Supplementary Methods and Analyses

Subjects

Subjects were recruited through referrals from the UCLA Autism Clinic, through flyers posted at UCLA and in the Los Angeles area, and from a pool of subjects who had previously participated in other research studies at UCLA. Inclusion criteria for the ASD group included: (1) a clinical diagnosis of ASD confirmed using the Autism Diagnostic Observation Schedule-Generic (ADOS-G) [1] and the Autism Diagnostic Observation Interview-Revised (ADI-R) [2], (2) no other diagnosed neurological disorders (e.g. cerebral palsy or epilepsy), (3) no structural brain abnormality, and (4) fluent language abilities. Typically developing subjects (TD) had no history of medical, psychiatric, or neurological disorders according to parental report. Neither age nor IQ differed significantly between ASD and TD groups. Shown in **Supplementary Table 1** are the mean Verbal, Performance, and Full-Scale IQ (assessed by the Wechsler Intelligence Scale for Children – Third Edition or the Wechsler Abbreviated Scale of Intelligence) for both ASD and TD groups. Also shown in this table are the mean scores on the social subscales of the ADOS and ADI for the ASD group.

Activation Paradigm

Subjects underwent two consecutive fMRI scans, each 4 min 54 sec in length. For each run, stimuli were 80 full-face, full-color pictures of young males and females displaying 5 different emotional states –angry, fearful, happy, neutral, sad– which were selected from the MacBrain Face Stimulus Set (http://www.macbrain.org/faces/index.htm). Each face was presented for 2 s with an average inter-stimulus interval of 3 s. The order of presentation of the stimuli was determined according to an optimized random sequence (Optimize Design II[3]) which included null events and temporal jittering (0ms, 125ms, 250ms, 375ms, and 500ms) to increase statistical efficiency.[4] Null events included fixation crosses at the level of the eyes to direct the subjects' attention to this key

region of the face in light of evidence of a positive relationship between gaze fixation and activity in the fusiform gyrus and amygdala.[5] In the two separate runs, subjects were instructed to either imitate or just observe the faces they saw via a set of high-resolution, magnet-compatible goggles (Resonance Technology Inc.). Both imitation and observation runs consisted of 96 events, lasting 3 s each, with 16 stimuli for each emotion and 16 null events. The order of presentation of the Imitate and Observe runs was counterbalanced within each group.

Behavioral and Eye-Tracking Data

In a separate video-taped testing session, half the children in each group performed the exact same tasks while gaze fixation was monitored with an eye tracker (Tobii 1750; Tobii Technology). Subject demographics for these children are presented in **Supplementary Table 2**. Notably, there were no between-group difference between ASD and TD children in this smaller sample; similarly, for both the ASD and TD groups, there were no differences between children who participated in the fMRI session only and those who also took part in the subsequent behavioral session in any of these variables. Analyses of the gaze fixation data (total amount of time spent fixating on the eye region, and on the face in general, summed across fixations lasting more than 40 ms) revealed no between-group differences in the amount of time spent fixating on the eyes during both observation (ASD: M = 67 s; TD: M = 48 s) and imitation (ASD: M = 61 s; TD: M = 43 s) of emotional facial expressions, suggesting that the fixation crosses during the null events were successful in drawing the subjects' attention to this region. Similarly, there were no group differences in the amount of time spent fixating on the face during both observation (ASD: M = 115 s; TD: M = 125 s) and imitation (ASD: M = 118 s; TD: M = 120 s) of facial expressions.

Possible between-group differences in imitative behavior were also examined. First, fourteen observers, all blind to diagnosis but familiar with autism symptomatology, watched the videotaped imitation sessions of all 10 children and classified each child as likely to be autistic or typically

developing, based on their imitation of emotional expressions. These data indicated that children with ASD could not reliably be distinguished from TD children, $\chi^2(1) = 2.9$, $p > .05$, by this sample of observers many of whom had considerable experience with typically developing and/or autistic children. More specifically, ASD and TD children were equally likely to be classified as autistic or typically developing, with very little consensus among the observers' judgments for any given child. Indeed, only one ASD child was correctly identified as such by the majority of the observers (13/14); somewhat surprisingly, a TD child was also identified as potentially autistic by the majority of the observers (13/14). Second, for each child, two additional observers — also blind to diagnosis — were asked to identify which emotion was being portrayed after the presentation of each stimulus face, as well as to rate how well the emotional expressions were portrayed by each child on a 5-point Likert scale (with 1 = very poorly, 3 = average, and 5 = very well). As expected, accuracy ratings were rather high and comparable across groups (TD: M = 82%, range: 72-93%,; ASD: M = 85%, range: 75-94%), with excellent inter-rater reliability (Cohen's Kappa = .81 and .72 for TD and ASD groups, respectively). Further, no between-group differences were observed in the mean ratings of how well children imitated the emotional expressions (TD: M = 4.6, ASD: M = 4.2, p=.28). Taken together with the lack of any between-group differences in primary motor regions, these behavioral findings then suggest that that any observed between-group difference in neural activity in mirror neuron regions do not reflect overt differences in imitative behavior.

fMRI Data Acquisition

Images were acquired on a Siemens Allegra 3 Tesla head-only scanner. For each subject, two sets of 96 functional images were collected over 36 axial slices covering the whole cerebral volume using an echo planar T2*-weighted gradient echo sequence (TR = 3000 ms; TE = 25 ms; flip angle = 90 degrees; matrix size = 64 x 64; FOV = 20 cm; in-plane resolution = 3 mm x 3 mm; slice

thickness = 3 mm/1 mm gap). A high-resolution EPI structural volume was also acquired coplanar with the functional images (TR = 5000 ms, TE = 33 ms, 128 x 128 matrix size, FOV = 20 cm).

fMRI Data Analysis

Using Automated Image Registration (AIR),[6, 7] all functional images were a) realigned to correct for head motion and co-registered to their respective high-resolution structural images using a 6-parameter rigid body transformation model, b) spatially normalized into a Talairach-compatible MR atlas[8] using polynomial non-linear warping, and c) smoothed using a 6-mm FWHM isotropic Gaussian kernel.

Statistical analyses were implemented in SPM99 (Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/) and MarsBaR (http://marsbar.sourceforge.net/), a region of interest (ROI) toolbox for SPM99). For each subject, condition effects were estimated according to the general linear model, using a canonical hemodynamic response function, high-pass filtering, and no global scaling. The resulting contrast images were then entered into second-level analyses using a random effects model to allow for inferences to be made at the population level.[9] For each group (ASD and TD), separate one-sample t-tests were implemented for each condition (Imitate and Observe) relative to baseline. Two-sample t-tests were used to examine between-group differences in each condition; these analyses were performed within regions where reliable activity was detected in either group during the Imitate condition ($t > 1.83$, $k = 15$ voxels).

Regression analyses were also conducted in the ASD group using the subjects' scores on the social subscales of the ADOS-G and ADI-R to investigate the relationship between symptom severity in the social domain and MNS activity during the Imitate condition. IQ was entered as a covariate in these analyses which were performed within our regions of interest (i.e., pars opercularis, insula, and limbic structures), functionally defined as the network of areas where

reliably greater activity was observed in TD children than in ASD children ($t > 1.83$, $p <. 05$, corrected for multiple comparisons at the cluster level).

To avoid the possibility of thresholding out any low-level activity in the ASD group, the results were initially explored at $p < .05$, uncorrected, for both magnitude and spatial extent. However, all reported activity survived correction for multiple comparisons at the cluster level ($p < . 05$, corrected), as well as a more stringent statistical threshold for magnitude ($p < .01$, uncorrected), unless otherwise noted for selected contrasts in a priori regions of interest (e.g., amygdalae).

Additional analyses were performed to rule out the possibility that the observed between-group differences were due to reduced gaze fixation in the ASD group. Based on prior evidence of a strong positive correlation between gaze fixation and right fusiform gyrus (rFG) activity in ASD,[5] for each condition we examined whether rFG activity modulated activity in MNS regions in the ASD group (raw parameter estimates of rFG activity were extracted for each condition and then regressed on brain activation voxel-wise). The results of these whole-brain regression analyses showed that rFG activity did not correlate with activity in MNS regions in either the Observe or Imitate conditions ($p < . 05$, uncorrected, for both magnitude and spatial extent). Further, we examined whether rFG activity correlated with activity in the MNS regions in the right pars opercularis where strong inverse correlations were observed with the social scores of the ADI and ADOS. The results of these analyses confirmed that rFG activity did not correlate with MNS activity in these regions (ADI: r (8) = -.29, p = .41; ADOS: r (8) = .26, p = .47). Finally, in the ASD children for whom eye-tracking data were available, we also explored whether activity in these same MNS regions was related to the amount of time they spent fixating the eye region during the eye-tracking session. Again, there was no indication of a positive relationship between MNS activity in the right pars opercularis and time spent fixating on the eyes (ADI: r (4) = -.02, p = .97; ADOS: r (4) = -.24, p = .69).

Lastly, analyses were also conducted to address the effects of head movement during scanning. For each subject, mean head motion was computed using AIR by averaging the displacements across all voxels in all functional images relative to their mean position[10] during the Imitate and Observe scans. In both groups, mean head movement was overall greater in the Imitate condition (ASD: M = .91 mm; TD: M = .62 mm) than in the Observe condition (ASD: M = .58 mm; TD: M = .37 mm), though these differences were not reliable in either the ASD, t (18) = 1.30, p = .21, or the TD group, t (18) = 1.71, p = .11. Most importantly, there were also no reliable differences between the ASD and TD groups in the amount of head movement detected in either the Imitate, t (13*) = 1.20, p = .25, or Observe condition, t (18) = 1.27, p = .22. Despite the non-reliable differences, since mean head motion was somewhat higher in the ASD group, we rerun our between-group comparisons using an analysis of covariance with motion parameters entered as a covariate. The results of these analyses virtually mirrored those reported, confirming our main findings of greater MNS and limbic system activity in TD vs. ASD children. Finally, voxel-wise regressions analyses confirmed that head motion was not correlated with activity in any of our regions of interest, in either condition (p<. 05, uncorrected, for both magnitude and spatial extent).

*df differ because of unequal variance between groups

References

1. Lord, C. *et al. J Autism Dev Disord* 30, 205-23 (2000).
2. Lord, C., Rutter, M. & Le Couteur, A. *J Autism Dev Disord* 24, 659-85 (1994).
3. Wager, T. D. & Nichols, T. E. *Neuroimage* 18, 293-309 (2003).
4. Dale, A. M. *Hum Brain Mapp* 8, 109-14 (1999).
5. Dalton, K. M. *et al. Nat Neurosci* 8, 519-26 (2005).
6. Woods, R. P. *et al. J Comput Assisted Tomogr* 22, 139-52 (1998).
7. Woods, R. P. *et al. J Comput Assisted Tomogr* 22, 153-65 (1998).
8. Woods, R. P. *et al. Hum Brain Mapp* 8, 73-9 (1999).
9. Friston, K. J. *et al. Neuroimage* 10, 385-96 (1999).
10. Woods, R. P. *Neuroimage* 18, 769-88 (2003).