 2 feature clusters comprised of 19 metrics, the classification accuracy reached a maximum accuracy of 95.9% (specificity 94.8%, and sensitivity 96.9%). It should be noted that all of the 19 features were effective connectivity paths. **Figure 2** demonstrates the increase in performance of classification with decreasing number of features (and removal of uninformative features). The p-values for all the accuracy values shown in **Figure 2** can be seen in **Table 2.**

Second, the causal connectivity weights of the 19 paths which led to maximum accuracy of 95.9% showed clear separation between participants with autism (blue) and typical control participants (green) as shown in **Figure 3**, with these paths showing significantly ($p < 0.05$ corrected using Bonferroni method for 18 paths; for one of the paths $p < 0.05$ uncorrected) weaker connectivity in participants with ASD compared to TD controls. Many of these connections are between regions that are part of the social brain network (LTPJ, RTPJ, LFFG, RFFG, LMTG, RMTG, RIFG) which may prove critical in accomplishing the ToM task used in this study. It is noteworthy that there may be other paths which

Table 2 | Classification accuracy values and the corresponding p-values obtained at each step of the RCE algorithm.

Accuracy	p-Value
0.625	0.100244
0.672	0.049369
0.724	0.008062
0.754	0.002611
0.778	0.002611
0.792	0.000715
0.805	0.000715
0.815	0.000715
0.82	0.000162
0.835	0.000162
0.855	2.97×10^{-05}
0.854	2.97×10^{-05}
0.862	2.97×10^{-05}
0.876	2.97×10^{-05}
0.892	4.22×10^{-06}
0.906	4.22×10^{-06}
0.906	4.22×10^{-06}
0.925	4.34×10^{-07}
0.932	4.34×10^{-07}
0.939	4.34×10^{-07}
0.945	4.34×10^{-07}
0.953	2.89×10^{-08}
0.959	2.89×10^{-08}

are significantly different between the groups. Here, we restrict ourselves to finding the statistical separation of features which have the highest ability for predicting the diagnosis of a given subject. We do so primarily because we are interested in features with predictive ability rather than those which just "differ" between the groups. Please refer to Appendix B **Figures B1**, **B2** in order to gain a qualitative understanding of the functional and effective
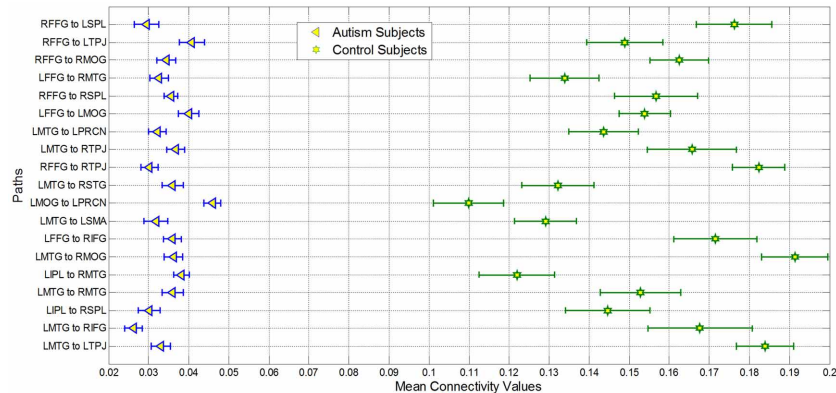
FIGURE 3 | Mean of nineteen paths which was most important for giving maximum classification accuracy for autism and control groups. All paths had significantly decreased connectivity ($p < 0.05$ corrected using Bonferroni method for 18 paths; for one of the paths $p < 0.05$ uncorrected) in the Autism group as compared to controls. The bars represent standard errors.
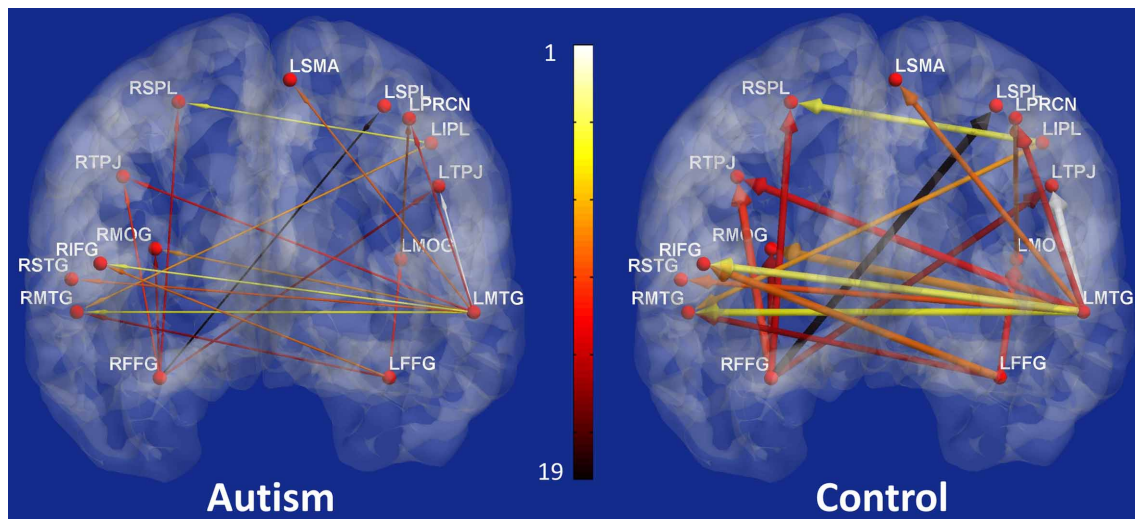

FIGURE 4 | The nineteen paths whose effective connectivity values were top-ranked features for classification of the two groups (Autism and Controls) with the maximum accuracy (Left panel: participants with autism; and right panel: control participants). The width of the arrows represents the path strength and the color of the path indicates its rank obtained during classification with 1 being the most significant and 19 being the least significant.

connectivity paths, respectively, between all 18 ROIs in both ASD and TD groups.

The 19 effective connectivity paths which were most important in classification are shown in **Figure 4**. The left panel shows these paths in ASD participants and the right panel in control participants. The width of the arrows illustrates the path weight in the corresponding group and the color represents the rank of the path obtained during classification.

Third, a correlation analysis was also performed between the features that were ranked highest in classification and gave rise to maximum accuracy, and assessment scores (AQ and RME). Given that the top-ranked features are not guaranteed to have normal distribution, we used Spearman's non-parametric correlation method to determine whether the top-ranked features were correlated with behavior. This analysis (including all participants in the study) revealed a significant negative correlation between

several effective connectivity paths and the AQ scores as well as a significant positive correlation between effective connectivity paths and RME scores (see **Table 3** for specific paths, correlation and p-values). These results suggest that as autism symptom severity increased, the effective connectivity of the top-ranked paths decreased; and as the theory-of-mind ability increased, effective connectivity of the top-ranked paths also increased. This provides a second-level test of the behavioral relevance of the top-ranked paths, which is to be expected given the fact that diagnosis was based on behavioral symptoms. As a cautionary note, these results should not be construed as a general discovery regarding brain connectivity features in autism which correlate with behavioral symptoms.

Neuroimaging data, especially brain connectivity analyses are prone to be influenced by head motion and signal quality. We conducted several different measures to make sure that our data

Table 3 | Paths correlated with Autism Quotient (AQ) scores and Reading Mind in Eye (RME) scores.

Paths correlated with Autism Quotient (AQ) scores					Paths correlated with Reading Mind in Eye (RME) scores				
Source region	→	Sink region	*p*-Value	Correlation coefficient	Source region	→	Sink region	*p*-Value	Correlation coefficient
LMTG	→	*LTPJ*	0.0003	−0.6137	*LMTG*	→	*LTPJ*	0.0311	0.3941
LMTG	→	*RIFG*	0.0007	−0.5837	*LMTG*	→	*RIFG*	0.0183	0.4279
LMTG	→	*RMTG*	0.0001	−0.6466	*LMTG*	→	*RMTG*	0.0455	0.3679
LMTG	→	*RMOG*	0.0006	−0.5901	*LMTG*	→	*RMOG*	0.0325	0.3912
LFFG	→	*RIFG*	0.0061	−0.4889	*LFFG*	→	*RIFG*	0.0428	0.3723
LMTG	→	*LSMA*	0.0006	−0.5930	*LMTG*	→	*LSMA*	0.0421	0.3734
RFFG	→	*RTPJ*	0.0028	−0.5274	*RFFG*	→	*RTPJ*	0.0373	0.3819
LFFG	→	*RMTG*	0.0119	−0.4531	*LFFG*	→	*RMTG*	0.0151	0.4395
RFFG	→	*RMOG*	0.0098	−0.4642	*RFFG*	→	*RMOG*	0.0035	0.5169
RFFG	→	*LSPL*	0.0187	−0.4266	*RFFG*	→	*LSPL*	0.0417	0.3741
LMTG	→	RSTG	0.0391	−0.3785	LIPL	→	RSPL	0.0082	0.4735
LMTG	→	RTPJ	0.0319	−0.3926	LFFG	→	LMOG	0.0213	0.4186
					RFFG	→	RSPL	0.0425	0.3727

The paths in red were correlated with both AQ and RME.

and the reported results were not influenced by quality related issues. First, the root mean square (RMS) values for each subject and each head motion parameter were obtained (see Appendix B **Table B1**). The RMS values were then submitted to a non-parametric Mann-Whitney U test, which also revealed no significant difference in motion in x [$U_{(28)} = 66$, $Z = −1.929$, $p = 0.06$], y [$U_{(28)} = 93$, $Z = −0.809$, $p = 0.42$], and z [$U_{(28)} = 96$, $Z = −0.684$, $p = 0.49$] translational directions. Nor was there a significant group difference in rotation in pitch [$U_{(28)} = 68$, $Z = −1.846$, $p = 0.06$], roll [$U_{(28)} = 107$, $Z = −0.228$, $p = 0.82$], and yaw [$U_{(28)} = 93$, $Z = −0.809$, $p = 0.42$]. These results indicate that there were no statistical differences in head motion between the two groups, assuming a p-value threshold of 0.05. However, there was a non-significant trend ($p = 0.06$) for translation in x direction and the degree of rotation in pitch to differ between the groups.

Further, we obtained the mean value of frame wise displacement (FD) for each subject as a quality control (QC) metric and investigated whether they correlated with any of the 19 top-ranked paths obtained from classification across the entire sample. The instantaneous motion of the head was expressed as a scalar quantity using the formula, $FD_i = |\Delta d_{ix}| + |\Delta d_{iy}| + |\Delta d_{iz}| + |\Delta \alpha_i| + |\Delta \beta_i| + |\Delta \gamma_i|$, where $\Delta d_{ik} = d_{(i-1)k} - d_{ik}$ and k is any of the 3 translational parameters (x, y, z) or rotational parameters (α, β, γ). We converted the rotational displacements from degrees to millimeters by calculating displacement on the surface of a sphere of radius 50 mm, assuming that the approximate mean distance from the center of the head to the cerebral cortex is 50 mm. The above procedure of calculating FD and correlating its mean with connectivity metrics obtained from individual subjects has been recommended recently for either confirming or ruling out the influence of head motion on connectivity measures (Power et al., 2012; Satterthwaite et al., 2012; Van Dijk et al., 2012; Satterthwaite et al., 2013; Yan et al., 2013). The QC-connectivity Spearman's correlations and corresponding p-values indicating their statistical significance are shown in **Table 4**. It is evident that none of the QC-connectivity correlations were statistically significant ($p > 0.05$). Given these

Table 4 | The Spearman's correlation between mean frame wise displacement (our quality control metric) and the Granger causality weights for the top ranked 19 paths.

Source region	→	Sink region	*p*-Value	Correlation coefficient
LMTG	→	LTPJ	0.869	−0.031
LMTG	→	RIFG	0.958	0.010
LIPL	→	RSPL	0.984	−0.003
LMTG	→	RMTG	0.701	−0.073
LIPL	→	RMTG	0.547	0.115
LMTG	→	RMOG	0.856	−0.035
LFFG	→	RIFG	0.309	0.192
LMTG	→	LSMA	0.741	−0.063
LMOG	→	LPRCN	0.827	0.042
LMTG	→	RSTG	0.725	−0.067
RFFG	→	RTPJ	0.599	0.099
LMTG	→	RTPJ	0.827	−0.042
LMTG	→	LPRCN	0.665	−0.083
LFFG	→	LMOG	0.702	−0.073
RFFG	→	RSPL	0.322	−0.187
LFFG	→	RMTG	0.623	0.094
RFFG	→	RMOG	0.404	−0.158
RFFG	→	LTPJ	0.703	−0.073
RFFG	→	LSPL	0.866	0.032

The paths shown in the table are ordered according to the rank obtained during classification with 1 being the most significant and 19 being the least significant (first path is Rank −1 and the last path is Rank −19).

evidence, any significant group differences for imaging metrics was probably not due to head motion. We did not use the scrubbing method described in Power et al. (2012), where removal of certain parts of the time series (scrubbing) creates an artificial discontinuity in the data. This may not be a problem while using Pearson's correlation coefficient as zero-lag synchronization in the data does not depend on the temporal ordering in the data as long as the correspondence between the variables being examined is preserved. However, other methods which are sensitive to temporal ordering in the data cannot use scrubbing.

Granger causality is one such method which is indeed sensitive to temporal ordering in the data and hence we did not use scrubbing.

Differences in signal to noise ratio (SNR) can also impact Granger causality estimates (Nalatore et al., 2007) when the SNR is low. On the other hand, when SNR = 2, which is typically the case for task-based fMRI, we have previously showed using simulations that Granger causality estimates are accurate in the absence of hemodynamic variability (which is the case here since we deconvolved the hemodynamic response) (Deshpande et al., 2010b). We calculated effective SNR of the deconvolved fMRI time series by estimating the variance of the entire deconvolved signal, i.e., the hidden neuronal variable, and divided it by the variance of the deconvolved signal during non-stimulation phases. We then populated the SNRs of each ROI in autism and control groups to two different samples and performed a non-parametric Wilcoxon ranksum to find statistical differences. The SNR was significantly higher ($p < 0.05$, z-value = 20.1) in the ASD group (SNR = 4.13 ± 0.01) as compared to the control group (SNR = 3.2 ± 0.03). The SNRs for both groups were high enough so that SNR differences between the groups will not impact Granger causality. SNR has an impact on Granger causality only when the SNR is low.

DISCUSSION

The goals of this study were: (1) to investigate effective connectivity among brain areas during intentional causal attribution in ASD and (2) to utilize machine learning techniques to classify participants based on effective connectivity weights from this study, and behavior assessment scores, functional connectivity, and fractional anisotropy obtained from DTI data from our previous study (Kana et al., 2012). Using SVM based classification, we found that the causal connectivity path weights had the highest discriminative power to separate groups by diagnosis with high accuracy. It was uncovered that the top-ranked causal connectivity paths were also significantly weaker between social brain regions in young adults with ASD as compared to their TD peers and correlated with the ASD symptom severity (AQ) scores and theory-of-mind ability as measured by the RME test.

An application of characterizing brain connectivity patterns is to test whether such patterns can differentiate individuals with ASD from typically developing control participants such that the diagnostic label of a new participant can be determined based on imaging data. Thus, in this study we conducted a classification analysis using the effective connectivity measures, functional connectivity values, fractional anisotropy obtained from DTI data and the causal attribution task performance scores to get a fair assessment of which metric possesses the highest discriminative power. A maximum classification accuracy of 95.9% was obtained with 2 clusters and 19 features, all of these being effective connectivity paths. These results suggest that significantly weaker causal influence between brain regions during ToM processing in ASD is sufficient to separate adults with ASD from typical control participants. The discriminative patterns found in this study using SVM may have clinical applications in the long-run. Accurate separation of ASD adults from TD peers may provide potential value for clinicians, particularly in cases when behavioral observation and clinical interviews are not sufficient enough to determine a

diagnosis. The key finding of differences in the causal influence of brain regions for ToM in ASD in this study adds to the relatively limited literature on effective connectivity in ASD. In addition, while previous studies explored effective connectivity in ASD during language processing, facial and emotional processing, and imitation, the current study examined effective connectivity in the context of a ToM task, which has not been studied in ASD to date. The current study expands what we know about inferring mental states in ASD, and provides insight into the causal relationships of brain regions during ToM processing. In addition, this study, to our knowledge, is the first to use effective connectivity measures for classification purposes in ASD. While this method will require some fine tuning, validation in a larger sample, and replication through multiple studies to be applied within clinical settings, the causal relationships between brain areas related to ToM holds promise for separating individuals with ASD from typical controls or from other disorders. Nevertheless, the current study marks the first attempt at using effective connectivity measures as inputs for a classification analysis of ASD subjects, therefore marking the first step in the direction of more accurate classification of the disorder.

Weaker effective connectivity of the 19 top-ranked paths found in participants with ASD in this study involved paths and regions that are found to be part of the social brain network. Several nodes, such as the TPJ, MTG, RIFG, IPL, FFG, and SMA have been associated with processing theory-of-mind, face processing, and the mirror neuron system. These findings are in line with previous studies of effective connectivity in ASD (Wicker et al., 2008; Shih et al., 2010). Our results also include significant functional alterations in social brain and visuospatial brain regions (e.g., TPJ, IFG, IPL, FFG, etc.) seen previously in functional connectivity findings (Kana et al., 2006, 2009, 2012; Just et al., 2007; Koshino et al., 2008; Mason et al., 2008), suggesting some consistency in disrupted connectivity across different modalities of connectivity and providing further support for disrupted connectivity accounts of ASD (Just et al., 2004; Kana and Just, 2011; Schipul et al., 2011; Kana et al., 2012). The findings here supplement the functional connectivity results in our previous study utilizing the same ToM stimuli, where ASD participants displayed significantly reduced functional connectivity between temporal and frontal regions, and weaker connectivity between networks made up of ventral premotor regions and TPJ (Kana et al., 2012). Our results in the current study further these previous findings by illustrating the directionality of connectivity. We found that, for ToM processing in TD participants, significantly stronger (compared to ASD group) causal connections existed among the 19 top-ranked paths which included the nodes that are associated with social cognition. So, here we find that the critical regions of the social brain are not as well coordinated with others, that they should be sharing information with, in participants with ASD. This lack of synchrony and reduced flow of information may represent a critical problem of bandwidth (maximal rate of data transfer supported by a communication channel) in ASD, where some information is getting by, but at a much lower rate than what would be needed for complex ToM connections (Just et al., 2012).

In a correlation analysis using assessment measures and effective connectivity paths for the entire sample of participants, we

found the paths that were among the top ranked features in the classification analysis were correlated with AQ and RME scores. While the AQ showed significant negative correlation, the RME showed significant positive correlation with connectivity paths. Similarly, participants with better ToM skills had stronger effective connectivity during this causal attribution task. It should be noted that most of these connection paths involved information transfer to different regions mainly from the temporal lobe (LMTG, and bilateral FFG). While FFG has been associated with face processing and processing socially salient stimuli (Schultz, 2005), middle and superior temporal areas have been found to be involved in social cognition, especially in taking intentional stance, as seen in the current study, on social scenarios (Mosconi et al., 2005). The correlations found in our study reveal how social abilities such as ToM skill can influence information transfer in the brain. In addition, it also points out that severe autism symptoms may have a neural basis in reduced causal brain connectivity from the temporal lobe. As noted earlier, the correlation analysis was performed across the entire sample and we restricted it to the top-ranked 19 paths because we feel that the covariance of a brain imaging based metric with a behavioral assessment

score is clinically meaningful only if the imaging metric under consideration has the power to predict the diagnostic label of a new subject. Therefore, there may very well be other connectivity paths in the brain which may be correlated with behavior (but which lack the discriminative ability) which we have not discussed here.

In conclusion, this study provides preliminary evidence to support a hypothesis that metrics based on directional brain connectivity obtained from a task engaging social brain areas may provide highly discriminative features for predicting whether a given subject has ASD or not. Studies involving larger sample size, and replication of these findings across multiple studies would be required to fully test the extent of this hypothesis and investigate its clinical implications.

REFERENCES

Abler, B., Roebroeck, A., Goebel, R., Hose, A., Schnfeldt-Lecuona, C., Hole, G., et al. (2006). Investigating directed influences between activated brain areas in a motor-response task using fMRI. *Magn. Reson. Imaging* 24, 181–185. doi: 10.1016/j.mri.2005.10.022

Alexander, A. L., Lee, J. E., Lazar, M., Boudos, R., DuBray, M. B., Oakes, T. R., et al. (2007). Diffusion tensor imaging of the corpus callosum in autism. *Neuroimage* 34, 61–73. doi: 10.1016/j.neuroimage.2006.08.032

Anderson, J. S., Nielsen, J. A., Froehlich, A. L., DuBray, M. B., Druzgal, T. J., Cariello, A. N., et al. (2011). Functional connectivity magnetic resonance imaging classification of autism. *Brain* 134(Pt. 12), 3742–3754. doi: 10.1093/brain/awr263

Asperger, H. (1944). "Die 'autistichen Psychopathen' im Kindersalter," in *Archive fur psychiatrie und Nervenkrankheiten*, Vol. 117, Reprinted in Autism and Asperger syndrome, ed U. Frith (Cambridge: Cambridge University Press), 76–136.

Assaf, M., Jagannathan, K., Calhoun, V., Miller, L., Stevens, M. C., Sahl, R., et al. (2010). Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *Neuroimage* 53, 247–256. doi: 10.1016/j.neuroimage.2010.05.067

Barnea-Goraly, N., Kwon, G., Menon, B., Eliez, S., Lotspeich, L., and Reiss, A. L. (2004). White matter structure

in autism: preliminary evidence from diffusion tensor imaging. *Biol. Psychiatry* 55, 323–326. doi: 10.1016/j.biopsych.2003.10.022

Barnea-Goraly, N., Lotspeich, L. J., and Reiss, A. L. (2010). Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract based spatial statistics. *Arch. Gen. Psychiatry* 67, 1052–1060. doi: 10.1001/archgenpsychiatry.2010.123

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001a). The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001b). The autism spectrum quotient (AQ): evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/A:1005653411471

Bird, G., Catmur, C., Silani, G., Frith, C., and Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders. *Neuroimage* 31, 1614–1624. doi: 10.1016/j.neuroimage.2006.02.037

Brunet, E., Sarfati, Y., Hardy-Baylé, M. C., and Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal

task. *Neuroimage* 11, 157–166. doi: 10.1006/nimg.1999.0525

Buchel, C., and Friston, K. J. (2000). Assessing interactions among neuronal systems using functional neuroimaging. *Neural Netw.* 13, 871–882. doi: 10.1016/S0893-6080(00)00066-6

Calderoni, S., Reticom, A., Biagi, L., Tancredi, R., Muratori, F., and Tosetti, M. (2012). Female children with autism spectrum disorder: an insight from mass-univariate and pattern classification analyses. *Neuroimage* 59, 1013–1022. doi: 10.1016/j.neuroimage.2011.08.070

Chen, G., Ward, D., Xie, C., Li, W., Wu, Z., Jones, J., et al. (2011). Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology* 259, 213–221. doi: 10.1148/radiol.10100734

Cherkassky, V., Kana, R., Keller, T., and Just, M. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport* 17, 1687–1690. doi: 10.1097/01.wnr.0000239956.45448.4c

Coutanche, M. N., Thompson-Schill, S. L., and Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *Neuroimage* 57, 113–123. doi:10.1016/j.neuroimage.2011.04.016

Craddock, R., Holtzheimer, I. P., Hu, X., and Mayberg, H. (2009). Disease state prediction from resting state functional connectivity. *Magn.*

Reson. Med. 62, 1619–1628. doi: 10.1002/mrm.22159

Cubillo, A., Rozmin, H., Ecker, C., Giampietro, V., Taylor, E., and Rubia, K. (2010). Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHS) and persisting symptoms during tasks of motor inhibition and cognitive switching. *J. Psychiatry Res.* 44, 629–639. doi: 10.1016/j.jpsychires.2009.11.016

Damarla, S. R., Keller, T. A., Kana, R. K., Cherkassky, V. L., Williams, D. L., Minshew, N. J., et al. (2010). Cortical underconnectivity coupled with preserved visuospatial cognition in autism: evidence from an fMRI study of an embedded figures task. *Autism Res.* 3, 273–279. doi: 10.1002/aur.153

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C., et al. (2008). Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biol.* 6:2683–2697. doi: 10.1371/journal.pbio.0060315

Deshpande, G., Hu, X., Stilla, R., and Sathian, K. (2008). Effective connectivity during haptic perception: a study using granger causality analysis of functional magnetic resonance imaging data. *Neuroimage* 40, 1807–1814. doi: 10.1016/j.neuroimage.2008.01.044

Deshpande, G., James, G., Craddock, R., Mayberg, H., and Hu, X. (2009a). "Predicting treatment in patients with major depression

using granger-based connectivity and support vector machines," in *Proceedings of ISMRM 17th Scientific Meeting 2009* (Honolulu, HI), 3362.

Deshpande, G., LaConte, S., James, G., Peltier, S., and Hu, X. (2009b). Multivariate Granger causality analysis of brain networks. *Hum. Brain Mapp.* 30, 1361–1373. doi: 10.1002/hbm.20606

Deshpande, G., Li, Z., Santhanam, P., Coles, C. D., Lynch, M. E., Hamann, S., et al. (2010a). Recursive cluster elimination based support vector machine for disease state prediction using resting state functional and effective brain connectivity. *PLoS ONE* 5:e14277. doi: 10.1371/journal.pone.0014277

Deshpande, G., Sathian, K., and Hu, X. (2010b). Effect of hemodynamic variability on Granger causality analysis of fMRI. *Neuroimage* 52, 884–896. doi: 10.1016/j.neuroimage.2009.11.060

Deshpande, G., Sathian, K., and Hu, X. (2010c). Assessing and compensating for zero-lag correlation effects in time-lagged granger causality analysis of fMRI. *IEEE Trans. Biomed. Eng.* 57, 1446–1456. doi: 10.1109/TBME.2009.2037808

Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361. doi: 10.1126/science.1194144

Duffy, F. H., and Als, H. (2012). A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls – a large case control study. *BMC Med.* 10, 64. doi: 10.1186/1741-7015-10-64

Friston, J. K., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D., and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Friston, K., Harrison, L., and Penny, W. (2003). Dynamic causal modeling. *Neuroimage* 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7

Garrity, A. G., Pearlson, G. D., McKiernan, K., Lloyd, D., Kiehl, K. A., and Calhoun, V. D. (2007). Aberrant "default mode" functional connectivity in schizophrenia. *Am. J. Psychiatry* 164, 450–457. doi: 10.1176/appi.ajp.164.3.450

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods.

Econometrica 37, 424–438. doi: 10.2307/1912791

Handwerker, D., Ollinger, J., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651. doi: 10.1016/j.neuroimage.2003.11.029

Havlicek, M., Friston, K., Jan, J., Brazdil, M., and Calhoun, V. (2011). Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering. *Neuroimage* 56, 2109–2128. doi: 10.1016/j.neuroimage.2011.03.005

Ingalhalikar, M., Parker, D., Bloy, L., Roberts, T. P. L., and Verma, R. (2011). Diffusion based abnormality markers of pathology: toward learned diagnostic prediction of ASD. *Neuroimage* 57, 918–927. doi: 10.1016/j.neuroimage.2011.05.023

Jones, T. B., Bandettini, P. A., Kenworthy, L., Case, L. K., Milleville, S. C., Martin, A., et al. (2010). Sources of group differences in functional connectivity: an investigation applied to autism spectrum disorder. *Neuroimage* 49, 401–414. doi: 10.1016/j.neuroimage.2009.07.051

Jou, R., Jackowski, A., Papademetris, X., Rajeevan, N., Staib, L., and Volkmar, F. (2011). Diffusion tensor imaging in autism spectrum disorders: preliminary evidence of abnormal neural connectivity. *Aust. N.Z. J. Psychiatry* 45, 153–162. doi: 10.3109/00048674.2010.534069

Just, M., Cherkassky, V., Keller, T., and Minshew, N. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain* 127, 1811–1821. doi: 10.1093/brain/awh199

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., and Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cereb. Cortex* 17, 951–961. doi: 10.1093/cercor/bhl006

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., and Varma, S. (2012). Autism as a neural systems disorder: a theory of frontal-posterior underconnectivity. *Neurosci. Biobehav. Rev.* 36, 1292–1313. doi: 10.1016/j.neubiorev.2012.02.007

Kana, R., Keller, T., Cherkassky, V., Minshew, N., and Just, M. (2006). Sentence comprehension in autism: thinking in pictures with decreased functional

connectivity. *Brain* 129, 2484–2493. doi: 10.1093/brain/awl164

Kana, R. K., and Just, M. A. (2011). "Autism as a disorder of functional brain connectivity," in *Handbook of Autism Spectrum Disorders*, eds D. G. Amaral, D. Geschwind, and G. Dawson (New York, NY: Oxford University Press), 981–989. doi: 10.1093/med/9780195371826.003.0061

Kana, R. K., Keller, T. A., Minshew, N. J., and Just, M. A. (2007). Inhibitory control in high-functioning autism: decreased activation and underconnectivity in inhibition networks. *Biol. Psychiatry* 62, 198–206. doi: 10.1016/j.biopsych.2006.08.004

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., and Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory-of-Mind regions in autism during mental state attribution. *Soc. Neurosci.* 4, 135–152. doi: 10.1080/17470910802198510

Kana, R. K., Libero, L. E., and Moore, M. S. (2011). Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. *Phys. Life Rev.* 8, 410–437. doi: 10.1016/j.plrev.2011.10.001

Kana, R. K., Libero, L. E., Hu, C. P., Deshpande, H. D., and Colburn, J. S. (2012). Functional brain networks and white matter underlying theory-of-mind in autism. *Soc. Cogn. Affect. Neurosci.* 7, 1–8. doi: 10.1093/scan/nss106

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child* 2, 217–250.

Keller, T. A., Kana, R. K., and Just, M. A. (2007). A developmental study of the structural integrity of white matter in autism. *Neuroreport* 18, 23–27. doi: 10.1097/01.wnr.0000239965.21685.99

Koshino, H., Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., and Just, M. A. (2008). fMRI investigation of working memory for faces in autism: visual coding and underconnectivity with frontal areas. *Cereb. Cortex* 18, 289–300. doi: 10.1093/cercor/bhm054

Kriegeskorte, N., Simmons, W., Bellgowan, P., and Baker, C. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303

Kuś, R., Kamiński, M., and Blinowska, K. (2004). Determination of EEG activity propagation: pair-wise versus multichannel estimate. *IEEE Trans. Biomed. Eng.* 51, 1501–1510. doi: 10.1109/TBME.2004.827929

Lawrie, S. M., Buechel, C., Whalley, H. C., Frith, C. D., Friston, K. J., and Johnstone, E. C. (2002). Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. *Biol. Psychiatry* 51, 1008–1011. doi: 10.1016/S0006-3223(01)01316-1

Lieberman, M., and Cunningham, W. (2009). Type I and Type II error concerns in fMRI research: rebalancing the scale. *Soc. Cogn. Affect. Neurosci.* 4, 423–428. doi: 10.1093/scan/nsp052

Marquand, A.F., Filippone, M., Ashburner, J., Girolami, M., Mourao-Miranda, J., Barker, G. L., et al. (2013). Automated, high accuracy classification of parkinsonian disorders: a pattern recognition approach. *PLoS ONE* 8:e69237. doi:10.1371/journal.pone.0069237

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., and Just, M. A. (2008). Theory-of-Mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. *Neuropsychologia* 46, 269–280. doi: 10.1016/j.neuropsychologia.2007.07.018

McIntosh, A., and Gozales-Lima, F. (1994). Structural equation modelling and its application to network analysis in functional brain imaging. *Hum. Brain Mapp.* 2, 2–22. doi: 10.1002/hbm.460020104

McIntosh, A. R., Kovacevic, N., Lippe, S., Garrett, D., Grady, C., and Jirsa, V. (2010). The development of a noisy brain. *Arch. Ital. Biol.* 148, 323–337.

Meyer-Lindenberg, A. S., Olsen, R. K., Kohn, P. D., Brown, T., Egan, M. D., Weinberger, D. R., et al. (2005). Regionally specific disturbance of dorsolateral prefrontal-hippocampal functional connectivity in schizophrenia. *Arch. Gen. Psychiatry* 62, 379–386. doi: 10.1001/archpsych.62.4.379.

Mirowski, P., Madhavan, D., LeCun, Y., and Kuzniecky, R. (2009). Classification of patterns of EEG synchronization for seizure prediction. *Clin. Neurophysiol.* 120, 1927–1940. doi: 10.1016/j.clinph.2009.09.002

Mizuno, A., Liu, Y., Williams, D. L., Keller, T. A., Minshew, N. J., and Just, M. A. (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. *Brain* 134(Pt. 8), 2422–2435. doi: 10.1093/brain/awr151

Mosconi, M. W., Mack, P. B., McCarthy, G., and Pelphrey, K. A. (2005). Taking an "intentional

stance" on eye-gaze shifts: a functional neuroimaging study of social perception in children. *Neuroimage* 27, 247–252. doi: 10.1016/j.neuroimage.2005.03.027

Murdaugh, D. L., Shinkareva, S. V., Deshpande, H. R., Wang, J., Pennick, M. R., and Kana, R. K. (2012). Differential deactivation during mentalizing and classification of autism based on default mode connectivity. *PLoS ONE* 7:e50064. doi: 10.1371/journal.pone.0050064

Nalatore, H., Ding, M., and Rangarajan, G. (2007). Mitigating the effects of measurement noise on Granger causality. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 75(3 Pt 1), 031123. doi: 10.1103/PhysRevE.75.031123

Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209. doi: 10.1016/j.neuroimage.2008.11.007

Power, J., Barnes, K., Snyder, A., Schlaggar, B., and Petersen, S. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230– 242. doi: 10.1016/j.neuroimage.2004.11.017

Sathian, K., Lacey, S., Stilla, R., Gibson, G., Deshpande, G., Hu, X., et al. (2011). Dual pathways for haptic and visual perception of spatial and texture information. *Neuroimage* 57, 462–475. doi: 10.1016/j.neuroimage.2011.05.001

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data.

Neuroimage 64, 240–256. doi: 10.1016/j.neuroimage.2012.08.052

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., et al. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage* 60, 623–632. doi: 10.1016/j.neuroimage.2011.12.063

Schipul, S., Keller, T., and Just, M. (2011). Inter-regional brain communication and its disturbance in autism. *Front. Syst. Neurosci.* 5:10. doi: 10.3389/fnsys.2011.00010

Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *Int. J. Dev. Neurosci.* 23, 125–141. doi: 10.1016/j.ijdevneu.2004.12.012

Shen, M. D., Shih, P., Öttl, B., Keehn, B., Leyden, K. M., Gaffrey, M. S., et al. (2012). Atypical lexicosematntic function of extrastriate cortex in autism spectrum disorder: evidence from functional and effective connectivity. *Neuroimage* 62, 1780–1791. doi: 10.1016/j.neuroimage.2012.06.008

Shih, P., Shen, M., Öttl, B., Keehn, B., Gaffrey, M., and Muller, R-A. (2010). Atypical network connectivity for imitation in autism spectrum disorder. *Neuropsychologia* 48, 2931–2939. doi: 10.1016/j.neuropsychologia.2010.05.035

Shinkareva, S. V., Wang, J., and Wedell, D. H. (2013). Examining similarity structure: multidimensional scaling and related approaches in neuroimaging. *Comput. Math. Methods Med.* 2013, 796183. doi: 10.1155/2013/796183

Solomon, M., Ozonoff, S. J., Ursu, S., Ravizza, S., Cumings, N., Ly, S., et al. (2009). The neural substrates of cognitive control deficits in autism spectrum disorders. *Neuropsychologia* 47, 2515–2526. doi: 10.1016/j.neuropsychologia.2009.04.019

Tian, L., Jiang, T., Wang, Y., Zang, Y., He, Y., Liang, M., et al. (2006).

Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder. *Neurosci. Lett.* 400, 39–43. doi: 10.1016/j.neulet.2006.02.022

Travers, B. G., Adluru, N., Ennis, C., Tromp do, P. M., Destiche, D., Doran, S., et al. (2012). Diffusion tensor imaging in autism spectrum disorder: a review. *Autism Res.* 5, 289–313. doi: 10.1002/aur.1243

Uddin, L., Vinod, M., Young, C. B., Ryali, S., Chen, T., Khouzam, A., et al. (2011). Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism. *Biol. Psychiatry* 70, 833–841. doi: 10.1016/j.biopsych.2011.07.014

Van Dijk, K. R. A., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438. doi: 10.1016/j.neuroimage.2011.07.044

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer. doi: 10.1007/978-1-4757-2440-0

Villalobos, M. E., Mizuno, A., Dahl, B. C., Kemmotsu, N., and Muller, R.-A. (2005). Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *Neuroimage* 25, 916–925. doi: 10.1016/j.neuroimage.2004.12.022

Wang, L. (2005). *Support Vector Machines: Theory and Applications*. New York, NY: Springer.

Wang, G., Chen, C., and Fushing, H. (2012). Extracting multiscale pattern information of fMRI based functional brain connectivity with application on classification of autism spectrum disorders. *PLoS ONE* 7:e45502. doi: 10.1371/journal.pone.0045502

Weygandt, M., Hackmack, K., Pfüller, C., Bellmann-Strobl, J., Paul, F., Zipp, F., et al. (2011). MRI pattern recognition in multiple sclerosis normal-appearing brain areas. *PLoS ONE* 6:e21138. doi: 10.1371/journal.pone.0021138

Wicker, B., Fonlupt, P., Hubert, B., Tardif, C., Gepner, B., and Deruelle, C. (2008). Abnormal cerebral effective connectivity during explicit emotional processing in adults with autism spectrum disorder. *Soc. Cogn. Affect. Neurosci.* 3, 135–143. doi: 10.1093/scan/nsn007

Yan, C. G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R. C., Di Martino, A., et al. (2013). A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage* 76, 183–201. doi: 10.1016/j.neuroimage.2013.03.004

Yang, X., Lin, D., Hao, Z., Liang, Y., Liu, G., and Han, X. (2003). A fast SVM training algorithm based on the set segmentation and k-means clustering. *Prog. Nat. Sci.* 13, 750–755. doi: 10.1080/10020070312331344360

APPENDIX A

IMAGE ACQUISITION

Structural Imaging: Acquisition of initial high-resolution T1-weighted scans was done using a 160-slice 3D MPRAGE volume scan with $TR = 200\,ms$, $TE = 3.34\,ms$, flip angle = 78, field of view (FOV) = 25.6 cm, matrix size = 256×256 and slice thickness = 1 mm.

Diffusion tensor imaging: A diffusion-weighted, single-shot, spin-echo, echo-planar imaging (EPI) sequence ($TR = 4400\,ms$, $TE = 85\,ms$, bandwidth = 1860 Hz/voxel, FOV = 240 mm and 128×128 matrix size, resulting in an in-plane resolution of $1.87 \times 1.87 \times 3\,mm$) was used to collect the images. Thirty-two 3 mm thick slices were imaged (no slice gap) with no diffusion weighting ($b = 0$ s/mm^2) and with diffusion weighting ($b = 1000$ s/mm^2) gradients applied in 12 orthogonal directions. Twenty-four images of each slice by gradient direction combination were acquired and averaged to produce the final diffusion imaging data set.

DATA ANALYSIS

The brain activation data were analyzed using Statistical Parametric Mapping (SPM8) software (Wellcome Department of Cognitive Neurology, London, UK). Images were corrected for slice acquisition timing, motion-corrected, normalized to the Montreal Neurological Institute (MNI) template, resampled to $2 \times 2 \times 2$ mm voxels, and smoothed with an 8-mm Gaussian kernel to decrease spatial noise. We performed statistical analysis on individual and group data by using SPM8's implementation of the general linear model (Friston et al., 1995). Within-group activation was analyzed for the ASD group, TD group, and for the whole group (ASD + TD) of participants. Activation data was analyzed for all trials with separate regressors defined for intentional causality, physical causality, and fixation baseline conditions. The within-group analyses used a cluster size of 80 mm^3 determined by 10,000 Monte Carlo simulations at an uncorrected p value of 0.001. According to Lieberman and Cunningham (2009), simulations can implicate cluster size thresholds that produce the best balance between Type I and Type II error. The between-group analyses used a cluster threshold of 10 contiguous voxels at an uncorrected p value of 0.005, as the effects did not survive a more conservative statistical threshold.

ROIs were defined on the group activation map for the whole group (ASD + TD) for the contrast Intention + Physical vs. Fixation, so that it best represents the study. Eighteen ROIs were identified: supplementary motor area (SMA), left and right inferior frontal gyrus (LIFG, RIFG), left and right ventral premotor cortex (LPMv, RPMv), left and right middle temporal gyrus (LMTG, RMTG), right superior temporal gyrus (RSTG), left and right inferior parietal lobule (LIPL, RIPL), left and right fusiform gyrus (LFFG, RFFG), left and right superior parietal lobule (LSPL, RSPL), left and right middle occipital gyrus (LMOG, RMOG), and left and right temporal parietal junction (LTPJ, RTPJ). A sphere was defined for each cluster (with a radius ranging from 8 to 12 mm) that best captured the cluster of activation in the contrast map for each group. The activation time-course extracted for each participant over the activated voxels within the ROI

originated from the normalized and smoothed images that were low-pass filtered and had the linear trend removed.

APPENDIX B

Table B1 | Root mean square values of head motion.

|	Translation			Rotation		
	X	Y	Z	Pitch	Roll	Yaw
ASD						
1	0.1456	0.1073	0.2415	0.0024	0.0013	0.0020
1	0.1234	0.0910	0.1028	0.0010	0.0013	0.0017
1	0.1009	0.1698	0.2096	0.0036	0.0023	0.0012
1	0.1521	0.2020	0.1704	0.0013	0.0015	0.0009
1	0.1216	0.1725	0.2679	0.0031	0.0014	0.0015
1	0.1090	0.1473	0.3443	0.0037	0.0011	0.0018
1	0.0639	0.0863	0.2077	0.0027	0.0013	0.0005
1	0.1082	0.0714	0.2311	0.0024	0.0012	0.0014
1	0.2747	0.2213	0.3942	0.0074	0.0023	0.0030
1	0.0426	0.0893	0.1033	0.0019	0.0006	0.0006
1	0.0788	0.1224	0.4748	0.0026	0.0010	0.0015
1	0.0527	0.1589	0.2260	0.0016	0.0012	0.0007
1	0.0322	0.2459	0.1560	0.0016	0.0004	0.0004
1	0.0452	0.0832	0.2903	0.0033	0.0009	0.0010
1	0.1833	0.5282	0.5859	0.0064	0.0027	0.0023
TD						
2	0.1191	0.0812	0.1415	0.0013	0.0022	0.0020
2	0.0657	0.1096	0.2497	0.0017	0.0022	0.0013
2	0.0998	0.0743	0.2858	0.0025	0.0009	0.0013
2	0.0965	0.1360	0.3082	0.0023	0.0016	0.0013
2	0.0231	0.1937	0.2289	0.0025	0.0010	0.0006
2	0.3144	0.3258	0.4231	0.0056	0.0021	0.0044
2	0.0376	0.1298	0.1110	0.0018	0.0003	0.0007
2	0.0359	0.1012	0.2624	0.0012	0.0010	0.0006
2	0.0381	0.1800	0.1977	0.0022	0.0023	0.0011
2	0.0281	0.0762	0.1060	0.0015	0.0008	0.0006
2	0.0433	0.0745	0.1944	0.0018	0.0006	0.0009
2	0.0969	0.1548	0.2143	0.0015	0.0023	0.0011
2	0.0566	0.0831	0.1272	0.0018	0.0006	0.0010
2	0.0564	0.1403	0.2922	0.0018	0.0022	0.0014
2	0.0828	0.2080	0.1962	0.0016	0.0012	0.0011
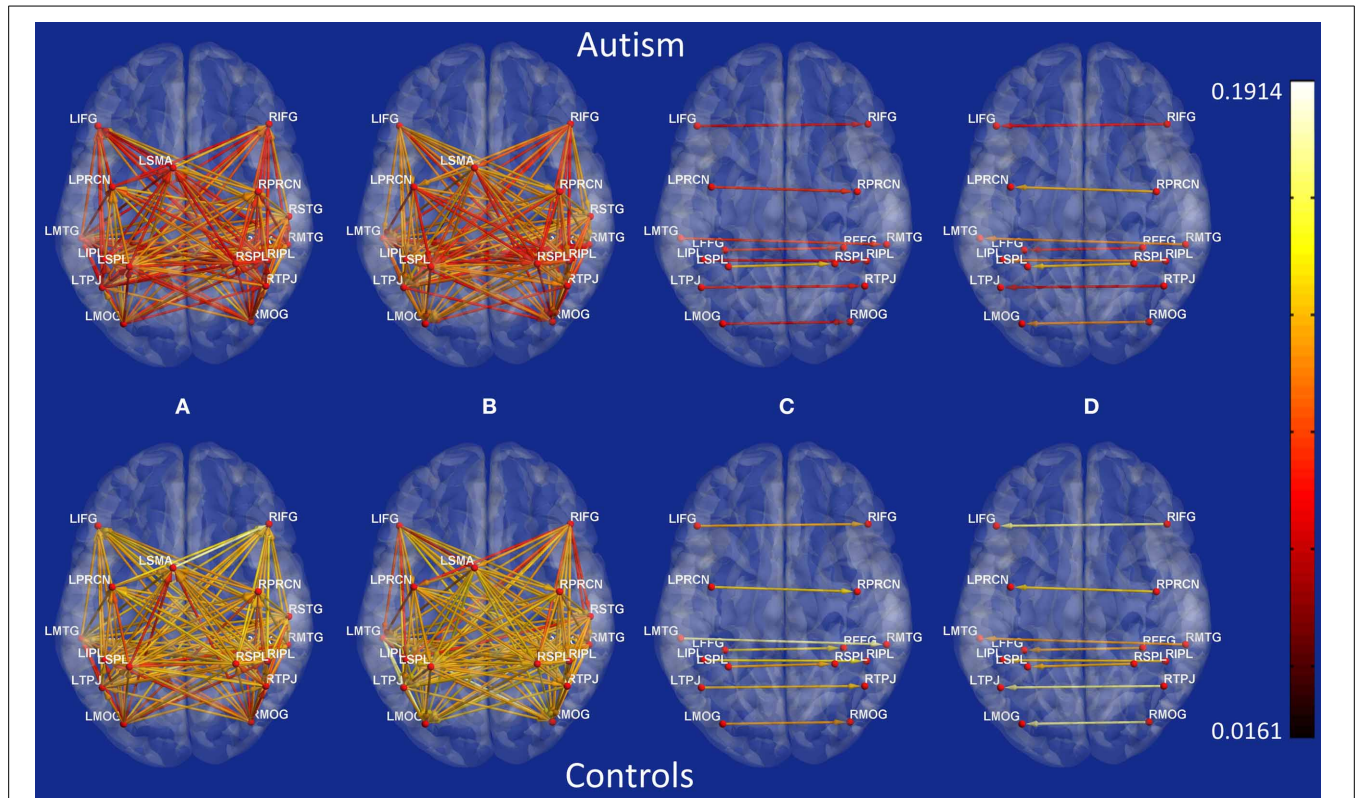
FIGURE B1 | Connectivity maps showing Granger causality path weights for all possible connections between 18 ROIs. Top: Autism, Bottom: Controls. **(A)** posterior to anterior paths, **(B)** anterior to posterior paths, **(C)** left to right paths, **(D)** right to left paths.
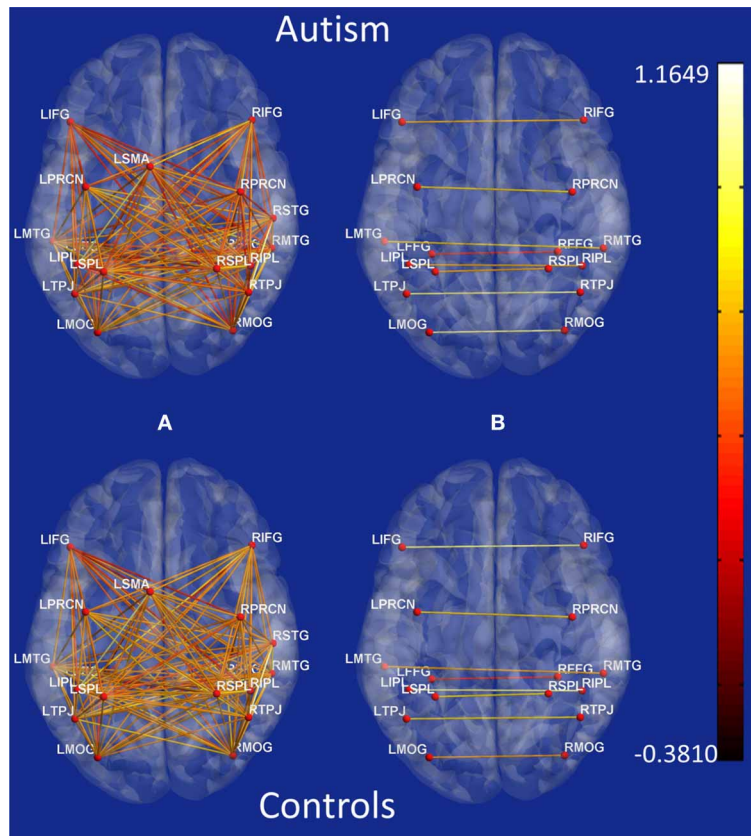
FIGURE B2 | Connectivity maps showing *z*-scores of functional connectivity path weights obtained using Pearson's correlation for all possible connections between 18 ROIs. Top: Autism, Bottom: Controls. **(A)** anterior–posterior paths, **(B)** bilateral paths.