NeuroImage 84 (2014) 648–656

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Revealing the neural networks associated with processing of natural social interaction and the related effects of actor-orientation and face-visibility

Manish Saggar *, Elizabeth Walter Shelly, Jean-Francois Lepage, Fumiko Hoeft, Allan L. Reiss

Center for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, Stanford University, USA

ARTICLE INFO

Article history:
Accepted 19 September 2013
Available online 29 September 2013

Keywords:
Social cognition
Event-related fMRI
Non-verbal communication
Actor-orientation
Face-visibility

ABSTRACT

Understanding the intentions and desires of those around us is vital for adapting to a dynamic social environment. In this paper, a novel event-related functional Magnetic Resonance Imaging (fMRI) paradigm with dynamic and natural stimuli (2 s video clips) was developed to directly examine the neural networks associated with processing of gestures with social intent as compared to nonsocial intent. When comparing social to nonsocial gestures, increased activation in both the mentalizing (or theory of mind) and amygdala networks was found. As a secondary aim, a factor of actor-orientation was included in the paradigm to examine how the neural mechanisms differ with respect to personal engagement during a social interaction versus passively observing an interaction. Activity in the lateral occipital cortex and precentral gyrus was found sensitive to actor-orientation during social interactions. Lastly, by manipulating face-visibility we tested whether facial information alone is the primary driver of neural activation differences observed between social and nonsocial gestures. We discovered that activity in the posterior superior temporal sulcus (pSTS) and fusiform gyrus (FFG) was partially driven by observing facial expressions during social gestures. Altogether, using multiple factors associated with processing of natural social interaction, we conceptually advance our understanding of how social stimuli is processed in the brain and discuss the application of this paradigm to clinical populations where atypical social cognition is manifested as a key symptom.

Introduction

Although Darwinian theory of evolution states that only the fittest will survive, cooperation within the same- and between-species is very common (Axelrod and Hamilton, 1981; Henrich et al., 2003). In particular, human societies outrank all other species based on our large-scale cooperation, even among genetically unrelated individuals (Fehr and Fischbacher, 2004; Nowak, 2012). The ability to efficiently communicate and interact is perhaps what makes humans such remarkable outliers relative to the rest of the animal kingdom. For a successful social interaction, we use a complex array of cues including facial expressions, gaze, gestures, speech, intonation and cadence to gauge the intentions of others around us. In comparison to the lower-level visual system, where the processing of stimuli is relatively linear and sequential (Gold et al., 2012; Shapley, 2009), processing of social stimuli is highly complex and non-linear, mainly due the inherent complexities of the stimuli itself. Thus, delineating the neural correlates of natural social interactions can be a daunting task. However, it is crucial that we understand how our brains process social cues, so that clinical populations where atypical social cognition is evident can be better understood and treated.

To understand and reveal the neural correlates of basic social interactions, this paper focuses on non-verbal communications, which can be used for obtaining and transmitting socially relevant information by performing different facial expressions, gaze movements, and hand gestures (Montgomery and Haxby, 2008). Several fMRI studies have investigated these communications in healthy adults (Conty and Grezes, 2012; Gallagher and Frith, 2004; Knutson et al., 2008; Lotze et al., 2006; Montgomery and Haxby, 2008; Montgomery et al., 2007; Morris et al., 2005; Schilbach et al., 2006; Villarreal et al., 2008). Although initial studies were limited to using static pictures of faces or bodies as stimuli (Montgomery and Haxby, 2008), several recent studies have explored new ways of presenting stimuli to better capture the neural correlates of realistic nonverbal communications. For example, researchers have started to appreciate the value of using short video clips to present dynamic and realistic stimuli to participants (Gallagher and Frith, 2004; Knutson et al., 2008; Lotze et al., 2006; Montgomery et al., 2007; Villarreal et al., 2008). Video clips are preferred, as compared to the static pictures, because the dynamic images tend to elicit higher activation in the brain regions responsible for processing and producing affective content (e.g., amygdala) as well as explicit body movements (e.g.,

* Corresponding author.
E-mail address: saggar@stanford.edu (M. Saggar).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.09.046

pSTS), thereby providing more information regarding social stimulus processing than static pictures (Grèzes et al., 2007). Other studies have used virtual immersion paradigms, e.g., Morris and colleagues contrasted trials where a virtual character was present with trials where just a framed picture was shown (Morris et al., 2005). Similarly, Schilbach and colleagues used virtual characters to differentiate between the neural networks associated with processing of one-on-one social interaction and processing of passively viewed social interaction between others (Schilbach et al., 2006). Using the virtual immersion paradigms, it is experimentally feasible to develop and study complex social interaction scenarios. However, it is unclear whether the neural networks engaged during a virtual interaction overlap (and to what degree) with the networks engaged during a *natural* interaction.

Given the complex nature of a social stimulus and its processing, several previous studies have attempted to fragment the stimuli into different categories and examine the associated neural networks in isolation. For example, researchers have examined the neural correlates of nonverbal interaction using just facial expressions (Montgomery and Haxby, 2008) or isolated hand movements (Villarreal et al., 2008). Further, recent studies have also attempted to differentiate and partition gestural stimuli into different categories, e.g., transitive versus intransitive (Villarreal et al., 2008), communicative versus object-based (Morris et al., 2005), or expressive versus body-referred (Lotze et al., 2006). Although these studies provide crucial information regarding how our brains process social cues in isolation, it is unclear what brain regions are engaged explicitly for processing cues involved in the "social" interaction gestalt. More importantly, because the social cognition system can be safely assumed to be more than just the sum of its parts, *fragmenting* social stimuli into different components and studying them in isolation could potentially obscure a more accurate and holistic understanding of this system.

To address some of the limitations in the extant literature and to investigate the neural basis of processing nonverbal interaction without finely fragmenting its components, we developed the Dynamic Social Gesture (DSG) task. This task uses dynamic and natural gesture stimuli (video clips of 2 s each) in an event related fMRI design. The DSG task is comprised of three factors (Fig. 1). First, sociability – short clips of actors performing interactive (e.g., "a friendly wave") or non-interactive gestures (e.g., "reaching for a cup"). Thus, in this paper, a gesture is deemed as social if it is *intended* to elicit a response from the participant and nonsocial otherwise. Very few studies to date have *directly* examined the neural networks associated with sociability (Mainieri et al., 2013; Morris et al., 2005). Of these studies, Morris and colleagues used a virtual immersion paradigm, where the social trials were defined as those in which a virtual character was present and the nonsocial trials were defined by the absence of the virtual character (and instead a framed picture of human face was shown). Furthermore, the virtual character always performed non-interactive gestures (e.g., scratching its face).

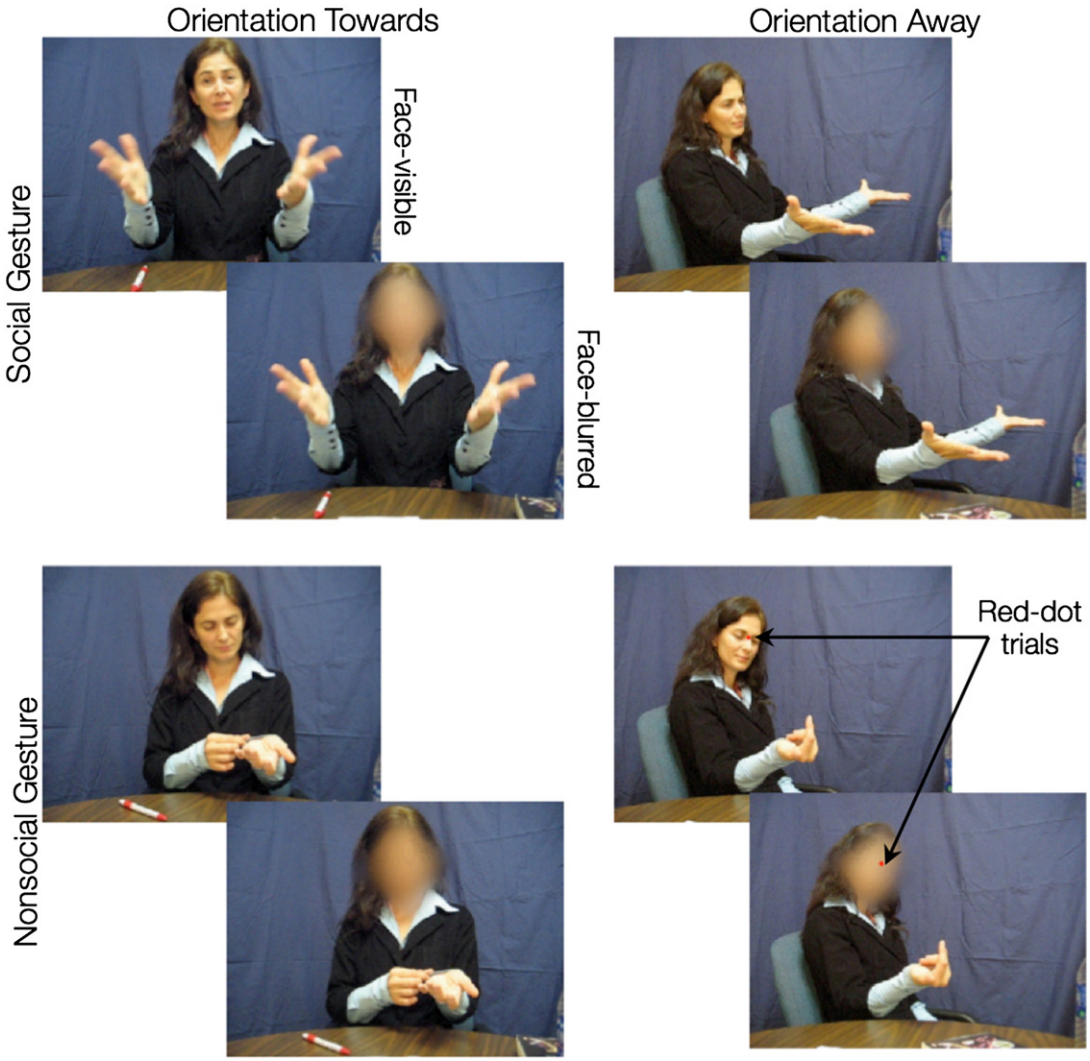

Fig. 1. DSG task contains three factors: sociability, actor-orientation, and face-visibility. Here we show sample snapshots from our dataset of 200 dynamic social gesture stimuli. Two gestures are shown here — implore (social) and looking at arms (nonsocial). To make sure that participants were attending to the stimuli, half of the gestures had a red dot that appeared on the actor's face (between the eyes) 1 s after the stimulus-onset. As a cover task, participants were instructed to press a button as soon as they detected the red dot.

In a more recent study, Mainieri and colleagues used hand gestures (with actor's face-covered) to examine the neural networks associated with sociability. Here, social trials were defined as gestures that were meaningful with high communicative intent (e.g., thumbs-up), whereas nonsocial trials were defined as meaningful gestures with low communicative intent (e.g., pointing towards left). In the current work, we used natural (as opposed to virtual) and integrated (as opposed to just hand gestures) stimuli to investigate the neural networks directly associated with processing of trials with social intent as compared to nonsocial intent.

The second factor built into the DSG task is actor's orientation, i.e., clips where actors perform an act towards the participant or away from the participant (i.e., towards a third person, who is not shown in the stimuli). The effects of another person's orientation during a social interaction are also not well understood. In particular, it is not clear how the neural mechanisms differ with respect to personal engagement during a social interaction versus passively observing another interaction. Gestures oriented towards the observer (as opposed to away from the observer) presumably also require higher social attention and gaze/facial expression processing (Nummenmaa and Calder, 2009). Further, investigating the brain areas responsible for processing personally engaging social interactions could shed light on our understanding of clinical populations that exhibit particular difficulty with such stimuli (e.g., fragile X syndrome (Garrett et al., 2004)). The previous study by Schilbach and colleagues used a virtual immersion paradigm to disentangle the effects of a virtual character's orientation during a social interaction. The virtual character was either oriented towards the participants (i.e. personally involved with participants) or oriented away (i.e. the participants were being passive observers) (Schilbach et al., 2006). In our work, we instead use natural stimuli to assess differences in activation patterns associated with actor's orientation during a dynamic social interaction.

The third factor built into the DSG task is that of face-visibility, i.e., clips where actors' faces are visible or blurred while they perform the gestures. In order to elucidate the neural correlates of natural social interactions, the actors were instructed to portray facial expressions associated with the performed gesture. Thus, the factor of face-visibility allowed us to test whether the facial information itself is the *primary driver* of neural activation differences between the social and nonsocial gestures.

In summary, we developed a novel fMRI paradigm that permits the study of multiple factors (sociability, orientation, and face-visibility) associated with processing of natural social interactions. Although in recent studies different components of social interaction have been brought together to better understand how interactions among these components facilitate social interaction processing (Conty et al., 2012; Ulloa et al., 2012), to our knowledge, no other study has directly investigated the neural networks associated with natural social versus nonsocial interaction processing and the related effects of actor-orientation and face-visibility.

Using the DSG task, we hypothesized that while comparing social to nonsocial gestures we would observe increased activation in the mentalizing (or theory of mind) network that is related to thinking about the internal states of others (Kennedy and Adolphs, 2012; Saxe, 2006a). We also hypothesized that our experimental design would help delineate the effects of an individual's orientation on gesture processing. Thus, comparing actor-orientation of towards versus away, we expected more activation in the visual attention areas, whereas in the reverse contrast we expected more activation in the dorsal visual stream related to detecting and analyzing components of movements (e.g., location, direction). Third, based on previous research focused on observing and imitating facial expressions (Carr et al., 2003; Pitcher et al., 2011), we predicted that the stimuli with dynamic face and gaze information would activate the fusiform area and the posterior STS more strongly than the blurred stimuli. Fourth, we predicted that the DSG task would also reveal brain areas where the factors of sociability, actor-orientation, and face-visibility interact while processing natural and dynamic social interactions.

Methods

Participants

Twenty healthy young adults (average age = 21.41, SD = 2.36, range = 16.9–25.7 years; 10 female) participated in this study after giving their written informed consent. All subjects were right-handed, had no contraindications for MRI scanning (e.g., metal implants or pacemakers), and had no self-reported history of past or current psychiatric or neurological condition. The university's research ethics board approved the experimental protocol and procedures.

Task and stimuli

The Dynamic Social Gesture (DSG) task (see Fig. 1) was comprised of a set of short (2 s) color video clips of a live actor either performing a social gesture ("friendly wave," "handshake," "beckoning," "joint attention," or "imploring") or a nonsocial gesture ("rubbing hand on table," "reaching for a cup," "brushing off a table," "looking at a book," and "looking at arms"). In half of the clips, the actor was oriented towards the participant (towards condition), and in the other half of the clips the actor was turned at an angle and oriented away from the participant, as if they were addressing an unseen individual just off camera (away condition). A total of five actors were used to generate movies in this study. Further, copies of all clips underwent smoothing of the entire face (using Elasty software; http://www.creaceed.com/elasty/mac), to create an additional blurred version of the stimulus. Thus our stimuli comprised of a 2 (sociability: social vs nonsocial) × 2 (actor-orientation: towards vs away) × 2 (face-visibility: visible vs blur) factorial design with 25 stimuli in each of these 8 experimental conditions and hence, a total of 200 different stimuli. The stimuli were not equated for the amount of motion or motion energy across conditions.

Stimuli were presented at the center of a viewing screen using a custom-built MR-compatible projection system and EPrime 2.0 software (http://www.pstnet.com/eprime.cfm). Participants lying in the scanner viewed stimuli via an angled mirror. The order and spacing between our stimuli were optimally randomized using optseq (http://surfer.nmr.mgh.harvard.edu/optseq/) in order to ensure orthogonality of our stimulus conditions. During the period between videos, the screen was all black for 500–9500 ms (based on optseq output).

Participants viewed all the movie clips during two runs (approximately 16 min in total) and, as an attentional cover task, were asked to press a button with their right index finger when they saw a red dot appear near the eyes and nose of the actor (Fig. 1). Red dots appeared on half of the gestures 1 s after the clip onset, and were equally likely to occur in every condition. This task was designed to be a simple control task, to ensure that participants were paying attention to the stimuli.

Before undergoing imaging, participants were familiarized with the task by playing a short (2 min) practice version of the task. Stimuli used for the practice version were distinct from those used in-scanner and comprised different actors and gestures. Participants were not told ahead of time about the gestures that they would be observing. They were told that their task was to detect the small dots that appeared near the eyes and nose area of the actors. Hence, the participants were naïve to the hypotheses of the study.

MRI acquisition

Participants were scanned on a 3 Tesla (GE Signa scanner, Milwaukee, WI) MRI at Stanford's Lucas Center for Neuroimaging using a custom-built single-channel birdcage head coil optimized for fMRI scans. Over two runs, a total of 469 whole-brain volumes were collected

on 30 axial-oblique slices (4.0 mm thick, 1.0 mm skip) prescribed parallel to the intercommissural (AC-PC) line, using a T2*-weighted gradient echo spiral pulse sequence (Glover and Lai, 1998) sensitive to blood oxygen level-dependence (BOLD) contrast with the following acquisition parameters: Echo Time (TE) = 30 ms, repetition time (TR) = 2000 ms, flip angle = 80°, FOV = 22 cm, acquisition matrix = 64 × 64, and approximate voxel size = 4.0 × 3.4 × 3.4 mm. To reduce blurring and signal loss arising from field in-homogeneities, an automated high-order shimming method based on spiral acquisitions was used before acquisition of functional MRI scans. A high-resolution T1-weighted three-dimensional inversion recovery spoiled gradient-recalled acquisition was acquired for co-registration with the following parameters: Echo Time (TE) = 6 s, repetition time (TR) = 35 s, flip angle = 45°, FOV = 24 cm, slice thickness = 1.5 mm, 124 slices in the coronal plane, matrix = 256 × 192, and acquired resolution = 0.94 × 1.25 × 1.5 mm. The images were reconstructed as a 256 × 256 × 124 matrix.

fMRI analysis

Functional MRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following pre-statistics processing was applied: motion correction using MCFLIRT, non-brain removal using BET, spatial smoothing using a Gaussian kernel of FWHM 5 mm, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 120.0 s). Additionally, sharp motion peaks were detected using *fsl_motion_outliers* script (supplied with FSL) and were regressed out in addition to the six motion parameters (from MCFLIRT). Registration to high-resolution structural and standard space images was carried out using FLIRT. Time-series statistical analysis was performed using FILM with local autocorrelation correction. Intra-subject individual runs were combined using a fixed effects model, by forcing the random effects variance to zero in FLAME (FMRIB's Local Analysis of Mixed Effects). Group-level analysis was carried out using threshold-free cluster enhancement (TFCE; height parameter = 2, extent parameter = 0.5) (Smith and Nichols, 2009) in combination with permutation testing (FSL Randomise with 5000 permutations) (Nichols and Holmes, 2002; Smith et al., 2004) to control for family-wise error (FWE) rate at $p < 0.05$. Featquery tool (supplied by FSL) was used to extract percent change in parameter estimates for anatomically defined (using Harvard–Oxford cortical and sub-cortical atlases) regions of interests. MRIcron (http://www.mricro.com/) was used to visualize neuroimaging results on the standard anatomical brain.

Results

Behavioral performance on the cover task

Participants achieved near-perfect accuracy in the cover task (M = 98.75%, SD = 1.9%). Response time in detecting dots was analyzed using a 3-factor repeated-measures ANOVA for sociability (2), actor-orientation (2), and face-visibility (2). Significant main-effects of sociability ($F(1,19) = 225.8$, $p < 0.001$) and actor-orientation ($F(1,19) = 194.9$, $p < 0.001$) were observed. Post-hoc analysis revealed that participants responded faster for social gestures as opposed to nonsocial ($t(19) = 15.02$, $p < 0.001$) and they were also faster in gestures with actor-orientation towards them, as opposed to away ($t(19) = 13.96$, $p < 0.001$) (Fig. 2A).

Further, we found two significant interactions, between: (a) sociability and actor-orientation ($F(1,19) = 378.35$, $p < 0.001$); and (b) actor-orientation and face-visibility ($F(1,19) = 7.34$, $p = 0.014$). These interactions were further explored using post-hoc Student's t-test analysis. The first interaction revealed that participants responded faster during the social gestures (as compared to the nonsocial gestures) when the acts were oriented towards them ($t(19) = 20.17$, $p < 0.001$) (Fig. 2B). The second interaction revealed that for the acts oriented-towards, participants responded faster when the actor's face was visible as opposed to when it was blurred ($t(19) = 3.88$, $p = 0.001$; Fig. 2C). Altogether, participants responded faster for detecting the red dot in conditions where either the gestures were social or the acts were oriented towards them.

Neuroimaging results

The fMRI data were analyzed to find the neural correlates for the three main effects (i.e., sociability, actor-orientation, and face-visibility) and four interactions between the main factors. The results are presented below.

Main effects

The main effect of sociability was examined using a 2-sample t-test to compare social with nonsocial gestures. In the subcortical regions, significant activations encompassed regions in the bilateral amygdala, hippocampus, and right parahippocampal gyrus. Bilateral cortical activations were observed in the angular gyrus of the inferior parietal lobule (IPL), lateral occipital cortex (LOC), cuneus, lingual, fusiform gyrus (FFG), and occipital pole, whereas right lateralized cortical activations were observed in the STS and medial temporal gyrus (MTG) (Fig. 3A and Table S1). The reverse contrast, however, comparing nonsocial with social gestures, failed to reveal any activation at the significance level.

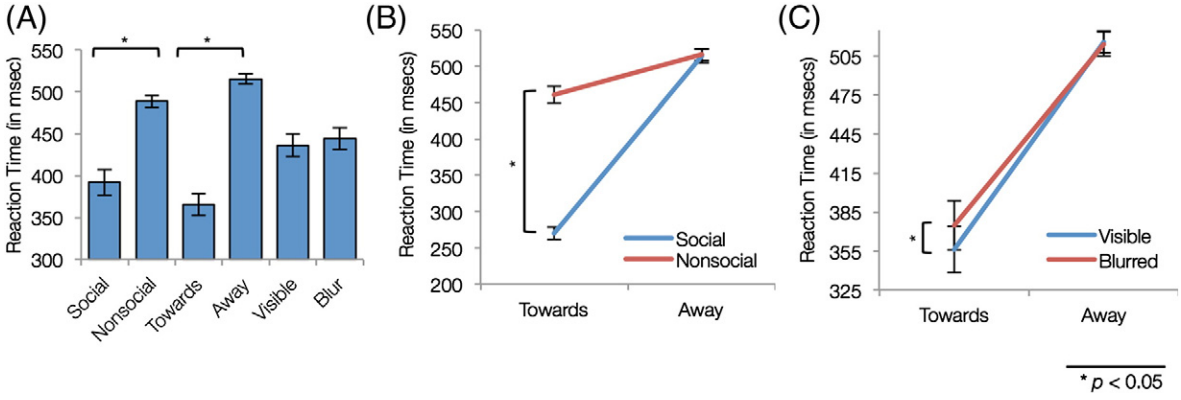

Fig. 2. Behavioral results: (A) Reaction time for each factor. Both social and oriented-towards gestures had lower reaction times. (B) Interaction between sociability and actor-orientation. Participants responded faster to social and oriented-towards gestures. (C) Interaction between face-visibility and actor-orientation. Within oriented-towards gestures, face-blurred gestures had higher reaction time than face-visible gestures.

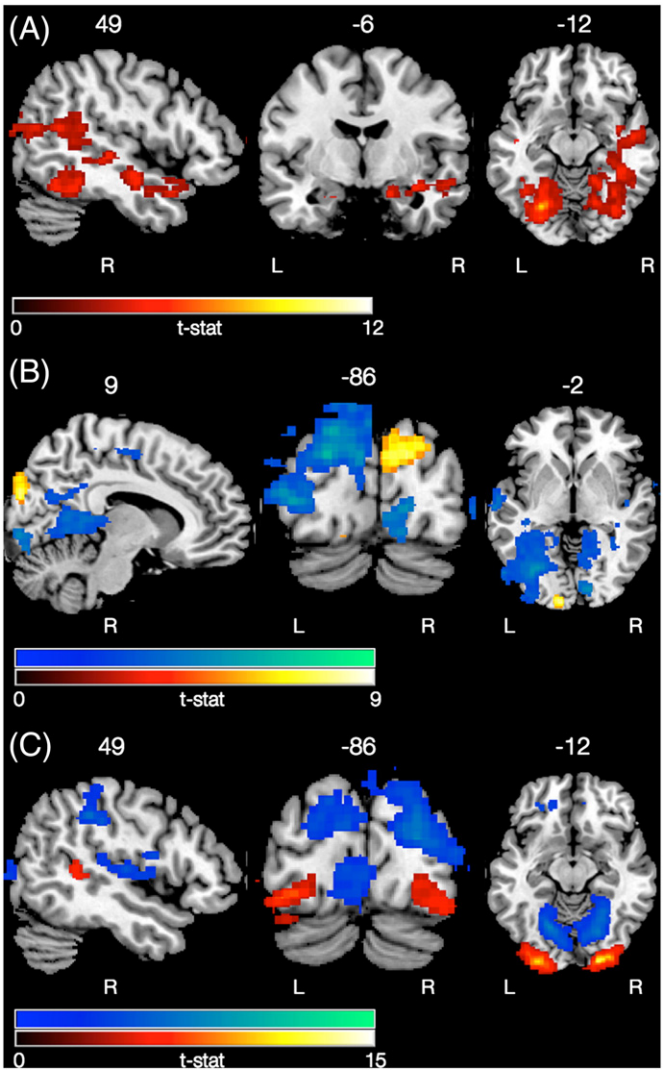

Fig. 3. Neural networks associated with the main effects of the task. (A) Effect of sociability, i.e., comparing all social gestures with nonsocial gestures (in hot color scale). None of the clusters reached significance in the reverse contrast. (B) Effect of actor-orientation, comparing all trials with oriented-towards with away gestures (in hot color scale). Reverse contrast is shown in the winter color scale. (C) Effect of face-visibility, i.e., comparing all face-visible with face-blurred gestures (in hot color scale). Reverse contrast is shown in the winter color scale.

The main effect of actor-orientation was also examined using a 2-sample t-test to compare gestures oriented towards with those oriented away from the participant. The results revealed activation in the bilateral occipital pole, where activation in the right occipital pole was more superior than the left. Further, activation in the left occipital pole also spread into the occipital fusiform gyrus (Fig. 3B and Table S2). In the reverse contrast, increased activity was revealed in the left lateralized dorsal occipital/parietal "where-pathway" (Ungerleider, 1994), parahippocampal gyrus, pre/post central gyrus, right lateralized insula, cingulate cortex, and bilateral central operculum (Fig. 3B and Table S3).

The main effect of face-visibility, i.e. comparing face-visible with face-blurred gestures, revealed activation in the bilateral occipital pole (also spreading into the occipital fusiform area) and right pSTS/MTG area (Fig. 3C and Table S4). The reverse contrast, revealed activations in the right lateralized dorsal occipital/parietal "where-pathway", insula, bilateral lingual, central operculum, and the anterior cingulate cortex (ACC) (Fig. 3C and Table S5).

Interactions

Out of the four possible interactions, only two were found to be significant. The first significant interaction was found between sociability (social or nonsocial) and actor-orientation (towards or away) in the bilateral lateral occipital cortex (LOC), left precuneus cortex, left precentral cortex, right intra-calcarine cortex, and right occipital pole (Fig. 4 and Table S6). This interaction was further explored by extracting beta-estimates from the significant clusters. Left and right LOC clusters revealed different interaction patterns. The left LOC cluster had similar activation levels for the social gestures oriented towards or away from the participant. However, for the nonsocial gestures, the left LOC cluster had higher activation for gestures oriented away as compared to gestures oriented towards participants. On the other hand, the right LOC cluster had similar activation levels for the nonsocial gestures oriented away or towards the participants. But, for social gesture the right LOC cluster had higher activation level for gestures oriented towards as compared to away from the participants. The right intra-calcarine cluster also differentiated between social gestures oriented towards versus away from the participants in a manner similar to the right LOC cluster. The interaction pattern in the right precentral cortex revealed higher activation for social gestures oriented towards versus away and higher activation for nonsocial gestures oriented away versus towards the participant.

The second significant interaction was found between sociability (social or nonsocial) and face-visibility (visible or blurred) in the left fusiform gyrus, such that social gestures with actor's face-visible elicited higher activation than social gestures with actor's face blurred. No such differentiation for face-visibility was found for the nonsocial gestures (Fig. 4 and Table S7).

Discussion

A novel event-related fMRI paradigm with dynamic and natural social gestural stimuli was developed to examine whether using natural stimuli as opposed to isolated face and hand movements would provide a more accurate and holistic understanding of how social interactions are processed in the brain. As hypothesized, while comparing social to nonsocial gestures we found increased activation in both the mentalizing (or theory of mind) and the amygdala networks. As a secondary aim, we examined how actor-orientation and face-visibility would affect social stimulus processing. For the effect of actor-orientation (collapsed across other factors), we found increased activation in the visual attention areas, while the reverse contrast showed more activation in the dorsal visual stream, perhaps related to detecting and analyzing movements in space. For the main effect of face-visibility, the presence of facial information alone was associated with increased activity in the pSTS, while the reverse contrast of face-blurred versus face-visible gestures revealed activation in the right dorsal visual stream, anterior cingulate, and central operculum. Further, two significant interactions were observed: sociability × actor-orientation and sociability × face-visibility. Both interactions were limited to the visual and primary motor areas, thereby suggesting that during a social interaction, orientation and face information are incorporated at the early stages of stimulus processing in the brain.

Behavioral performance during the cover task

Participants were faster for social and oriented-towards gestures. It is possible that during oriented-towards gestures, participants responded faster because of the location of the actor, which was more towards the center of the screen as compared to the oriented-away gestures. However, previous work on measuring the effect of head orientation on gaze processing (Pageler et al., 2003), using static pictures presented in the center of the screen, also showed that participants were behaviorally fastest in conditions when both actor's face and gaze direction were oriented towards them (similar to our oriented-towards condition).

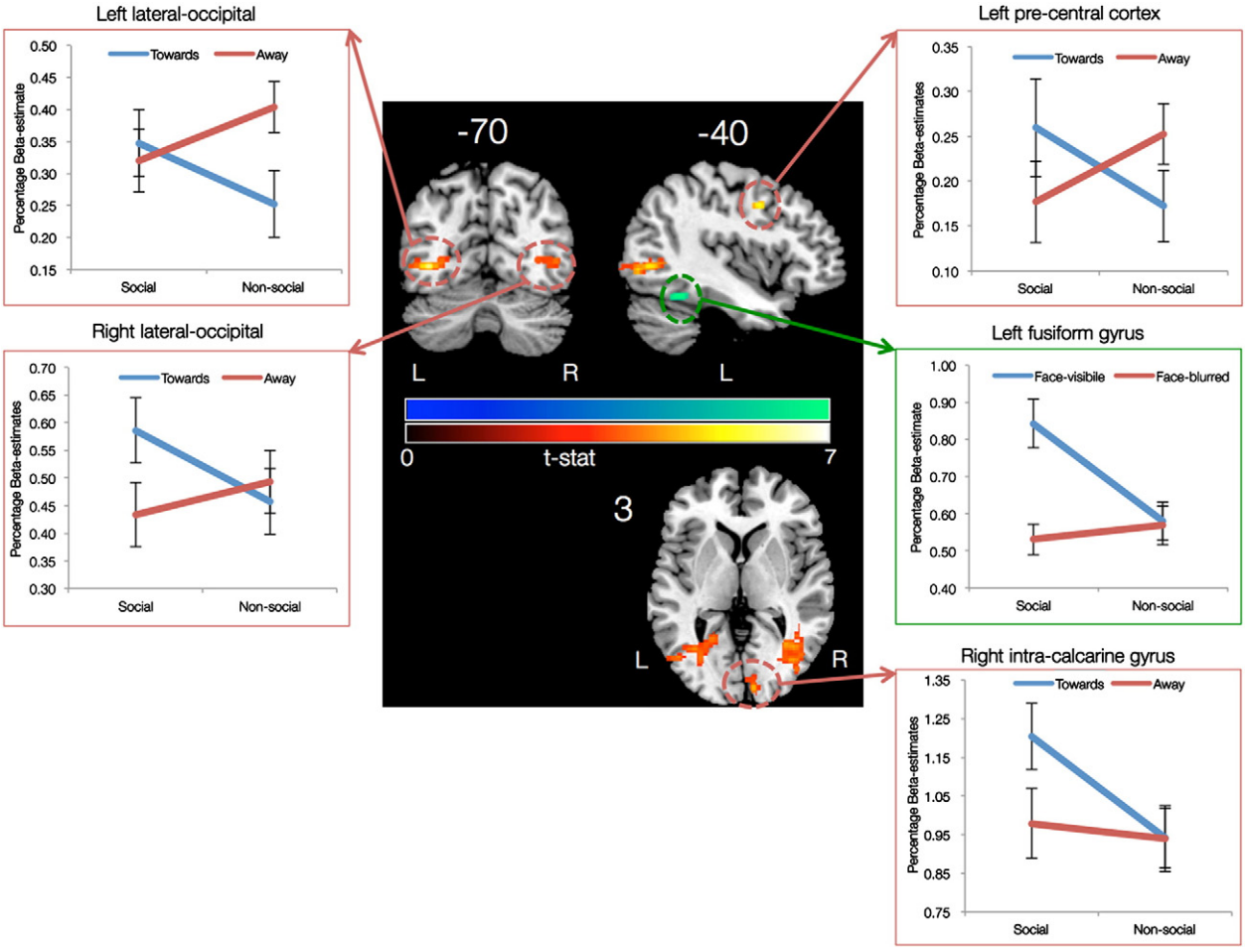

Fig. 4. Interaction between the main factors of the DSG task. The first interaction was observed between sociability and actor-orientation (in hot color scale) in the regions of left and right lateral occipital cortex, right pre-central gyrus, and right intra-calcarine gyrus. The second interaction was observed between sociability and face-visibility (in winter color scale) in the left fusiform gyrus.

Additionally, within the oriented-towards gestures, participants were faster during face-visible as opposed to face-blurred gestures. Taking these findings into account, it could be argued that facilitation in reaction time for act-oriented towards gestures in this study could be, in part, due to faster processing of socially salient information relative to nonsocially relevant stimuli.

The neural networks associated with natural social gesture processing

In the current work, natural stimuli were used to investigate the neural networks directly associated with processing of trials with social intent as compared to nonsocial intent. Very few previous studies have directly examined such neural networks (Mainieri et al., 2013; Morris et al., 2005). Morris and colleagues used a virtual immersion paradigm and contrasted trials where a virtual character was present (termed as "social" trials) with trials where the character was absent but other objects were shown (termed as "nonsocial" trials). It was observed that the presence of a virtual character engages processing bilaterally in the lateral occipital/temporal areas and FFG. Further, any socially neutral gestures (e.g., scratching its face) performed by the character led to specific engagement of the right pSTS and precentral gyrus.

In a more recent study, Mainieri and colleagues used hand movements (with actor's face covered) to contrast social (defined as meaningful and highly communicative gestures) with nonsocial (meaningful, but less communicative gestures) trials. The experimental paradigm also included an imitation phase, where the participants were asked to imitate the observed gestures. The results of this study showed increased engagement of medial/superior frontal cortex and cingulate gyrus during social as compared to nonsocial trials (Mainieri et al., 2013). In addition to the STS, medial prefrontal cortex is also considered a part of the mentalizing network (Kennedy and Adolphs, 2012). In the current work we did not observe medial frontal activation while comparing social to nonsocial trials (collapsed across other conditions of actor-orientation and face-visibility). However, as a post-hoc exploratory analysis, comparing social with nonsocial gestures *specifically* in the trials where the actors' faces were visible and their orientation was towards the observer (i.e., social-towards-visible > nonsocial-towards-visible) we did find activations in the bilateral ventromedial and dorsomedial prefrontal cortex, left frontal-orbital cortex, left cingulate gyrus (anterior division) and left inferior frontal gyrus (pars triangularis) (Fig. S1).

In lieu of a direct comparison, other studies have indirectly investigated the neural correlates of sociability (Knutson et al., 2008; Lotze et al., 2006; Montgomery and Haxby, 2008). For example, Montgomery and colleagues examined the differences in patterns of brain activation for processing static pictures portraying emotional facial expressions versus isolated but social hand gestures (Montgomery and Haxby, 2008). Observing social hand gestures elicited increased activation bilaterally in the STS and IPL, whereas observing just facial expressions elicited increased activation in the frontal operculum.

Altogether, using natural stimuli, our current work builds upon previous results and goes further to suggest the engagement of a wider network (that includes STS, IPL, FFG, LOC, amygdala and hippocampus) during the processing of natural social gestures as opposed to nonsocial ones. Further, given that aberrant activation in these regions, especially

in the STS, has been linked with atypical social cognition (e.g., autism (Boddaert et al., 2004), fragile X syndrome (Garrett et al., 2004), Williams syndrome (Golarai et al., 2010), Turner syndrome (Molko et al., 2004)), we anticipate that paradigms that use natural stimuli could be used in the future to better predict and/or detect early onset of such abnormalities and perhaps as a biomarker to follow response to intervention.

Influence of actor-orientation on the neural networks associated with social gesture processing

A factor of actor-orientation was included in the DSG task to examine the differences in neural mechanisms associated with interactions oriented-towards the participant (i.e., personally engaging) versus interactions oriented-away from the participant (i.e., passively observing the actor interact with a third person). Contrasting oriented-towards with away trials resulted in increased activation bilaterally in the occipital pole, suggesting increased visual processing when the acts were oriented-towards the participants. Response time data on the cover task (of detecting red dots) also showed that participants were faster in responding during social and oriented-towards trials. Overall, these findings suggest that during oriented-towards trials, participants were more engaged and attentive so as to infer the actor's goals and intentions.

We also found an interaction between sociability and actor-orientation in the bilateral LOC, left precentral cortex, and right intracalcarine/occipital pole areas. The extracted beta estimates from these regions suggest engagement of visual areas specifically in the right hemisphere for social trials that were oriented towards the participants. Previous work has shown that the lateral occipital complex (especially extra-striate body area) is sensitive to the orientation of the stimuli (i.e., allocentric versus egocentric) (Chan et al., 2004; Saxe et al., 2006) and takes into account the social meaning of actions (Jeannerod, 2004). Further, the precentral cortex has also been shown to be involved in the processing of social interaction, perhaps due to its proximity to the "mirror regions" (Morris et al., 2005). Given these results, in the future, our DSG task could be employed to shed light on our understanding of clinical populations that exhibit particular difficulty with socially engaging stimuli (e.g., fragile X syndrome (Garrett et al., 2004)).

The reverse contrast of oriented away versus towards trials, activated the left dorsal visual stream or the "where" pathway. This result is in line with the idea that participants were trying to discern the direction or orientation of the actor's movements in space, or to whom the action was intended (not shown in the stimuli). In addition, we also found activation in the right insular region. During the oriented-away trials, a component of social rejection could have come into play, as the actor's behavior displayed in the clip did not include social interaction with the participant. Previous work provides a basis for this conjecture in that brain regions involved in affective processing (especially posterior cingulate and insula) are activated during social rejection (Kross et al., 2007; Zucker et al., 2011).

The only other study that directly examined the influence of actor's orientation on perceiving social interactions was done using a virtual immersion paradigm by Schilbach et al. (2006). In this study, the participant either was gazed at by a virtual character (similar to our orientation-towards condition) or the participant observed the virtual character looking at someone else (orientation-away condition). Further, these virtual characters either portrayed socially relevant facial expressions or arbitrary facial movements. When contrasting oriented-towards versus away trials Schilbach et al. observed increased activation in the ventromedial prefrontal cortex (VMPFC), where as the reverse contrast revealed activation in the left dorsal visual stream and precuneus. Using natural stimuli, we did not observe engagement of VMPFC during orientation-towards trials. Further, we observed that the interaction between sociability and actor-orientation was also limited to the visual processing areas. One plausible explanation for the differences in activation pattern could be that in the Schilbach et al.'s study participants were shown a close-up view of the virtual character's face and the character portrayed social facial expressions during social orientation-towards trials (as compared to a half-body view of the actor with facial expressions and hand movements in our DSG task). Several previous studies using fMRI have shown that the VMPFC is implicated in processing of emotional empathy during interactions (Hynes et al., 2006; Saxe, 2006b; Völlm et al., 2006). Thus, it is possible that the natural social stimuli in our DSG task could not have differentially engaged empathy in the orientation towards versus away trials. Future work is required in this area to dissociate the effects of emotional empathy and social attention orienting. It is also important to note that although virtual immersion paradigms provide a powerful tool in neuroscience to systematically study different constructs, the assurance that virtual characters actually evoke the same experiences as the natural human stimuli is an essential prerequisite.

Is face-processing the primary driver for the neural networks associated with social gesture processing?

To keep the stimuli natural and realistic, the actors in the DSG task were instructed to portray facial expressions associated with the performed social gesture. Thus, to examine whether facial information itself is the primary driver for the neural mechanisms engaged during social gesture processing, a face-visibility factor was included in the DSG task. Contrasting face-visible with face-blurred gestures revealed activation bilaterally in the occipital poles (spreading into the FFG) and right pSTS/MTG area. Previous work using static pictures as well as dynamic stimuli has shown increased activation in the right pSTS when an actor's face was visible (irrespective of other body parts) (Morris et al., 2006; Pitcher et al., 2011).

We also found an interaction between sociability and face-visibility in the left FFG such that when the actor's face was visible, more activation was found in FFG for social trials as compared to nonsocial trials. This finding suggests that the FFG encoded facial expressions associated with social gestures. Previous studies in clinical populations with atypical social functioning have shown that the FFG plays a crucial role in social stimulus processing (Fung et al., 2012). For example, in Williams syndrome, where elevated social approach is manifested, there is an abnormally greater response of the FFG to facial expressions (Golarai et al., 2010), whereas in fragile X syndrome, where increased social avoidance is manifested, relatively reduced response of the FFG to facial expressions is observed (Garrett et al., 2004). Our findings in a healthy population using dynamic and natural stimuli are in line with previous research, suggesting a role for fusiform gyrus in detecting salient social stimuli in addition to processing of facial information. Altogether, building upon previous work, our results suggest that the activation in the pSTS and FFG during social versus nonsocial comparison could be partly associated with facial expression processing.

In the reverse contrast-face-blurred versus face-visible gestures, we found activations in the right lateralized dorsal visual stream, anterior cingulate, bilateral lingual gyrus and central operculum. Activation in the dorsal visual stream could be due to visuospatial search for other crucial features of social interaction, since the facial features were not available (Ungerleider, 1994). Previous work in perceiving social stimuli has identified engagement of the lingual gyrus when participants perceived body-only, as opposed to face or face and body of actors (Morris et al., 2006). Several other studies have pointed out a role for the lingual gyrus in processing higher-order motion information/discrimination and spatial processing (Collignon et al., 2011; Servos et al., 2002). Building upon this previous work, our finding suggests that during face-blurred gestures, participants were attending to visuospatial aspects of stimuli, especially motion-related information from gestures. Lastly, given that the perception of blurred faces can potentially be interpreted as an error or conflict, processing of such stimuli could explain activation of the anterior cingulate (Garavan et al., 2003).

Limitations

The theoretical constructs used to characterize gesture types have differed from study to study, with some researchers conceptualizing differences between "transitive" versus "intransitive" gestures (Villarreal et al., 2008), and others considering differences between "expressive" versus "instrumental" gestures (Gallagher and Frith, 2004), and so on. It is not yet clear whether an overarching set of categories can be found to encompass all gestures in humans. In this study, we attempted to avoid fine fragmentation of gesture stimuli and conceived of gesture differences broadly in terms of "social" versus "nonsocial", with added factors of actor-orientation and facial information. Our conceptualization, which overlaps in part with the dichotomies previously proposed, provides a foundation towards constructing a holistic understanding of complex social processing.

Prior studies designed to understand the brain bases of gesture processing have frequently presented the same gestural stimuli multiple times throughout the course of the experiment (Gallagher and Frith, 2004; Knutson et al., 2008; Villarreal et al., 2008). Such repetitions could lead to behavioral and neural habituation. To counteract this possibility, we employed five different actors, each performing ten different gestures. However, it is still possible that repetition of the same gesture type across actors could have lead to similar habituation, although to a lesser degree.

Previous gesture processing studies have often employed explicit instructions (e.g., press a button to indicate direction of gaze). These direct measures have the disadvantage of making the participants aware of the variables of interest in the experiment, which may lead to results that do not reflect rapid implicit brain processes harnessed unconsciously in everyday life. We instead used a cover task of detecting red-dots to avoid any overlap with social processing. We expect that the neural activations uncovered by our paradigm should reflect and overlap with activations that would also be found in an explicit version of our task. In the future, however, an eye-tracker could be used to better capture the implicit processes of social cognition.

Lastly, although we attempted to match the gestures in social and nonsocial condition on the basis of overall actor-movement in the videos, significant differences were found post-hoc between conditions such that the social gestures had higher overall movement than nonsocial gestures (Fig. S2). Thus, some of the differences in activation pattern between social and nonsocial gestures could be associated with overall higher actor-movement in the social video clips.

Conclusions

In this paper we directly examined the neural networks that are engaged in processing natural and realistic social interactions. The data generated suggest that the use of natural stimuli resulted in engagement of a wider network during the processing of social gestures as opposed to nonsocial ones. Further, the right LOC was found to differentially activate for social and oriented-towards trials, thereby suggesting a role of right LOC in processing stimuli that warrant higher social attention. Lastly, by manipulating face-visibility we found that observing facial expressions during social gestures partially droves the activations in the pSTS and FFG areas. Overall, using a novel event related fMRI paradigm these results mark a conceptual advance in understanding the complex dance of human social cognition.

Acknowledgments

This work was supported by the Canal Family Research Fund and NIH grants MH050047, MH064078, MH019908 and HD049653 to A.L.R. The authors thank the reviewers, Drs. Nalini Ambady and Kalanit Grill-Spector for their valuable comments on earlier drafts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.09.046.

References

Axelrod, R., Hamilton, W., 1981. The evolution of cooperation. Science 211, 1390–1396.
Boddaert, N., Chabane, N., Gervais, H., Good, C.D., Bourgeois, M., Plumet, M.-H., Barthélémy, C., Mouren, M.-C., Artiges, E., Samson, Y., Brunelle, F., Frackowiak, R.S.J., Zilbovicius, M., 2004. Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. NeuroImage 23, 364–369.
Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proc. Natl. Acad. Sci. U. S. A. 100, 5497–5502.
Chan, A.W.-Y., Peelen, M.V., Downing, P.E., 2004. The effect of viewpoint on body representation in the extrastriate body area. NeuroReport 15, 2407–2410.
Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., Lepore, F., 2011. Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. Proc. Natl. Acad. Sci. U. S. A. 108, 4435–4440.
Conty, L., Grezes, J., 2012. Look at me, I'll remember you: the perception of self-relevant social cues enhances memory and right hippocampal activity. Hum. Brain Mapp. 33, 2428–2440.
Conty, L., Dezecache, G., Hugueville, L., Grezes, J., 2012. Early binding of gaze, gesture, and emotion: neural time course and correlates. J. Neurosci. 32, 4531–4539.
Fehr, E., Fischbacher, U., 2004. Social norms and human cooperation. Trends Cogn. Sci. 8, 185–190.
Fung, L.K., Quintin, E.-M., Haas, B.W., Reiss, A.L., 2012. Conceptualizing neurodevelopmental disorders through a mechanistic understanding of fragile X syndrome and Williams syndrome. Curr. Opin. Neurol. 25, 112–124.
Gallagher, H.L., Frith, C.D., 2004. Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures. Neuropsychologia 42, 1725–1736.
Garavan, H., Ross, T.J., Kaufman, J., 2003. A midline dissociation between error-processing and response-conflict monitoring. NeuroImage 20, 1132–1139.
Garrett, A.S., Menon, V., MacKenzie, K., Reiss, A.L., 2004. Here's looking at you, kid: neural systems underlying face and gaze processing in fragile X syndrome. Arch. Gen. Psychiatry 61, 281–288.
Glover, G.H., Lai, S., 1998. Self-navigated spiral fMRI: interleaved versus single-shot. Magn. Reson. Med. 39, 361–368.
Golarai, G., Hong, S., Haas, B.W., Galaburda, A.M., Mills, D.L., Bellugi, U., Grill-Spector, K., Reiss, A.L., 2010. The fusiform face area is enlarged in Williams syndrome. J. Neurosci. 30, 6700–6712.
Gold, J.M., Mundy, P.J., Tjan, B.S., 2012. The perception of a face is no more than the sum of its parts. Psychol. Sci. 23, 427–434.
Grèzes, J., Pichon, S., de Gelder, B., 2007. Perceiving fear in dynamic body expressions. NeuroImage 35, 959–967.
Henrich, J., Bowles, S., Boyd, R., Hopfensitz, A., 2003. Genet. Cult. Evol. Coop. 2003.
Hynes, C.A., Baird, A.A., Grafton, S.T., 2006. Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. Neuropsychologia 44, 374–383.
Jeannerod, M., 2004. Visual and action cues contribute to the self-other distinction. Nat. Neurosci. 7, 422–423.
Kennedy, D.P., Adolphs, R., 2012. The social brain in psychiatric and neurological disorders. Trends Cogn. Sci. 16, 559–572.
Knutson, K.M., McClellan, E.M., Grafman, J., 2008. Observing social gestures: an fMRI study. Exp. Brain Res. 188, 187–198.
Kross, E., Egner, T., Ochsner, K., Hirsch, J., Downey, G., 2007. Neural dynamics of rejection sensitivity. J. Cogn. Neurosci. 19, 945–956.
Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., Halsband, U., 2006. Differential cerebral activation during observation of expressive gestures and motor acts. Neuropsychologia 44, 1787–1795.
Mainieri, A.G., Heim, S., Straube, B., Binkofski, F., Kircher, T., 2013. Differential role of mentalizing and the mirror neuron system in the imitation of communicative gestures. NeuroImage 81, 294–305.
Molko, N., Cachia, A., Riviere, D., Mangin, J.F., Bruandet, M., LeBihan, D., Cohen, L., Dehaene, S., 2004. Brain anatomy in Turner syndrome: evidence for impaired social and spatial-numerical networks. Cereb. Cortex 14, 840–850.
Montgomery, K.J., Haxby, J.V., 2008. Mirror neuron system differentially activated by facial expressions and social hand gestures: a functional magnetic resonance imaging study. J. Cogn. Neurosci. 20, 1866–1877.
Montgomery, K.J., Isenberg, N., Haxby, J.V., 2007. Communicative hand gestures and object-directed hand movements activated the mirror neuron system. Soc. Cogn. Affect. Neurosci. 2, 114–122.
Morris, J.P., Pelphrey, K.A., McCarthy, G., 2005. Regional brain activation evoked when approaching a virtual human on a virtual walk. J. Cogn. Neurosci. 17, 1744–1752.
Morris, J.P., Pelphrey, K.A., McCarthy, G., 2006. Occipitotemporal activation evoked by the perception of human bodies is modulated by the presence or absence of the face. Neuropsychologia 44, 1919–1927.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
Nowak, M.A., 2012. Why we help. Sci. Am. 307, 34–39.

Nummenmaa, L., Calder, A.J., 2009. Neural mechanisms of social attention. Trends Cogn. Sci. 13, 135–143.
Pageler, N.M., Menon, V., Merin, N.M., Eliez, S., Brown, W.E., Reiss, A.L., 2003. Effect of head orientation on gaze processing in fusiform gyrus and superior temporal sulcus. NeuroImage 20, 318–329.
Pitcher, D., Dilks, D.D., Saxe, R.R., Triantafyllou, C., Kanwisher, N., 2011. Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56, 2356–2363.
Saxe, R., 2006a. Why and how to study theory of mind with fMRI. Brain Res. 1079, 57–65.
Saxe, R., 2006b. Uniquely human social cognition. Curr. Opin. Neurobiol. 16, 235–239.
Saxe, R., Jamal, N., Powell, L., 2006. My body or yours? The effect of visual perspective on cortical body representations. Cereb. Cortex 16, 178–182.
Schilbach, L., Wohlschlaeger, A.M., Kraemer, N.C., Newen, A., Shah, N.J., Fink, G.R., Vogeley, K., 2006. Being with virtual others: neural correlates of social interaction. Neuropsychologia 44, 718–730.
Servos, P., Osu, R., Santi, A., Kawato, M., 2002. The neural substrates of biological motion perception: an fMRI study. Cereb. Cortex 12, 772–782.
Shapley, R., 2009. Linear and nonlinear systems analysis of the visual system: why does it seem so linear? A review dedicated to the memory of Henk Spekreijse. Vis. Res. 49, 907–921.
Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44, 83–98.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 (Suppl. 1), S208–S219.
Ulloa, J.L., Puce, A., Hugueville, L., George, N., 2012. Sustained neural activity to gaze and emotion perception in dynamic social scenes. Soc. Cogn. Affect. Neurosci. (Online ISSN 1749–5024).
Ungerleider, L., 1994. 'What' and "where" in the human brain. Curr. Opin. Neurobiol. 4, 157–165.
Villarreal, M., Fridman, E.A., Amengual, A., Falasco, G., Gerscovich, E.R., Ulloa, E.R., Leiguarda, R.C., 2008. The neural substrate of gesture recognition. Neuropsychologia 46, 2371–2382.
Völlm, B.A., Taylor, A.N.W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F.W., Elliott, R., 2006. Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. NeuroImage 29, 90–98.
Zucker, N.L., Green, S., Morris, J.P., Kragel, P., Pelphrey, K.A., Bulik, C.M., LaBar, K.S., 2011. Hemodynamic signals of mixed messages during a social exchange. NeuroReport 22, 413–418.