

Running head: NEURAL RESPONSE TO AWKWARDNESS IN ASD

A specific hypoactivation of right temporo-parietal junction/posterior superior temporal sulcus in response to socially awkward situations in autism

Peter C. Pantelis,^{1*} Lisa Byrge,¹ J. Michael Tyszka,² Ralph Adolphs,^{2,3} Daniel P. Kennedy^{1*}

¹Department of Psychological and Brain Sciences, Indiana University - Bloomington
1101 E. 10th Street, Bloomington, IN 47405

²Division of Humanities and Social Sciences, California Institute of Technology
1200 E. California Boulevard, Pasadena, CA 91125

³Computation and Neural Systems Program, California Institute of Technology
1200 E. California Boulevard, Pasadena, CA 91125

**Corresponding authors*, e-mail: pcpanTEL@indiana.edu, dpk@indiana.edu; phone: 812-856-7800

Abstract

People with autism spectrum disorders (ASD) often have difficulty comprehending social situations in the complex, dynamic contexts encountered in the real world. To study the social brain under conditions which approximate naturalistic situations, we measured brain activity with fMRI while participants watched a full-length episode of the sitcom *The Office*. Having quantified the degree of social awkwardness at each moment of the episode, as judged by an independent sample of controls, we found that both individuals with ASD and control participants showed reliable activation of several brain regions commonly associated with social perception and cognition (e.g., those comprising the “mentalizing network”) during the more awkward moments. However, individuals with ASD showed less activity than controls in a region near right temporo-parietal junction (RTPJ) extending into the posterior end of the right superior temporal sulcus (RSTS). Further analyses suggested that, despite the free-form nature of the experimental design, this group difference was specific to this RTPJ/RSTS area of the mentalizing network; other regions of interest showed similar activity across groups with respect to both location and magnitude. These findings add support to a body of evidence suggesting that RTPJ/RSTS plays a special role in social processes across modalities and may function atypically in individuals with ASD navigating the social world.

A specific hypoactivation of right temporo-parietal junction/posterior superior temporal sulcus in response to socially awkward situations in autism

Introduction

The largely effortless and automatic manner with which most people navigate the social world belies the underlying complexity of social interaction. Individuals with autism spectrum disorders (ASD) have notable difficulty in social situations (American Psychiatric Association, 2013). Yet, paradoxically, high-functioning adolescents and adults with ASD may sometimes exhibit no impairment in experimental tasks and assessments designed to test their social abilities in the laboratory, while nevertheless demonstrating the level of social impairment in everyday life necessary to have received such a clinical diagnosis (Klin et al., 2007).

A possible reason for this observed discrepancy is that laboratory tests may not faithfully reproduce several critical aspects of the naturalistic conditions under which social situations are encountered. First, laboratory tests are typically presented via only one isolated modality, versus via multiple modalities simultaneously. Second, laboratory stimuli are often presented in a static (e.g., text, pictures) as opposed to dynamic (e.g., video) form. Third, whereas typical experimental designs invite subjects to direct their attention to particular aspects of a stimulus (e.g., “How does person X feel?”), inference in the real world is often open-ended and without prompt. Fourth, the complexity of laboratory stimuli is typically reduced (by design), narrowing down for the subject the number of possible dimensions of a scene that may be salient even when there is no explicit task. For these reasons, to assess the psychology and neuroscience of social abilities in high-functioning individuals with ASD, it may be important to employ experimental tasks that better approximate real-world social scenarios—in all of their dynamism and complexity—for greater ecological validity.

We therefore aimed to present subjects with experimental stimuli that tap into social cognitive and perceptual processes and are comparatively naturalistic in manner of presentation

and richness of complexity. High-functioning adults with ASD and neurotypical (NT) controls watched a complete episode of a television sitcom (the American version of *The Office*) while undergoing functional magnetic resonance imaging (fMRI) scanning. *The Office* features uncomfortable and socially inappropriate interactions among its characters, and we selected this television program in particular for this socially awkward quality (in addition to several other reasons, described in *Methods*). Social awkwardness overlaps conceptually with the understanding of faux pas (Baron-Cohen et al., 1999) and violation of social norms (Berthoz et al., 2002), and is the kind of social cognitive process which Happé's (1994) "strange stories" are meant to tap. We assume that complex social abilities like perspective taking, mentalistic reasoning, and a rich understanding of what is socially appropriate are likely essential to the appreciation of these situations. We asked whether individuals with ASD exhibit neural differences at socially awkward moments, not because we are interested in the neural correlates of social awkwardness *per se*, but because we expect that a gamut of social cognitive and perceptual processes are especially likely to be recruited at these moments, and as they typically would be recruited in similar situations encountered in the real world.

Importantly, subjects had no explicit task to perform while watching *The Office* in the scanner (other than to watch and listen to the show and remain still), nor were experimental sessions divided into discrete blocks or trials across which an independent variable was systematically manipulated. We argue that this approach allows for a more naturalistic manner of stimulus presentation. Nevertheless, an experiment like this stands in stark contrast to another prevailing approach to stimulus design in psychophysics and neuroscience, which aims to systematically reduce the complexity of natural stimuli in order to gain better experimental control over the relevant perceptual variables defining a stimulus set (Johansson, 1973; Tremoulet and Feldman, 2000; Wilson et al., 2002; McAleer and Pollick, 2008). This reductionist approach allows for greater precision and confidence in the interpretation of resultant experimental data. Yet these advantages may trade off with ecological validity; for example, it can be difficult to systematically reduce the complexity of social situations to more basic components without disrupting the essential social quality of the stimuli or task, or making the independent variable

unnaturally salient to the participant. Employing more naturalistic social stimuli (Heavey et al., 2000; Moran et al., 2004; Dziobek et al., 2006; Hasson et al., 2009; Lahnakoski et al., 2012; Salmi et al., 2013) helps to bridge the more reductionist science to the ultimate object of generalization—the operation of the system *in situ*.

We are particularly interested in a constellation of brain regions that includes medial prefrontal cortex (MPFC), right and left temporo-parietal junctions (RTPJ and LTPJ), right superior temporal sulcus (RSTS) and temporal pole, and posterior medial cortices (posterior cingulate, precuneus). This “mentalizing network” is relevant to the study of ASD because it has been implicated in social perceptual and cognitive processes (Fletcher et al., 1995; Saxe and Kanwisher, 2003; Iacoboni et al., 2004; Saxe and Wexler, 2005; Schultz et al., 2005; Frith and Frith, 2006; Van Overwalle, 2009; Dufour et al., 2013; Takahashi et al., 2014, see bottom row of Fig. 2 for a visualization of brain areas found by the automated meta-analytic tool Neurosynth to be associated with social tasks across 800 previous studies, when queried with “social: reverse inference”, Yarkoni et al., 2011), which tend to be impaired in autism. Furthermore, abnormal activity in these brain regions has sometimes been found in autism during performance of social tasks (Castelli et al., 2002; Pelphrey et al., 2005; Wang et al., 2007; Kennedy and Courchesne, 2008; DiMartino et al., 2009; Lombardo et al., 2011; Murdaugh et al., 2012), while engaging in non-social tasks (Müller et al., 2001, 2003) and at rest in the absence of any explicit task instructions (Kennedy et al., 2006). When examining the brain’s response to socially awkward moments, and comparing the response of individuals with ASD with that of NT controls, several different patterns of results could be observed: 1.) no group differences (as in, for instance, Dufour et al., 2013), 2.) a substantially overlapping qualitative pattern of neural response, but with attenuated responses in one or more regions in the ASD group, 3.) a qualitatively different pattern of active brain regions in ASD, suggesting systematically atypical neural processes at socially awkward moments.

Methods

Subjects

Each subject's consent was obtained according to the Declaration of Helsinki (1991), and the institutional review board of the California Institute of Technology approved the fMRI study protocol. A total of 45 subjects participated in the fMRI experiment, including 20 individuals with ASD and 25 neurotypical (NT) controls. DSM-IV-R diagnoses of ASD were made by a clinical psychologist following administration of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2001) and, when a parent was available (i.e., in all but one case), either the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994) or the Social Communication Questionnaire (SCQ; Rutter et al., 2003).

Three subjects (1 ASD and 2 NT) were excluded prior to inspection of the data (1 ASD subject fell asleep during scanning; 1 NT subject quit in the middle of the experiment; 1 NT subject's behavior was noted as strange by the experimenters). After preprocessing of the fMRI data, 4 more subjects' data (2 ASD and 2 NT) were excluded based on high levels of a signal artifact metric associated with head motion (temporal **d**erivative of **v**ariance across voxels, or DVARS, Power et al., 2012); see *fMRI analysis* below for more details.

Our fMRI analyses thus included data from 38 total subjects (17 individuals with ASD and 21 NT controls). Demographic and behavioral/clinical data describing the two samples are provided in Table 1. There were no statistically significant group differences in gender breakdown, age, or IQ.

In addition, an independent sample of 46 neurotypical undergraduate subjects participated in a behavioral study (described below, *Quantifying Social Awkwardness*) for course credit (Mean age = 20.8, SD = 1.6; 24 males/22 females). The institutional review board of Indiana University approved this behavioral study.

Stimuli

Each subject watched the same episode of the television show *The Office* (Season 1, Episode 6, "Hot Girl"; Production Code 1003; Copyright NBC Universal), complete and unedited with the

exception of the opening title sequence being removed. This episode had a total running time of 20:52, and was divided into two halves of approximately equal length. This particular television program was chosen for several reasons. It conveys many socially awkward moments, utilizes limited cinematic features (e.g., zooming, panning, and camera cuts) that can direct viewers attention, is not accompanied by a laugh track, and features often subtle expressions of emotion compared to other shows. See *Supplementary Material* for a description of the television series and the specific episode.

Quantifying Social Awkwardness

An independent sample of 46 neurotypical undergraduate subjects watched the same complete episode in order to provide a moment-by-moment rating of its level of social awkwardness. These subjects were instructed to push a spring-loaded joystick (Pro Logitech Extreme 3D Pro; Model # 963290-0403) forward to the extent that they deemed what was happening within the scene to be awkward. A bar on the right side of the screen provided a real-time indicator of the degree to which they were moving the joystick, corresponding to a level between “not at all awkward” (baseline; resting state of joystick) to “extremely awkward” (pushed all the way forward). Subjects were instructed to continue pressing the joystick forward (to whatever extent they deemed appropriate) for as long as the scene continued to be awkward; thus, their moment-to-moment positioning of the joystick was meant to reflect their perceived degree of awkwardness of the video at any given time. These awkwardness ratings were sampled at the presentation of each frame of the episode (24 samples per second); Figure 1 shows each of the 46 subjects’ responses, along with the mean awkwardness calculated across all subjects. Because the repetition time (TR) of the fMRI scanner was 2.5 s, subjects’ mean awkwardness rating over each 2.5 s interval was calculated to be used as our regressor of interest.

Most of the fMRI subjects (12 of the subjects with ASD and 19 of the NT controls) rewatched the episode after the main experiment was complete, while still in the scanner. During the second half of the episode, these subjects pressed a button when they deemed something socially awkward to be happening. These raw ratings are presented as *Supplementary Material*.

Because we were unable to collect these data for many of the subjects, because this was a fairly coarse behavioral measure taken during a repeated viewing of the episode, and because subjects were explicitly instructed during this repeated viewing to rate the awkwardness, we do not place great emphasis on these awkwardness ratings as a faithful representation of how the fMRI subjects perceived the episode during their first, undirected viewing. Nevertheless, both the ASD and NT groups appear to show good correspondence with the ratings that were provided by the independent sample of 46 NT subjects, with no obvious group differences in the pattern of responses.

Imaging

All MRI data were acquired using a 3T Tim Trio system (Siemens Medical Solutions) with a 32-channel phased array head coil. High-resolution T1-weighted MP-RAGE volumetric datasets (1 mm isotropic voxel size; TR/TE/TI = 1500/2.91/800 ms) were acquired as anatomical references for each participant. Functional MRI image acquisition consisted of 2 sessions of approximately 10.5 minutes of T2*-weighted echo planar images (EPI) (TR/TE = 2500/30 ms, flip angle = 85°, 3 mm isotropic voxels and 47 slices acquired with ascending ordering, covering the whole brain). These 2 sessions corresponded to the first and second half of the episode (247 and 258 volumes; see Stimuli) and data from the two halves were collapsed together for the fMRI analyses reported here. The first two volumes (5 s) in each session were discarded prior to preprocessing to minimize magnetization equilibration effects. Gradient echo field mapping data were acquired with identical geometry to the EPI data for EPI off-resonance distortion correction (TR/TE = 500/2.5 ms, 5 ms, flip angle = 60°).

fMRI preprocessing and analysis

Preprocessing and most statistical analyses were performed using FEAT (v6.00) within FSL (the FMRIB Software Library, available at <http://fsl.fmrib.ox.ac.uk/>; Jenkinson et al., 2012). Preprocessing consisted of rigid-body motion correction, slice-timing correction, field map-based geometric distortion correction, nonbrain removal, temporal smoothing (100 s high pass filter cutoff), and spatial smoothing (5 mm full width at half maximum [FWHM] Gaussian kernel). BOLD EPI data was registered in two steps: affine registration to each subject's

T1-weighted anatomical image using Boundary-Based Registration (BBR; FMRIB's Linear Image Registration Tool, FLIRT; <http://fsl.fmrib.ox.ac.uk/flirt/>) and nonlinear registration of the T1-weighted anatomical image to the standard Montreal Neurological Institute (MNI152 T1-weighted; FMRIB's Non-linear Image Registration Tool, FNIRT; <http://fsl.fmrib.ox.ac.uk/fnirt/>).

For each subject, we calculated DVARS (temporal derivative of variance across voxels, Power et al., 2012) for each volume in the session as an indicator of gross artifact affecting the overall BOLD signal. Any volume with DVARS (i.e. root mean square intensity difference of volume N to volume $N + 1$) greater than .6% was effectively excluded from the analysis via a nuisance “spike” regressor in the GLM. There was no significant difference in the mean number of volumes removed for ASD ($M = 7.7, SD = 8.1$) versus NT control participants ($M = 8.0, SD = 12.4; t[36] = 0.10, p = 0.92$), nor was there was a significant group difference in mean DVARS for the remaining volumes included in the analyses (**ASD**: $M = 33.3, SD = 3.0$; **NT**: $M = 35.1, SD = 3.4; t[36] = 1.75, p = 0.09$).

A stability problem with 2 of the 32 elements of the head coil used for this study resulted in intermittent Nyquist ghost artifacts in many subjects' scans. We created an algorithm for detecting these artifacts within each horizontal slice at each TR (i.e. each volume) based on the measured signal outside of the brain. This detection algorithm yielded a continuous output for each of ~ 30 horizontal slices within each volume, generating a set of ~ 30 continuous nuisance regressors that were entered into the GLM for each subject. There was no significant group difference in the mean output yielded by the Nyquist ghost detection algorithm across all brain slices ($t[36] = 0.63, p = 0.53$), nor was there a significant group difference in the mean output yielded by this algorithm for each subject's worst slice ($t[36] = 1.18, p = 0.24$). Cerebrospinal fluid and white matter signals and their temporal derivatives were also included in the GLM as nuisance regressors, as were 6 head motion regressors ($x, y,$ and z translations and rotations) and their temporal derivatives.

We convolved our social awkwardness measure (see *Quantifying Social Awkwardness*) with a gamma function (6 s mean lag, 3 s half width, 0 s phase), and included this convolved, continuous regressor (along with its temporal derivative) in the GLM as our variable of interest.

The GLM parameters were estimated with FILM (FMRIB's Improved Linear Modeling; part of FSL), a method which includes a pre-whitening step. The parameter estimates of GLMs fit to each subject were the inputs for higher-level whole brain analyses.

A recent study (Dufour et al., 2013) found typical levels of activity in ASD across several brain regions often connected to social cognition, using a verbal false belief task (presented in text form or auditorily). We modeled region of interest (ROI) analyses after theirs, using a Bayesian approach that, like theirs, allowed us to report evidence in favor of *no* group differences where the data supported such a claim. We used the ROIs they make available (at saxelab.mit.edu/hypothesis_spaces.zip; see *Supplemental Material* for additional information about these ROIs), defined with respect to their large and independent dataset. This allowed us to examine the same approximate brain regions thought to be important for social cognition, and explore whether our comparatively naturalistic stimuli elicited a different pattern of brain activation with respect to this network.

Results

Whole brain analysis: Main effects of social awkwardness

Using FSL's mixed effects algorithm (FMRIB's Local Analysis of Mixed Effects, 1+2), we thresholded voxels positively correlated with our measure of social awkwardness (at $Z > 2.33, p < .01$) and applied a cluster-level significance threshold of $p < 0.05$ (Worsley, 2001). Collapsed across groups, we observed significant clusters of activity in several regions of the brain predicted *a priori* to be important for social cognition—including RTPJ, RSTS, precuneus (PC), and MPFC—in addition to a number of other brain regions—including orbitofrontal cortices, occipital cortex, and bilateral occipital fusiform gyri (Figure 2; see *Supplementary Material* for a table of information about these clusters and local maxima within them). Excluding brain stem voxels, 11% of the brain was significantly correlated with the social awkwardness regressor. These clusters found to be significantly correlated with social awkwardness bore resemblance to brain areas found by the automated meta-analytic tool Neurosynth to be recruited in social tasks across 800 studies (see Fig. 2, bottom row; indexed with “social: reverse inference” at

www.neurosynth.org; Yarkoni et al., 2011).

Figure 2 (top row) shows the overlap between those clusters we found to be significantly correlated with our social awkwardness regressor via whole brain analysis and a predefined network of seven region-of-interest masks Dufour et al. (2013) derived empirically ($RTPJ \cup RSTS \cup LTPJ \cup PC \cup DMPFC \cup MMPFC \cup VMPFC$). We hypothesized that socially awkward moments would recruit this predefined brain network; however, given the uncontrolled nature of the stimulus we did not expect that these moments of this particular television program would recruit *only* these regions. Unsurprisingly, the neural response was not entirely specific to these ROIs: at this statistical threshold, 75% of the voxels that were correlated with socially awkward moments fell outside of the Dufour et al. (2013) predefined network. Some voxel clusters corresponded well with these ROIs, but extended outside of them, some voxels were not contained within the Dufour et al. ROIs, but nevertheless showed excellent correspondence to “social” regions culled from Neurosynth (e.g., orbitofrontal cortices), and some voxels belonged to regions entirely distinct from these regions of *a priori* interest (e.g., occipital fusiform gyri). Nevertheless, after accounting for volume (these ROIs comprised only 9% of total brain voxels), the Dufour et al. ROIs were correlated with social awkwardness at a rate far higher than the rest of the brain: 33% of voxels within the predefined Dufour et al. network were significantly correlated with social awkwardness with respect to the whole brain cluster analysis, compared to just 9% of voxels lying outside of this network (excluding brain stem).

In order to compare how well various masks discriminated the social awkwardness clusters—in a manner that accounts for both sensitivity and specificity, and theoretically is not dependent on the arbitrary thresholds used to generate the masks—we here employ d' . We estimate *hit rate* from the proportion of the total number of voxels significantly correlated with social awkwardness falling within the candidate mask. We estimate *false alarm rate* from the proportion of the total number of voxels *not* significantly correlated with social awkwardness falling within the same candidate mask. The following formula is then used to estimate d' : $norminv(\text{hit rate}) - norminv(\text{false alarm rate})$, where $norminv$ is the inverse cumulative distribution function of the standard normal. Using this metric, a binary mask of the Dufour et al. ROIs discriminated

voxels significantly correlated with social awkwardness ($d' = .83$) better than many other binary masks generated from Neurosynth queries related to social perception and cognition (“social”: $d' = .65$, “biological”: $d' = .59$, “tom”: $d' = .53$, “intentions”: $d' = .52$, “mentalizing”: $d' = .49$, “perspective”: $d' = .49$, “gaze”: $d' = .25$, “social cognition”: $d' = .23$, “eyes”: $d' = -.01$, “mirror neuron”: $d' = -.38$, “faces”: $d' = -.62$).

Whole brain analysis: Group differences

We also examined whether any brain region across the whole brain showed group differences in response to social awkwardness (voxel-level threshold, $p < .01$; cluster-level threshold, $p < .05$). We found one contiguous cluster of voxels (2.8 cm^3 in size) spanning across RTPJ and RSTS for which there was significantly greater correlation for the NT group compared to the ASD group (Fig. 3). The MNI coordinates of voxels of peak group difference within this cluster were ($52 -36 2$; $z = 3.62$), ($56 -44 18$; $z = 3.57$), and ($56 -42 20$; $z = 3.49$); its center of mass was ($57 -40 14$).

75% of this cluster’s voxels were found within a region we predefined as RTPJ, and 25% of the voxels were found within the posterior portion of predefined RSTS. However, many of the voxels of peak difference were located very near the border between these two regions—a border which is very approximate to begin with. We therefore cannot determine with confidence whether we observed group differences at two dissociable loci, or one locus at the border of these two areas.

As an exploratory analysis, we examined whether activity in this RTPJ/RSTS cluster was correlated with autistic symptom severity among the 14 subjects with ASD for whom we were able to obtain AQ scores. We did not observe a significant correlation ($r[12] = -.045, p = .88$); however, we acknowledge that this small sample underpowers this analysis for the reliable detection of any but the strongest effects.

Across the whole brain, we found no significant group differences in the opposite (ASD > NT) direction.

ROI analysis

Given the high degree of overlap between regions we expected *a priori* to be particularly active at socially awkward moments, the regions that were indeed correlated with this social awkwardness regressor, and a network of several ROIs identified independently by Dufour et al. (2013) in a recent fMRI study of autism, we focused a more detailed analysis on this particular network, and limit the following analysis to this subset of brain regions.

For each subject, we extracted the peak t-stat voxel within each of the seven respective ROIs defined by Dufour et al. (2013), each representing the voxel most reliably correlated with the subject's observation of social awkward moments. For each ROI, we then tested whether there was a group difference (ASD vs. NT) in the location of this peak voxel. We calculated the Euclidean distance between the two group mean peak locations (in 3-dimensional space), and performed a Monte Carlo approximation (10^6 samples) of a non-parametric permutation test to derive the probability of observing a group mean difference of this size or larger given random rearrangement of group labels (Table 2). Only when examining the RTPJ region of interest did we observe even a marginally significant group difference ($p = .06$).

We then performed full Bayesian analyses (Gallistel, 2009, using Matlab functions provided at <http://cognitivegenetic.rutgers.edu/ptn/> and uniform priors across the predefined spatial limits of each particular ROI) to weigh the evidence for (or against) a group difference in peak voxel location in the x , y , or z dimensions, respectively. For most ROIs, we found moderate evidence in favor of *no* group difference in all spatial dimensions (Table 2). The strongest evidence we found for a group difference in location was along the y dimension in RTPJ, where there was moderate to strong evidence that ASD peak voxels within this ROI were comparatively more posterior than NT peak voxels.

For each subject and for each ROI, we then defined a sphere of voxels within a 5 mm radius of this peak location. This sphere was allowed to extend outside of the predesignated ROI, but not outside of the predefined network (i.e., $\text{RTPJ} \cup \text{LTPJ} \cup \text{RSTS} \cup \text{PC} \cup \text{DMPFC} \cup \text{MMPFC} \cup \text{VMPFC}$). We derived AICc-adjusted likelihood ratios (Burnham and Anderson, 2002, 2004) to again weigh the evidence for (or against) a group difference (ASD vs. NT), this time with respect

to the mean t-value within the sphere defined around the individual's peak voxel.¹ In Table 3, we also provide conventional (two-tailed) p-values to complement these less commonly-used statistics.

For most of the predefined ROIs (LTPJ, PC, DMPFC, MMPFC, and VMPFC), we found weak to moderate evidence in favor of the null hypothesis (i.e., no difference between group means). The exceptions were RTPJ and RSTS, for which we found some weak evidence of a group difference. These trends are consistent with our initial findings of a group difference in a (corrected) cluster straddling across these two predefined regions, and no significant group differences elsewhere. For this analysis, we initially assumed 5 mm radius spheres surrounding peak voxels in each ROI, but as illustrated in Figure 4, the choice of assumption (i.e., radius of sphere) does not affect the pattern of results.

Functional connectivity analysis

Our initial whole brain analysis of the fMRI data (the analysis described above) with respect to our social awkwardness regressor yielded evidence of a group difference in an area spanning RTPJ and RSTS (see Fig. 3). Having isolated this region, we searched for other regions with potentially high functional connectivity to the RTPJ/RSTS cluster using a conventional whole-brain temporal correlation analysis. To achieve this, we reran the full GLM including all nuisance regressors, but replaced the convolved social awkwardness regressor and its temporal derivative with the mean activity of voxels within the RTPJ/RSTS cluster shown in Figure 3. This regressor was generated by masking each subject's time series with the supra-threshold cluster region from the group level contrast and taking the unweighted mean across these voxels for each TR.

Collapsed across groups, a whole brain analysis (voxel-level threshold $p < .01$; cluster-level threshold $p < .05$) revealed clusters of activity significantly correlated with this seed region in each of the ROIs that were earlier found to have been correlated with observation of socially awkward moments—RTPJ, RSTS, precuneus (PC), and MPFC. But additionally, in this case we observed strongly correlated activity in LTPJ and LSTS, mirroring the activity of corresponding regions in the right hemisphere (Fig. 5).

We did not observe group differences in functional connectivity within any of the regions for

which we had strong *a priori* interest (i.e. the Dufour et al. ROIs). However, a 3.4 cm³ cluster of voxels, with local maxima in left orbitofrontal cortex ($z = 3.86$, MNI: -36 20 -14) and left temporal pole ($z = 4.68$, MNI: -26 14 -36), showed greater co-activity with RTPJ/RSTS in the NT group than the ASD group. In the opposite direction (ASD > NT), a 5.4 cm³ cluster of primarily white matter voxels in the left hemisphere showed greater co-activity with RTPJ/RSTS (peak: $z = 4.09$, MNI: -10 -12 36), likely reflecting a false positive. Neither of these clusters of group difference overlapped with any of our predefined ROIs (RTPJ, RSTS, LTPJ, PC, DMPFC, MMPFC, or VMPFC). Though we report these results for completeness, we view these results as tentative and preliminary, especially since we had no *a priori* hypotheses about them.

Discussion & Conclusions

For both individuals with ASD and their neurotypical (NT) counterparts, several brain regions previously associated with social cognition and perception tracked the socially awkward moments of an episode of *The Office*, and for the most part, to similar extents. However, we also found quantitative differences between the ASD and NT groups, most notably near the right temporo-parietal junction (RTPJ), spanning into the posterior end of the right superior temporal sulcus (RSTS). Using a Bayesian approach, we demonstrated that this observed group difference was likely specific to this particular region; analyses of several other regions of interest (ROIs) instead provided some evidence in favor of the null hypothesis (i.e., no difference between groups). Furthermore, we found no group differences in functional connectivity between this RTPJ/RSTS region and these other ROIs, suggesting that the abnormal activity in this region was likely not the result of extensive disruption of the functional connectivity of this network.

The cluster we isolated (MNI coordinates of center of mass: 57 -40 14; peak voxel: 52 -36 2), for which we observed a significant group difference (NT > ASD), was at the border of predefined RTPJ and RSTS ROIs. Although this area of the brain has been implicated in many social processes, the functional dissociability of the two regions remains a subject of debate in the literature. Very similar loci of activity have been functionally isolated across a wide variety of social tasks, some of which could be characterized as tapping lower-level social perceptual

processes, some of which might be described as tapping into a more cognitive “theory of mind,” and some of which are difficult to characterize with the perceptual-cognitive dichotomy.

Though Han et al. (2013) found this region to be sensitive to point-light animations of humans or animals in motion (MNI: 55 -41 16), others have argued that it is not the intrinsic motion of the object *per se* to which this region is sensitive, but how both the object’s motion and its interaction with its environmental context reveals its underlying intentionality (Saxe et al., 2004; Lee et al., 2014, MNI: 54 -42 9 and MNI: 56 -54 16, respectively). Similarly, although Villarreal et al. (2008) found this area to be sensitive to both meaningful and seemingly meaningless gestures, Pelphrey et al. (2004) and Vander Wyk et al. (2009) argued that this region was especially sensitive to the emotional or intentional context of grasping behavior (MNI: 57 -47 11 and MNI: 59 -47 -1, respectively). Takahashi et al. (2014) had NT adults engage in a game with various types of agents, and a principal component of the agent space associated with intelligence or rationality modulated a cluster of voxels at RTPJ/RSTS (MNI: 58 -46 0). In all of these cases, this region at RTPJ/RSTS was sensitive to agent behavior that implied an underlying intentionality—whether it be gaze, grasp, locomotion, or face-to-face interaction.

Lahnakoski et al. (2012) characterized this region (MNI: 58 -44 14) as being a social “hub” of sorts, because it responded to many different types of social features presented audiovisually to subjects (e.g., faces, bodies, and speech), to similar extents. Even if one only *hears* a person walking, this region is apparently evoked (Bidet-Caulet et al., 2005, MNI: 58 -47 15). Thus, the precise role of this area of cortex in social processing has not been defined. Even narrowing this role down along the perceptual-cognitive spectrum has proven difficult, because if a region is reliably active during a particular social perceptual task, this does not necessarily imply that this region is *for* social perception and *not for* higher-level “theory of mind.” As Saxe and Kanwisher (2003) argue, showing someone videos or pictures of people (or animals, mouths, hands, or bodies) invites the participant to think about their minds, implicitly tapping mentalizing processes. In other words, once one *perceives* another agent, it is natural to then start *thinking* about the agent’s mental states. Our study, with its naturalistic, but uncontrolled stimuli and open-ended experimental “task,” is unable to resolve these debates.

We instead note that this RTPJ/RSTS region, for which we find a group difference, bears very strong resemblance to an area that has been associated with social tasks in a number of other studies of neurotypical adults. Perhaps the most comparable empirical result with neurotypical adults was reported by Paulus et al. (2014), who found a similar region (MNI: 51 -49 7/66 -43 13) to be selective to scenes (in this case, static sketches drawn in black and white) in which a character unintentionally violates social norms, creating an awkward and embarrassing social situation. But it is clear that a better understanding of the computational functions each brain region contributes to the appreciation of socially awkward moments will be necessary to determine what is systematically different (and the same) in ASD in response to these moments. For example, is the reduced activation observed in RTPJ/RSTS correlative or causal in these individuals' social cognitive impairment? In other words, is RTPJ/RSTS dysfunctional in ASD, causing social impairment? Or might individuals with ASD exhibit atypical gaze or attentional patterns at socially awkward moments, indirectly resulting in less engagement of this region?

In individuals with ASD, Castelli et al. (2002) and Lombardo et al. (2011), like us, found attenuated activity at or near RTPJ (MNI: [52 -46 24] and [52 -62 28], respectively)—and in the latter study, only there. These two experiments presented subjects with tasks in quite different modalities. Castelli et al. (2002) used dynamic visual stimuli (in the style of Heider and Simmel, 1944, and not unlike those in Lee et al., 2014) and asked subjects open-ended questions. Lombardo et al. (2011), on the other hand, used verbal stimuli, asking subjects to make judgments about the mental ongoings of others. Therefore, an atypical response in autism of RTPJ/RSTS may transcend the visual presentation modality, tracking higher-level social features. That we would find a similar pattern of results with our dynamic stimulus presented both aurally and visually adds support to a body of evidence suggesting that RTPJ/RSTS plays a special role in social processes across modalities and may function atypically in individuals with ASD navigating the social world.

This discussion of group differences in RTPJ/RSTS should not obscure another prevalent and noteworthy pattern of results: evidence for a *lack* of group difference in other regions of the mentalizing network during socially awkward moments. These null findings partially replicate those of Dufour et al. (2013), who found no ASD versus NT group differences in seven ROIs

associated with mentalizing using a well-established false belief (versus false photo) task. Although RTPJ/RSTS may not be engaged to the same extent during socially awkward moments, in general, the brains of individuals with ASD respond similarly to those of NT controls. This interpretation is further supported by the finding that the functional connectivity of the mentalizing network seems to be largely intact; though we did find subtle group differences in two regions located outside of mentalizing ROIs, it does not appear to be the functional anatomy or coupling of this network that is grossly atypical, but the specific engagement of specific areas at particular times.

Given the presumed ecological validity of the stimulus presentation, we suggest that these experimental findings likely reflect how the brain of someone with high-functioning ASD would process socially complex and awkward situations as encountered in the real world. However, we also readily acknowledge that much of the benefit of a semi-naturalistic approach also creates challenges. Most notably, the stimuli are uncontrolled along a great many random but potentially meaningful variables, and it is not possible to control for all of them, making it difficult to discern which features of the stimuli are critical drivers of the neural response. Therefore, the use of semi-naturalistic stimuli cannot replace the use of more controlled stimuli, but should continue to complement or validate this approach (Hasson and Honey, 2012). When experiments employing naturalistic stimuli are embedded into a larger literature containing more controlled experiments, it is difficult to deny cases in which reorienting neuroscience toward ecological relevance has resulted in genuine insight and scientific progress (Rieke et al., 1995; Geisler et al., 2001; Graziano et al., 2002; Quiroga et al., 2005). In the present study, complex and dynamic social stimuli served as powerful tools, allowing for the observation of subtle and highly specific neural differences in ASD.

Funding

This work was supported by the National Institute of Mental Health at the National Institutes of Health (R00-MH094409 to D.P.K.) and the Brain & Behavior Research Foundation (previously NARSAD) Young Investigator Award (to D.P.K.).

References

- American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders. American Psychiatric Publishing, Arlington, VA, 5th edition.
- Baron Cohen S, O’Riordan M, Stone V, Jones R, Plaisted K. 1999. Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *J Autism Dev Disord.* 29:407–418.
- Berthoz S, Armony JL, Blair RJR, Dolan RJ. 2002. An fmri study of intentional and unintentional (embarrassing) violations of social norms. *Brain.* 125:1696–1708.
- Bidet Caulet A, Voisin J, Bertrand O, Fonlupt P. 2005. Listening to a walking human activates the temporal biological motion area. *NeuroImage.* 28:132–139.
- Burnham KP, Anderson DR. 2002. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach.* Springer, New York.
- Burnham KP, Anderson DR. 2004. Multimodel inference: Understanding aic and bic in model selection. *Sociol Method Res.* 33:261–304.
- Castelli F, Frith C, Happé F, Frith U. 2002. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain.* 125:1839–1849.
- Declaration of Helsinki. 1991. *BMJ.* 302:1194.
- DiMartino A, Ross K, Uddin LQ, Sklar AB, Castellanos FX, Milham MP. 2009. Functional brain correlates of social and nonsocial processes in autism spectrum disorders: An activation likelihood estimation meta-analysis. *Biol Psychiat.* 65:63–74.
- Dufour N, Redcay E, Young L, Mavros P, Moran JM, Triantafyllou C, Gabrieli JDE, Saxe R. 2013. Similar brain activations during false belief tasks in a large sample of adults with and without autism. *PLoS One.* 8:1–13.

- Dziobek I, Fleck S, Kalbe E, Rogers K, Hassenstab J, Brand M, Kessler J, Woike JK, Wolf OT, Convit A. 2006. Introducing MASC: A movie for the assessment of social cognition. *J Autism Dev Disord.* 36:623–636.
- Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RSJ, Frith CD. 1995. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition.* 57.
- Frith CD, Frith U. 2006. The neural basis of mentalizing. *Neuron.* 50:531–534.
- Gallistel CR. 2009. The importance of proving the null. *Psychological Review.* 116:439–453.
- Geisler WS, Perry JS, Super BJ, Gallogly DP. 2001. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Res.* 41:711–724.
- Graziano MSA, Taylor CSR, Moore T. 2002. Complex movements evoked by microstimulation of precentral cortex. *Neuron.* 34:841–851.
- Han Z, Bi Y, Chen J, Chen Q, He Y, Caramazza A. 2013. Distinct regions of right temporal cortex are associated with biological and human-agent motion: Functional magnetic resonance imaging and neuropsychological evidence. *J Neurosci.* 33:15442–15453.
- Happé F. 1994. An advanced test of theory of mind: Understanding of story characters’ thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J Autism Dev Disord.* 24:129–154.
- Hasson U, Avidan G, Gelbard H, Vallines I, Harel M, Minshew N, Behrmann M. 2009. Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Res.* 2:220–231.
- Hasson U, Honey CJ. 2012. Future trends in Neuroimaging: Neural processes as expressed within real-life contexts. *NeuroImage.* 62:1272–1278.
- Heavey L, Phillips W, Baron Cohen S, Rutter M. 2000. The awkward moments test: A naturalistic measure of social understanding in autism. *J Autism Dev Disord.* 30:225–236.

- Heider F, Simmel M. 1944. An experimental study of apparent behavior. *Am J Psychol.* 57:242–259.
- Iacoboni M, Lieberman MD, Knowlton BJ, Molnar Szakacs I, Moritz M, Throop CJ, Fiske AP. 2004. Watching social interactions produces dorsomedial prefrontal and medial parietal bold fmri signal increases compared to a resting baseline. *NeuroImage.* 21:1167–173.
- Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. 2012. Fsl. *Neuroimage.* 62:782–790.
- Johansson G. 1973. Visual perception of biological motion and a model for its analysis. *Percept Psychophys.* 14:201–211.
- Kennedy DP, Courchesne E. 2008. Functional abnormalities of the default network during self- and other-reflection in autism. *Soc Cogn Affect Neur.* 3:177–190.
- Kennedy DP, Redcay E, Courchesne E. 2006. Failing to deactivate: Resting functional abnormalities in autism. *P Natl Acad Sci USA.* 103:8275–8280.
- Klin A, Saulnier CA, Sparrow SS, Cicchetti DV, Volkmar FR, Lord C. 2007. Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: The Vineland and the ADOS. *J Autism Dev Disord.* 37:748–759.
- Lahnakoski JM, Glerean E, Salmi J, Jääskeläinen IP, Sams M, Hari R, Nummenmaa L. 2012. Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front Hum Neurosci.* 6:1–14.
- Lee SM, Gao T, McCarthy G. 2014. Attributing intentions to random motion engages the posterior superior temporal sulcus. *Soc Cogn Affect Neur.* 9:81–87.
- Lombardo MV, Chakrabarti B, Bullmore ET, MRC AIMS Consortium, Baron Cohen S. 2011. Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *NeuroImage.* 56:1832–1838.

- Lord C, Rutter M, DiLavore P, Risi S. 2001. Autism Diagnostic Observation Schedule (ADOS) Manual. Western Psychological Services, Los Angeles, CA.
- Lord C, Rutter M, Le Contour A. 1994. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord.* 24:659–685.
- McAleer P, Pollick FE. 2008. Understanding intention from minimal displays of human activity. *Behav Res Methods.* 40:830–839.
- Moran JM, Wig GS, Adams Jr. RB, Janata P, Kelley WM. 2004. Neural correlates of humor detection and appreciation. *NeuroImage.* 21:1055–1060.
- Müller R, Kleinhans N, Kemmotsu N, Pierce K, Courchesne E. 2003. Abnormal variability and distribution of functional maps in autism: An fMRI study of visuomotor learning. *Am J Psychiat.* 160:1847–1862.
- Müller R, Pierce K, Ambrose JB, Allen G, Courchesne E. 2001. Atypical patterns of cerebral motor activation in autism: A functional magnetic resonance study. *Biol Psychiat.* 49:665–676.
- Murdaugh DL, Shinkareva SV, Deshpande HR, Wang J, Pennick MR, Kana RK. 2012. Differential deactivation during mentalizing and classification of autism based on default mode network connectivity. *PLoS ONE.* 7:1–11.
- Paulus FM, Müller-Pinzler L, Jansen A, Gazzola V, Krach S. 2014. Mentalizing and the role of the posterior superior temporal sulcus in sharing others' embarrassment. *Cerebral Cortex.* .
- Pelphrey KA, Morris JP, McCarthy G. 2004. Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J Cogn Neurosci.* 16:1706–1716.
- Pelphrey KA, Morris JP, McCarthy G. 2005. Neural basis of eye gaze processing deficits in autism. *Brain.* 128:1038–1048.

- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*. 59:2142–2154.
- Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. 2005. Invariant visual representation by single neurons in the human brain. *Nature*. 435:1102–1107.
- Rieke F, Bodnar DA, Bialek W. 1995. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *P Roy Soc Lond B Bio*. 262:259–265.
- Rutter M, Bailey A, Lord C. 2003. *Social Communication Questionnaire (SCQ)*. Western Psychological Services, Los Angeles, CA.
- Salmi J, Roine U, Glerean E, Lahnakoski J, Nieminen-von Wendt T, Tani P, Leppämäki S, Nummenmaa L, Jääskeläinen IP, Carlson S, Rintahaka P, Sams M. 2013. The brains of high functioning autistic individuals do not synchronize with those of others. *Neuroimage Clin*. 3:489–497.
- Saxe R, Kanwisher N. 2003. People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*. 19:1835–1842.
- Saxe R, Wexler A. 2005. Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*. 43:1391–1399.
- Saxe R, Xiao DK, Kovacs G, Perrett DI, Kanwisher N. 2004. A region of right posterior temporal sulcus responds to observed intentional actions. *Neuropsychologia*. 42:1435–1446.
- Schultz J, Friston KJ, O’Doherty J, Wolpert DM, Frith CD. 2005. Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron*. 45:625–635.
- Takahashi H, Terada K, Morita T, Suzuki S, Haji T, Kozima H, Yoshikawa M, Matsumoto Y, Omori T, Asada M, Naito E. 2014. Different impressions of other agents obtained through social interaction uniquely modulate dorsal and ventral pathway activities in the social human brain. *Cortex*. .

- Tremoulet PD, Feldman J. 2000. Perception of animacy from the motion of a single object. *Perception*. 29:943–951.
- Van Overwalle F. 2009. Social cognition and the brain: A meta-analysis. *Human Brain Mapping*. 30:829–858.
- Vander Wyk BC, Hudac CM, Carter EJ, Sobel DM, Pelphrey KA. 2009. Action understanding in the superior temporal sulcus region. *Psychol Sci*. 20:771–777.
- Villarreal M, Fridman EA, Amengual A, Falasco G, Gerscovich ER, Ulloa ER, Leiguarda RC. 2008. The neural substrate of gesture recognition. *Neuropsychologia*. 46:2371–2382.
- Wang AT, Lee SS, Sigman M, Dapretto M. 2007. Reading affect in the face and voice: Neural correlates of interpreting communicative intent in children and adolescents with autism spectrum disorders. *Arch Gen Psychiat*. 64:698–708.
- Wilson HR, Loffler G, Wilkinson F. 2002. Synthetic faces, face cubes, and the geometry of face space. *Vision Res*. 42:2909–2923.
- Worsley KJ. 2001. Statistical analysis of activation images. *In* P Jezzard, P. M Matthews, and S. M Smith (eds.), *Functional MRI: An Introduction to Methods*, chapter 14. Oxford University Press.
- Yarkoni T, Poldrack R, Nichols T, Van Essen D, Wager T. 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 8:665–670.

Footnotes

¹We intend for these statistics to be interpreted similarly as one would interpret Bayes Factors, though the two approaches are not strictly the same. Here we choose to use AICc-adjusted likelihood ratios instead of a full Bayesian analysis to avoid the problem of defining an *a priori* hypothesis space in a situation where the size of the possible effect is unconstrained.

Table 1

Descriptions of the autism spectrum disorder (ASD) and neurotypical (NT) samples. For age, IQ, and AQ, group means are provided with standard deviations in parentheses. The AQ was only administered to 14 ASD and 17 NT subjects. For the three ADOS subscores, means are shown with the range (min–max) in parentheses. The ADOS was only administered to the ASD subjects.

	<i>n</i>	M/F	Age	IQ	AQ	ADOS _{comm.}	ADOS _{soc.}	ADOS _{behav.}
ASD	17	12/5	28.9(8.9)	111(12)	28.4(6.7)	4.3(2–7)	8.6(4–14)	1.6(0–4)
NT	21	18/3	26.9(5.5)	112(9)	13.8(4.9)	-	-	-

Table 2

In the first row, we report the Euclidean distance between the group means of locations of peak voxels (in 3D space), in each predefined region of interest. The displayed p-value (probability of observing a group mean difference of this size or greater under the null hypothesis) is derived from a 10^6 sample Monte Carlo approximation of a non-parametric permutation test. Bayes Factors (BF) are calculated separately for each spatial dimension. BFs > 1 indicate evidence for the null hypothesis, BFs < 1 indicate evidence for a group difference, and BF=1 represents no evidence in either direction.

	Region of Interest						
	RTPJ	RSTS	LTPJ	PC	DMPFC	MMPFC	VMPFC
Distance (mm)	8.6	11.1	1.9	7.6	3.1	3.5	2.1
<i>p-value</i>	.06	.22	.12	.95	.68	.66	.86
BF-X	1.4	3.8	0.26	5.3	3.8	4.2	3.9
BF-Y	0.11	1.7	5.7	4.7	3.8	2.6	4.0
BF-Z	4.1	3.9	2.1	5.2	4.6	3.8	3.8

Table 3

Mean *t*-statistic within specified spherical regions (of 5 mm radius) centered in predefined ROIs (standard deviation in parentheses). The *p*-value displayed is from a two-tailed, two-sample *t*-test; a Monte Carlo approximation of a non-parametric permutation test yielded identical *p*-values, to at least the two decimal places displayed. In order to approximate a Bayes Factor, we calculated the ratio of the maximum likelihoods of the data under a null model (under which the data are sampled from one Gaussian distribution) and an alternative model (under which the data are sampled from two Gaussian distributions with different means but the same variance), and adjusted for the more complex model via the Akaike Information Criterion (corrected for finite sample sizes, AICc). AICc-adjusted likelihood ratios > 1 indicate evidence for the null hypothesis, AICc adj. LR < 1 indicate evidence for a group difference, and AICc adj. LR=1 represents no evidence in either direction.

	Region of Interest						
	RTPJ	RSTS	LTPJ	PC	DMPFC	MMPFC	VMPFC
Neurotypical mean (s.d.)	2.58(.76)	3.14(1.14)	1.85(1.02)	2.54(.69)	2.86(1.35)	3.08(1.08)	1.81(.99)
Autism mean (s.d.)	2.17(.73)	2.53(1.12)	1.75(.77)	2.13(1.02)	2.53(.93)	2.79(1.32)	2.15(.80)
Effect size (Cohen's <i>d</i>)	0.55	0.54	0.11	0.48	0.29	0.24	0.38
<i>p</i>-value	.10	.11	.75	.14	.38	.46	.26
AICc adj. LR	0.81	0.82	3.09	1.03	2.16	2.46	1.69

Figure Captions

Figure 1. Top: Raw data for social awkwardness rating task. Each row represents an individual subject. Subjects' ratings are represented on a gradient from white (not awkward) to black (maximally awkward). *Bottom:* The continuous regressor of interest for fMRI analysis: A smoothed running average of subjects' mean awkwardness rating. The "halfway" point, at which the episode was divided, is shown with a dashed line. **Note:** A measure of variance is intentionally absent in this bottom plot, since the mean signal reflects the actual regressor used in the fMRI analysis. A sense of variance at each point can be obtained by viewing the raster plot above.

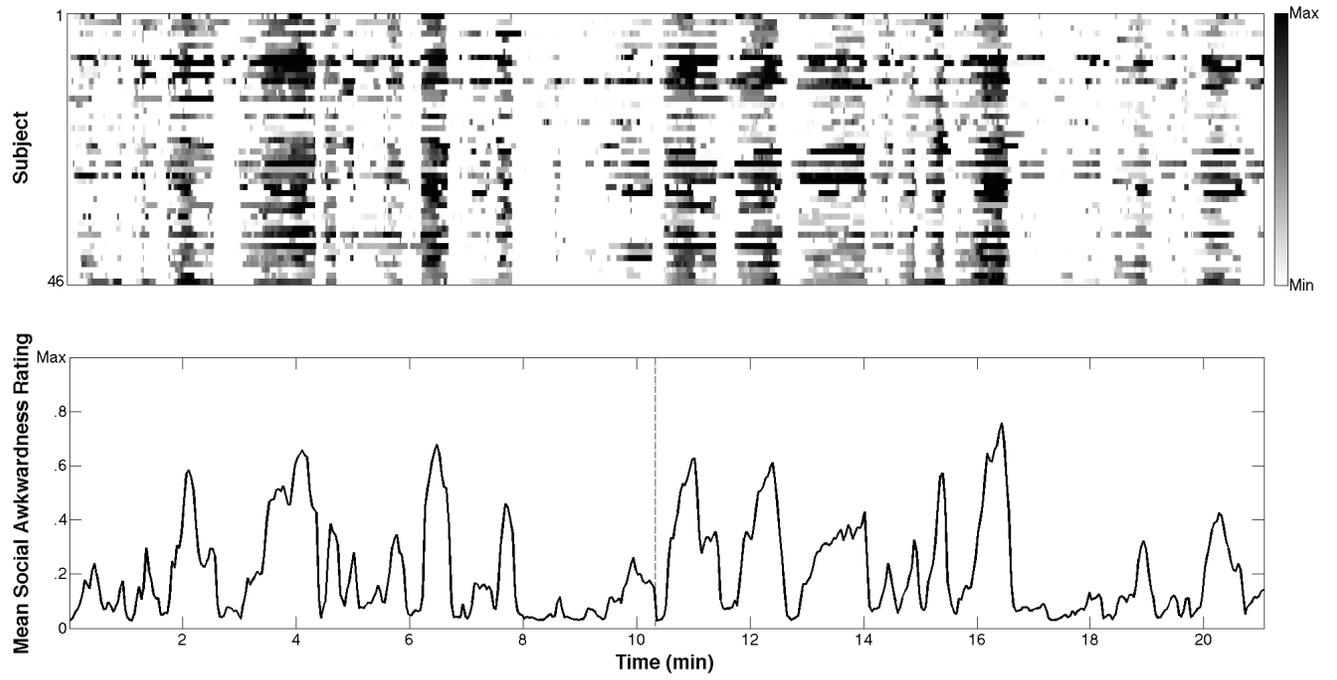
Figure 2. Top row: Voxels positively correlated with observation of socially awkward moments (red-yellow) across all subjects (after voxel-level threshold of $p < .01$ and cluster-level threshold of $p < .05$). Outline of "mentalizing network," as defined *a priori* by an independent study (Dufour et al., 2013), is shown in green. **Bottom row:** Areas (red) correlated with social tasks, via the meta-analytic tool Neurosynth. The Neurosynth "decoder" tool makes it possible to compare unthresholded maps of brain activation with statistical maps generated by the Neurosynth across a wide variety of keywords. Of the 92 possible keyword "features" made available for comparison, the "social" brain map correlated best with the pattern of activation observed here (i.e. correlation with socially awkward moments). *Left:* Lateral view ($x=52$) of RTPJ and RSTS regions. *Center:* Coronal view ($y=-58$), with RTPJ, PC, and LTPJ visible. *Right:* Medial view ($x=0$), with PC and MPFC visible. Aside from LTPJ, there was substantial overlap between these previously defined regions of interest and those which appear to have been active with respect to our variable of interest.

Figure 3. The cluster of voxels for which we observed a significant group difference (Neurotypical > ASD), shown on a gradient from red to orange. These sagittal cross sections go from lateral (left) to medial (right): $x = 62, 58, 56, 50$. The cluster straddles predefined RTPJ and RSTS regions; outlined here in blue and yellow, respectively.

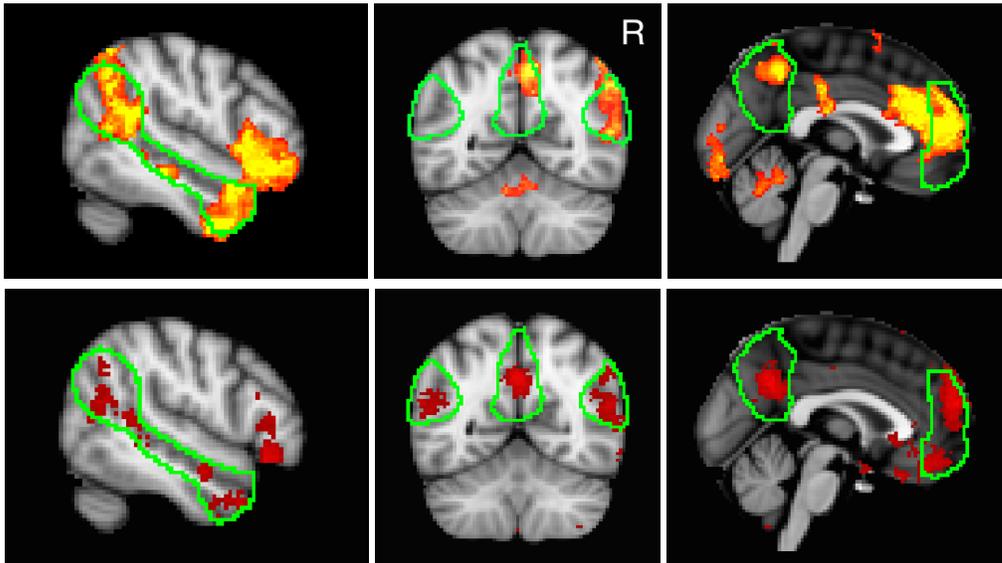
Figure 4. Weight of Evidence (\log_{10} odds), either in favor of a difference between group means (above dotted line), or for no group difference (below dotted line). For each subject, for each region of interest, a sphere of voxels was defined around the voxel most correlated with social awkwardness. The radius of this sphere could be made to vary; here, the weight of evidence is calculated under varying assumptions for the radius of this sphere, from 3 mm (smallest circles) to 9 mm (largest circles). The choice of assumptions does not seem to qualitatively affect the result of the analysis; for most regions, we find weak to moderate evidence *in favor* of the null hypothesis, with the exceptions of RTPJ and RSTS, for which we find weak evidence for a group difference.

Figure 5. Clusters co-active, across all subjects, with the region of RTPJ/RSTS for which a group difference (NT > ASD) had been previously observed in response to social awkwardness. Outline of “mentalizing network,” as defined *a priori* by an independent study (Dufour et al., 2013), is shown in green. *Left:* Lateral view ($x = -52$) of LTPJ and left temporal lobe. *Center:* Coronal view ($y = -56$). *Right:* Lateral view ($x = 52$) of RTPJ and right temporal lobe.

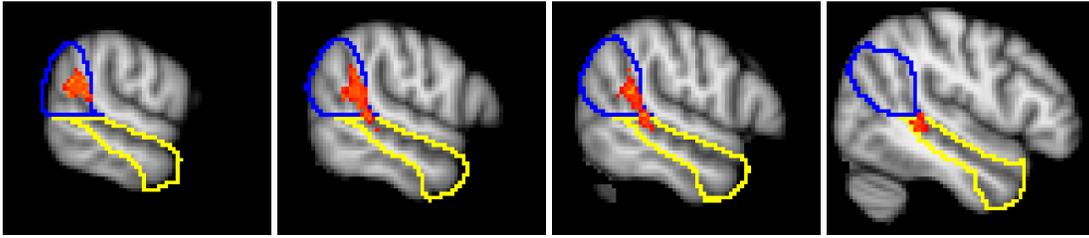
Neural response to awkwardness in ASD, Figure 1



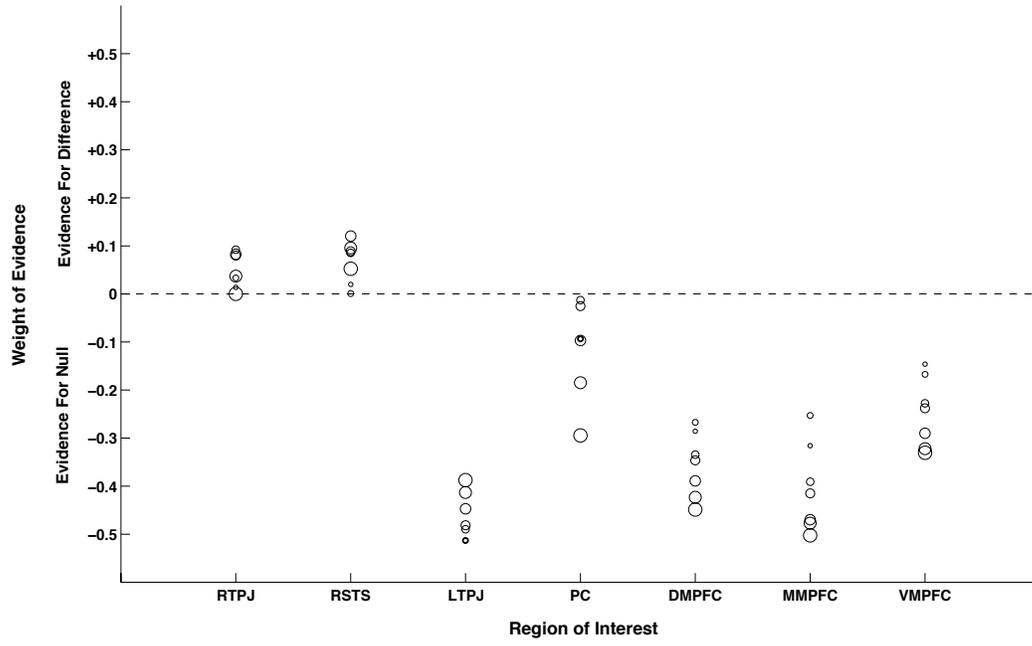
Neural response to awkwardness in ASD, Figure 2



Neural response to awkwardness in ASD, Figure 3



Neural response to awkwardness in ASD, Figure 4



Neural response to awkwardness in ASD, Figure 5

