Interaction matters: A perceived social partner alters the neural processing of human speech

CrossMark

Katherine Rice *, Elizabeth Redcay

Department of Psychology, University of Maryland, College Park, MD 20742, USA

ABSTRACT

Mounting evidence suggests that social interaction changes how communicative behaviors (e.g., spoken language, gaze) are processed, but the precise neural bases by which social-interactive context may alter communication remain unknown. Various perspectives suggest that live interactions are more rewarding, more attention-grabbing, or require increased mentalizing—thinking about the thoughts of others. Dissociating between these possibilities is difficult because most extant neuroimaging paradigms examining social interaction have not directly compared live paradigms to conventional "offline" (or recorded) paradigms. We developed a novel fMRI paradigm to assess whether and how an interactive context changes the processing of speech matched in content and vocal characteristics. Participants listened to short vignettes—which contained no reference to people or mental states—believing that some vignettes were prerecorded and that others were presented over a real-time audio-feed by a live social partner. In actuality, all speech was prerecorded. Simply believing that speech was live increased activation in each participant's own mentalizing regions, defined using a functional localizer. Contrasting live to recorded speech did not reveal significant differences in attention or reward regions. Further, higher levels of autistic-like traits were associated with altered neural specialization for live interaction. These results suggest that humans engage in ongoing mentalizing about social partners, even when such mentalizing is not explicitly required, illustrating how social context shapes social cognition. Understanding communication in social context has important implications for typical and atypical social processing, especially for disorders like autism where social difficulties are more acute in live interaction.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Two friends are out for a walk. One turns to the other, points to the sky and says "Look," and both crane their necks upwards to see a hot air balloon. This interchange, and the thousands of moments of their interaction before and after, can be decomposed into constituent parts: speech processing, gesture, gaze following. Each piece alone, however, is neither necessary nor sufficient to create full, experienced interaction, with its emergent proprieties and attunement between social partners that extends beyond input and output stimuli. During real-time social interactions, social partners create a shared psychological state (Tomasello et al., 2005) necessary for successful communication. In this framework, communication is defined not by the medium (e.g., language), but by the desire in the sender to create this shared psychological state, a recognition of that desire by the receiver, and the "shared intentionality" that the two partners create together (Clark, 1996; Sperber and Wilson, 1996; Tomasello et al., 2005). Characterizing the neural bases of communication will improve understanding of this core human behavior, with implications for social disabilities such as

autism, where impairments are most acute in social interaction (Klin et al., 2003).

Past neuroimaging research has almost exclusively examined the component pieces of communication (e.g., emotion understanding, language processing, thinking about others' mental states), despite behavioral evidence that these processes operate differently in interactive, communicative contexts (Gallotti and Frith, 2013; Sebanz et al., 2006). For example, infant interaction with an adult promotes language learning more than observing matched recorded stimuli (Kuhl et al., 2003; Goldstein and Schwade, 2008). In adults, even the belief that one is interacting with another person's avatar (versus an otherwise identical computer agent) improves learning (Okita et al., 2007; see Fox et al., 2015 for a review of avatar vs. agent literature). Individuals also show different gaze patterns when viewing a potential interactive partner as compared to a video (Freeth et al., 2013; Laidlaw et al., 2011; Risko et al., 2012), and naturalistic stimuli better distinguish typical and autism groups (Speer et al., 2007). Thus, although extant neuroimaging research has yielded important insight into the social brain, non-interactive methods cannot capture whether and how the social brain functions differently in social interaction (Schilbach et al., 2013).

Although evidence suggests real-time interaction may affect behavior, the question of *how* real-time interaction alters communication

* Corresponding author. Fax: +1 301 314 9566.
E-mail address: krice@umd.edu (K. Rice).

remains unanswered. Several theories suggest that a fundamental aspect of communication is making rapid inferences about a social partner's communicative intentions and mental states even in simple interactions (e.g., gaze following; Teufel et al., 2009), although the exact nature of these inferences is debated (Clark, 1996; Klin et al., 2003; Shintel and Keysar, 2009; Sperber and Wilson, 1996; Teufel et al., 2009; Tomasello et al., 2005). Despite the possibility that communication and mental state inference—or mentalizing—are intertwined, both processes are typically studied separately in non-interactive contexts. For example, fMRI research identifying a "mentalizing network" has predominately focused on explicit reasoning about the mental state contents (e.g., beliefs) of characters in stories. The role of this brain network in ongoing interaction is unknown. One possibility is that although not all social interaction contains explicit mentalizing demands, implicit mentalizing processes—including tracking mental state content (Schneider et al., 2014b; Senju et al., 2009) and monitoring communicative intent (Kampe et al., 2003)—are more engaged when processing communicative cues from a real-time social partner versus recorded cues.

Greater demand on mentalizing systems is just one theorized difference between processing real-time communication and processing offline, component communicative behaviors. Other possibilities are that live, contingent interactions are more rewarding (Mundy and Neal, 2000; Pfeiffer et al., 2014; Schilbach et al., 2010, 2013), provide unique sources of information compared to recordings (e.g., responsive eye gaze; Kuhl, 2011), or capture more attention (Kuhl, 2007; Risko et al., 2012), perhaps due to increased arousal (Okita et al., 2007), than recorded stimuli. Finally, a tightly controlled experiment may reveal no differences in the processing of matched communicative behaviors in offline versus interactive contexts, suggesting that communication is the sum of its parts. Neuroimaging paradigms may be especially well suited to dissociate these possibilities and to identify the implicit, ongoing processes that are hypothesized to underlie communication and that are difficult to assess via self-report (Schneider et al., 2014a).

Despite the promise of neuroimaging to illuminate communicative processes, extant interactive neuroimaging paradigms have not directly addressed whether and how interactive context changes the processing of well-matched input stimuli, either because that is not the direct question of interest or due to methodological limitations (Redcay et al., 2013b). Past interactive neuroimaging studies have contrasted contingent interaction to non-contingent recordings (Redcay et al., 2010, 2012; Schilbach et al., 2010) and compared direct to averted gaze (Holler et al., 2015), but such paradigms do not directly address whether live context alone alters the processing of well-matched input stimuli. Research in which participants play either human or computer opponents in various games (Coricelli and Nagel, 2009; Gallagher et al., 2002; McCabe et al., 2001) examines engagement with human versus non-human actors, but not how human communication differs in interactive versus non-interactive contexts. Further, such games often demand explicit mentalizing (e.g., deliberate reflection on an actor's motive) and thus cannot determine the extent to which mentalizing is automatically engaged in day-to-day human communication (e.g., pointing to a balloon in the sky). Determining whether and how the neural systems underlying communication differ when communicative acts occur in interactive contexts requires targeted and well-controlled paradigms.

In this study, we employed a novel fMRI paradigm to address the vital gap in our understanding of the neural systems underlying interactive communication. Participants completed trials in which they listened to a short vignette presenting two options, then heard about someone's preference, and made a choice for that person. For some trials, participants believed they were listening to a live social partner over a real-time audio-feed, whereas for other trials they believed they were listening to a recording of another person. Crucially, all stimuli were prerecorded, to ensure matched stimuli within and across participants. To match contingency and attention across conditions, participants saw contingent positive or negative feedback based on their answer to each question. At the end of the experiment, participants completed ratings of liveness, likeability, and engagement for the live and recorded speakers.

Our main analysis examined neural responses during the short vignette portion (i.e., story), which contained no explicit mentalizing demands. Our aim was to identify the neural systems supporting social interaction beyond the processes supporting the interaction's component parts (e.g., processing human speech). We hypothesized that the brain would be differentially engaged when participants perceived the speaker to be talking to them in real-time (i.e., Live) as compared to a matched recording, specifically in each participant's mentalizing regions. We identified each individual's mentalizing regions using a highly replicated language-based theory of mind localizer, which isolates reasoning about mental states from reasoning about physical inferences that are matched on representational and attentional demands (Dodell-Feder et al., 2011). Using a functional localizer is especially important in isolating the regions involved in mentalizing, as portions of the mentalizing network (e.g., TPJ) have also been implicated in more general attentional processes (Decety and Lamm, 2007; Mitchell, 2008) and overlap with portions of the default mode network (Mars et al., 2012; Schilbach et al., 2008; Spreng et al., 2009). As a control, we also compared two recorded conditions: a friendly, engaging voice (Social, which was contrasted with the Live condition), and a less engaging voice (Standard). This contrast between recorded conditions was designed to ensure that effects of audio characteristics, likeability, or attention were not driving differences between perceived live and recorded conditions.

Finally, we examined whether autistic-like personality traits were related to neural sensitivity to live interaction, given that autism has been associated with diminished response to live social partners (Klin et al., 2003) and with preserved explicit mentalizing but impaired on-line implicit mentalizing (Senju et al., 2009). Although all participants in the current study were typical adults, previous research has indicated that typical variability in autistic-like traits is related to behavioral and brain measures of social cognition (e.g., Barman et al., 2015; Bayliss and Tipper, 2005; Hasegawa et al., 2013; Nummenmaa et al., 2012; Miller and Saygin, 2013; Poljac et al., 2013), and there is evidence that the relation between autistic-like traits and social behavior in stronger in live than recorded contexts (Laidlaw et al., 2011). In particular, we theorized that individuals with higher levels of autistic-like traits would show diminished behavioral responsiveness to the live social partner (as measured by self-report) and diminished sensitivity to live versus recorded speech, especially in regions associated with mentalizing, which would provide evidence that sensitivity to social partners is a core social process.

Methods

Participants

Thirty-one adults (13 males), aged 18–27 years, participated in the neuroimaging study in exchange for course credit or payment. All participants were native English speakers, had normal hearing, normal or corrected-to-normal vision, no first-degree relatives with autism or schizophrenia, and no personal history of any neurological impairments or psychological disorders. Two participants were excluded because they did not believe that the interaction in the live condition was live, yielding a final sample of 29 adults. A subset ($n = 23$) of the final sample completed a mentalizing localizer task, and analyses with the localizer regions were restricted to this subset. A separate sample of twenty adults (13 males), aged 18–29 years, completed pilot testing of the audio stimuli, and twenty other adults (8 males), aged 18–28 years completed a behavioral version of the social interaction experiment. The local Institutional Review Board approved all study protocols.

Social interaction experiment

Creating the live illusion

Participants were told that they would listen to short stories and answer questions, and that sometimes the stories and questions would be presented by a social partner via a real-time audio-feed, and that, for those items, they would receive real-time video feedback from the social partner (e.g., thumbs up). Participants were told that other trials would be prerecorded audio with computer-generated feedback. In actuality, all stimuli were prerecorded.

To maintain the live illusion, the experimenter wore the same clothing and had the same appearance as in the prerecorded feedback videos. Further, before the experiment, participants participated in a truly live video chat (Supplementary Fig. 1). Participants only met the live speaker. Participants were debriefed at the end of the experiment.

Task design

The task was programmed and presented using the Psychophysics Toolbox Extension for MATLAB 7.6 (PTB-3; Brainard, 1997). Participants viewed 36 individual trials across 4 runs, 12 from each condition: Live, Social, and Standard (Fig. 1). Each condition had a different female speaker: Live was the live social partner with a friendly tone; Social was a recorded, friendly voice; and Standard was a recorded, neutral voice. Each trial consisted of the story, answering a question by selecting one of two options and receiving feedback. Live feedback was a silent video of the live speaker, Social feedback was a standardized picture of a happy or sad female (Tottenham et al., 2009), and Standard feedback was a gold star or red "x." During all audio, and for 2 s before the start of the story, a screen displayed either LIVE VOICE (in green text) or RECORDED (in orange text).

A 2- to 4-s jittered fixation cross was present between the story and the question period and before feedback. There were also 20 s of baseline (fixation cross) at the start, middle, and end of each run. Trial distribution and timing was determined by OptSeq (http://surfer.nmr.mgh.harvard.edu/optseq/). The event of interest was the story, and collinearity analysis using AFNI's 3dDeconvolve (Cox, 1996) revealed that all beta values of interest were estimable.

Post-test procedure

Participants completed a 7-point Likert-scale questionnaire to assess perceptions of each speaker. For each speaker, participants were asked two questions to assess likeability ("How much did you like this speaker?" and "How much do you think you would like interacting in real life with this speaker?"), two questions to assess engagement ("How much did you pay attention when this speaker was talking?" and "How motivated were you to get the questions asked in her voice right?"), and three questions to assess liveness ("How much did your experiences with her feel live?" "How much did it feel like this speaker was talking directly to you?" and "How much did it feel like this speaker was interacting with you in real-time over a direct connection versus sounding like a recording?"). For each participant, we averaged together these scores on individual items to create composite liveness, likeability, and engagement scores for each speaker.

Participants also completed the Autism Quotient (AQ; Baron-Cohen et al., 2001), which is a self-report measure of autistic-like personality traits. For each item (e.g., "I enjoy social situations") participants answered on a scale of 1 (definitely agree) to 4 (strongly disagree), such that higher scores indicate more autistic-like traits. The AQ captures variability in ASD traits in the typical population (Baron-Cohen et al., 2001), and in typical individuals, higher AQ scores (more autistic-like traits) are related to diminished or atypical social cognitive abilities (e.g., perception of biological motion; Miller and Saygin, 2013; gaze cueing; Bayliss and Tipper, 2005; emotion perception; Poljac et al., 2013), atypical behavioral response to live but not recorded interaction (Laidlaw et al., 2011), and to neural activity in response to social stimuli (Barman et al., 2015; Hasegawa et al., 2013; Nummenmaa et al., 2012). The current sample showed variability in AQ scores, ranging from 8 to 29 (mean = 16.5, SD = 5.4), with all participants scoring below the conventional cutoff for clinical concern (a score of 32; Baron-Cohen et al., 2001).

Stimuli

Each of the three trial types (i.e., Live, Social, Standard) had a unique speaker. Given the current study's within-subjects design, it was critical that participants were immediately aware of which condition they were in. To ensure immediate and reliable recognition of condition, each condition contained both a salient bottom-up cue (speaker voice) and top-down cue (background screen) as to condition.
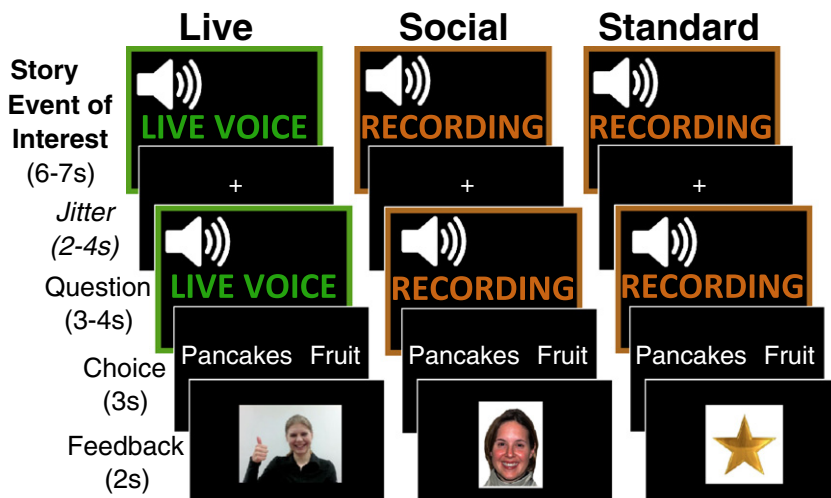

Fig. 1. Experimental trial structure. The three conditions (Live, Social, and Standard) are depicted in the three columns. Participants believed that audio in the Live condition was presented over a live audio-feed, and the other two conditions were recorded. The Live and Social speakers both had a friendly tone and the Standard speaker had a neutral tone. The content of the story portion was identical across conditions and consisted of a two-sentence description of two different objects that contained no reference to people or mental states (e.g., "There are two things on the breakfast menu. One is pancakes and one is a bowl of fruit."). For Live trials, participants answered a question about the live social partner and for the Social and Standard conditions, participants answered question about a third party character (e.g., "I/Megan am/is trying to eat healthy. Which food should I/she eat?"). Participants saw contingent feedback based on their answer to the question. Experimental analyses were restricted to the story portion. s = seconds. (See also Supplementary Fig. 1 & Supplementary Table 1).

All audio was normalized to 60 dB. We compared speakers on pitch and length of their audio clips using Praat 5.3 (Boersma, 2002; Supplementary Table 1). There were no length differences and, as intended, the Standard speaker was lower-pitched than the other speakers. For the story, the Social speaker had higher pitch than the Live speaker. To determine baseline ratings of liveness, likeability, and engagement for all three speakers, we conducted pilot testing ($N = 20$) in which participants listened to audio but were not told any stimuli were live. Unlike in the main experiment, participants did not meet any of the speakers, only heard third-person audio recordings (e.g., "What should Mary do?"), and saw the same feedback (gold star) for each speaker. There were significant effects of speaker on ratings of liveness, likeability, and engagement ($ps < .01$; Supplementary Table 3C). The Social speaker was rated as more live and likeable than the Live speaker, and the Live speaker and Social speaker were more live, likeable, and engaging than the Standard speaker ($ps < .05$).

Each participant was assigned one of three stimuli sets, which differed on which 12 short vignette and question pairs were assigned to each condition, and ensured that the total amount of time for each condition was matched. The order of the items was randomized within condition and the order of runs was counterbalanced across participants.

Social interaction control behavioral study

In addition to the Live speaker being presented as live, there were three other differences between the Live and Social conditions: first-person language (e.g., "I like"), video rewards, and meeting the speaker before the experiment. Given that this novel paradigm represents an initial attempt to understand the neural mechanisms of social interaction, we wanted to ensure that the social-interactive context was sufficiently believable (i.e., meeting the live partner before the experiment) and salient (i.e., being visually reminded of the live interaction via short videos of the live social partner, making a choice about the social partner). Although only the content-matched audio portion was analyzed, these three potential confounds were included in this initial study in order reinforce the live illusion and to create a more ecologically valid social context.

In order to determine if these three factors alone could produce perceived liveness, we conducted a separate behavioral study with $N = 20$ participants (who did not participate in the pilot behavioral experiment or scan) who completed the same task as the fMRI participants, including the post-test questionnaire (Supplementary Table 3B), but were told all stimuli were prerecorded.

Mentalizing localizer

A subset of participants ($n = 23$) completed a mentalizing localizer after the main social interaction experiment, in which they saw written stories about false beliefs and false physical representations (e.g., photographs; Dodell-Feder et al., 2011; http://saxelab.mit.edu/superloc.php). Participants completed two runs, each with five false belief and five false photograph stories (presented for 10 s), followed by a true-false question presented for 4 s. The 14-second block was analyzed as one event.

Image acquisition and preprocessing

Imaging data were collected using a 12-channel head coil on a single Siemens 3.0-T scanner at the Maryland Neuroimaging Center (MAGNETOM Trio Tim System, Siemens Medical Solutions). The scanning protocol for each participant consisted of four runs of the main experiment (T2-weighted echo-planer gradient-echo; 36 interleaved axial slices; voxel size = $3.0 \times 3.0 \times 3.3$ mm; repetition time = 2200 ms; echo time = 24 ms; flip angle = 90°; pixel matrix = 64×64), two runs of the theory of mind localizer (T2-weighted echo-planer gradient-echo; 36 interleaved axial slices; voxel size = $3.0 \times 3.0 \times 3.3$ mm; repetition time = 2000 ms; echo time = 24 ms; flip angle = 90°; pixel matrix = 64×64), and a single structural scan (three-dimensional T1 magnetization-prepared rapid gradient-echo sequence; 176 contiguous sagittal slices, voxel size = $1.0 \times 1.0 \times 1.0$ mm; repetition time = 1900 ms; echo time = 2.52 ms; flip angle = 9°; pixel matrix = 256×256). The parameters for the functional scans for the social interaction experiment were selected after piloting with four typical adults in order to best allow for signal preservation while maximizing specificity in regions prone to signal dropout.

For both the social interaction experiment and the mentalizing localizer, fMRI preprocessing was performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm). Data were slice-time corrected, realigned to the original volume from the first functional run, and then normalized to the Montreal Neurological Institute (MNI) template brain using both linear and nonlinear transformations. Data were high-pass filtered (1/128 Hz) and spatially smoothed using a 5 mm full-width half-maximum (fwhm) Gaussian kernel.

Outlying volumes (i.e., due to motion or global signal) for each participant were detected using the artifact detection toolbox (http://nitrc.org/projects/artifact_detect). Motion outliers were defined as the difference between two consecutive volumes exceeding 1 mm (across translational and rotational movements), and global signal outliers were defined as being three SD from the mean global signal. For the social interaction experiment, participants were to be excluded if, on two or more runs, their number of motion or global signal outliers exceeded 15% of collected volumes or if total motion exceeded 4 mm. No runs, however, met these criteria. For the theory of mind localizer, two participants had runs with between 15% and 20% outliers, but this motion was not correlated with the task and thus all subjects were included in analyses.

Data analysis

Social interaction experiment analysis

General linear models were used to estimate parameters for each condition's story event. The model included each story event convolved with the canonical hemodynamic response function, and, as regressors of no interest, the anticipatory periods, the question periods, and the feedback periods. All six motion parameters (roll, pitch, yaw, x, y, and z) and each individual outlier time point were included as regressors.

We analyzed two specific contrasts: first, to examine the effect of live interaction we compared Live Story vs. Social Story; second, to isolate the effects of prosody and likeability, we compared Social Story vs. Standard Story. Given that the Standard Story was not well matched to the Live Story, that comparison was less theoretically interesting, and we did not analyze that contrast. Contrast maps were thresholded at a two-tailed $p < .001$ and cluster-corrected for multiple comparisons (overall alpha = .05, $k = 93$) using SPM's false discovery rate algorithm. In order to examine individual differences in sensitivity to live interaction, we extracted each individual's contrast value from the clusters identified in the group-level analysis. Specifically, we extracted each individual's Social > Standard contrast values for the clusters showing a significant effect of group for Social > Standard and similarly, Live > Social values for the clusters showing a significant effect of group for Live > Social. We examined the correlations between these individual contrast values and both AQ scores and perceived liveness of the Live speaker. We specifically examined liveness (rather than likeability or engagement) as we designed this rating to quantify the participant's experience of the live illusion.

Mentalizing localizer analysis

The regression model included the same motion and outlier nuisance regressors as the social interaction experiment and additional regressors for the belief and photo story conditions. To identify each

individual's regions associated with explicit mentalizing, we examined each participant's whole-brain contrast for false belief vs. false photo stories at $p < .001$, $k = 10$ voxels, and identified peak coordinates for eight regions of interest within each participant (Young et al., 2007): bilateral TPJ, bilateral aSTS, precuneus, dMPFC, middle MPFC (mMPFC), and ventral MPFC (vMPFC). All participants with the localizer scan ($n = 23$) had identifiable activation in all regions except for mMPFC ($n = 21$) and vMPFC ($n = 18$). Voxels that were significantly active within a 9 mm radius sphere surrounding each participant's peak coordinate were used as regions of interest. Again, within each region, we analyzed two specific contrasts: first, Live Story vs. Social Story to examine the effects of live social interaction on activation in mentalizing regions; and second, Social Story vs. Standard Story, to examine the effects of prosody and likeability in mentalizing regions. Given that our specific a priori hypotheses related to these pairwise comparisons, in order to minimize comparisons, we did not compare the effects of Live Story to Standard Story in these regions nor did we conduct a repeated-measures ANOVA across all three conditions.

Results

Perceived interactive context modulates brain response to speech

In order to examine how live interaction modulates the neural correlates of language processing, we compared BOLD activation during the story for the Live condition to the recorded Social condition. Whole-brain results demonstrated that processing speech from a live social partner modulated neural activity compared to processing matched recorded human speech (Fig. 2A; Supplementary Table 2 & Supplementary Fig. 2), specifically in regions often associated with social cognition (i.e., mentalizing) and social engagement (Frith and Frith, 2006), including dorsomedial prefrontal cortex (dMPFC) and temporal parietal junction (TPJ).

Effect of live condition reflects perceived interactive communication

Perceived liveness is driven by the live illusion

To ensure that these whole-brain effects were due to perceived live interaction and were not reducible to other between-condition differences, we investigated several potential confounds. First, our condition manipulation relied on a participant's belief that the Live condition was live. In order to confirm and quantify belief in the illusion, all participants filled out rating scales after the fMRI session. All participants believed the live illusion. To further check the top-down manipulation of

perceived liveness, we examined participant ratings of speaker liveness (e.g., "How much did it feel like this speaker was talking directly to you?"). Participants rated the Live speaker as significantly more live than the Social speaker ($p < .001$; Fig. 3A; Supplementary Table 3A). Greater perceived liveness of the Live speaker was also marginally correlated with greater activation for Live vs. Social speech in rDMPFC ($r = .34$, $p = .07$), as measured by individually extracted contrast values from the group-defined cluster, although correlations with perceived liveness were not significant for any of the other identified regions. Further, there was also no correlation between perceived liveness and Social vs. Standard speech activation in right STG, the region defined by that group-level contrast ($r < .1$).

In addition to evidence from the neuroimaging study, an additional behavioral experiment with a separate group of participants ($N = 20$) was conducted to analyze factors that were confounded with the live condition. These potential confounds were included to enhance the illusion of liveness: participants met only the live speaker briefly before the start of the experiment, saw the live speaker give ostensibly live video feedback, and heard the live speaker use first-person language in the unanalyzed question portion. In this separate study, where live stimuli were presented as recorded, these three confounding factors failed to produce significantly higher liveness ratings for the Live compared to Social speaker (Supplementary Table 3B). That is, only when participants were told that they were interacting with a live social partner did they actually rate that speaker as feeling significantly more live.

Effects of live context are not reducible to other differences between conditions

One factor that could account for differences between live and recorded interactions is differences in attention. To examine this, we compared reaction time across conditions for the question that followed the story and found no significant differences ($F < 1$, Supplementary Table 4). There were also no significant between-condition differences in accuracy, but the very high accuracy for all questions precludes using that measure to make interpretations about levels of attention across conditions.

Self-reported social engagement and speaker likeability, however, did differ across conditions. Participants rated the Live speaker as more likeable and socially engaging than the Social speaker ($ps < .01$). Thus, to ensure that whole-brain Live > Social effects were not driven by likeability or engagement, we compared Social and Standard speakers, who also differed on likeability and engagement ($ps < .01$). This comparison also helped isolate any low-level effects of audio characteristics, as both Live versus Social and Social versus Standard
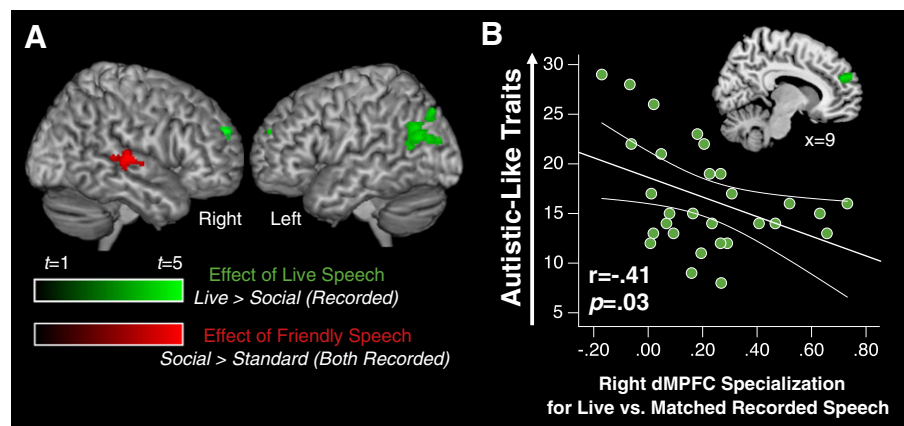

Fig. 2. Neural correlates of processing live versus recorded speech. (A) Whole-brain comparisons ($N = 29$) of Live Story versus Social Story and Social Story versus Standard Story (corrected $p < .05$). Listening to live speech versus a matched recording activated regions of the brain associated with processing social stimuli, and this activation was dissociable from the effect of listening to a friendly and engaging speaker (Social > Standard). (B) The relation between autistic-like traits, as measured by the Autism Quotient, and individual contrast values for Live Story > Social Story ($N = 29$). Right dorsomedial prefrontal cortex (dMPFC) was defined based on the active voxels for the group-level contrast in (A). Higher levels of autistic-like traits were associated with lower levels of differential activation for live stimuli. (See also Table 2 and Supplementary Fig. 2).
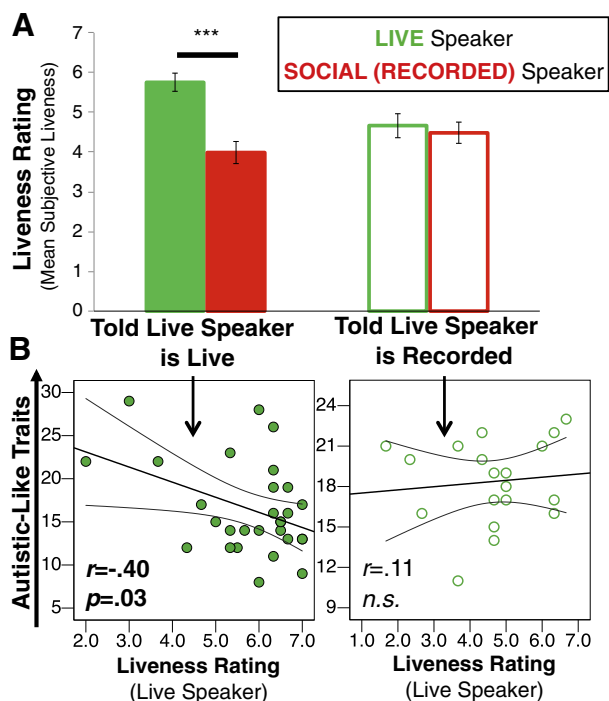
Fig. 3. Behavioral impressions of speaker liveness. (A) Perceived liveness is driven by the illusion of a live setup. All scanned participants ($N = 29$) were told that the Live condition was live, and they rated the live and matched recorded (i.e., Social) speakers as significantly different on a post-test questionnaire of perceived liveness. Perceived liveness was assessed using a composite of Likert-scale items scored on a 1 to 7 scale (e.g., how direct the speaker seemed, how much it seemed as though the speaker was in the room with the participant). This difference disappeared on an identical behavioral version of the task in which participants ($N = 20$) were told all audio—including audio from the scan's Live speaker—was prerecorded. (B) Perceptions of liveness for the Live condition speaker were significantly correlated with autistic-like traits (as measured by the Autism Quotient) only for the group that thought they were participating in a live interaction. For both groups, correlations with Autism Quotient scores and perceived liveness of the Social speaker were not significant. ***$p < .001$ (see also Supplementary Table 3).

conditions differed on mean pitch. The comparison of the two not-live conditions (Social > Standard) only implicated regions frequently associated with auditory and speech processing (right superior temporal gyrus), and the specific activation was consistent with pitch responsivity (www.neurosynth.org; posterior probability for pitch = 0.87). There was no overlap in regions identified by the Live > Social and Social > Standard comparisons.

Ongoing mentalizing supports communicative interaction

Given that additional analyses revealed that the Live > Social effects were driven by the experimental manipulation of a simulated live interaction and not by potential confounding effects, we next examined the Live > Social contrast in each participant's individually identified mentalizing network (defined using an independent localizer; Dodell-Feder et al., 2011). All eight mentalizing regions showed the hypothesized pattern of results: increased activation for live versus recorded social speech. Specifically, this difference was significant in six of the eight identified regions: bilateral TPJ (right: $t(22) = 3.67$, $p = .001$; left: $t(22) = 4.21$, $p < .001$), precuneus ($t(22) = 2.30$, $p = .032$), dMPFC ($t(22) = 2.33$, $p = .029$), middle MPFC ($t(20) = 3.97$, $p < .001$), and ventral MPFC ($t(17) = 4.25$, $p < .001$). The difference between Live and Social was in the same direction but did not reach significance for the remaining two regions: right ($t(22) = 1.99$, $p = .059$) and left ($t(22) = 1.82$, $p = .083$) anterior superior temporal sulcus (aSTS). After applying a Holm–Bonferroni correction for multiple comparisons (Holm, 1979), significant differences between Live and Social remained in bilateral TPJ, middle MPFC, and ventral MPFC (corrected $ps < .01$).

Social speech, as compared to Standard speech, did not show significantly greater activation in any of the eight mentalizing regions ($ps > .1$; Supplementary Fig. 3).

Autistic-like personality traits relate to processing live stimuli

Autism Quotient (AQ; Baron-Cohen et al., 2001) scores, which measure autistic-like traits, were significantly correlated with perceived liveness for the Live condition (Fig. 3B), but not the Social ($r(27) = -.27$, $p = .17$) nor Standard ($r(27) = .13$, $p = .49$) conditions, such that greater autistic-like personality traits were related to diminished perceptions of live social interaction during the Live condition. This correlation between AQ scores and perceived liveness remained significant after controlling for both engagement and likeability ratings ($r(25) = -.54$, $p = .003$). AQ scores were not correlated with perceived liveness for any of the speakers in the control behavioral experiment when participants were told that all conditions were prerecorded.

Given that activation in right dMPFC for Live vs. Social speech was correlated with perceived liveness, we next examined the relation between AQ scores and specialization for live speech in this region. AQ scores were also significantly correlated with specialization in right dMPFC (defined by the group-level contrast) for Live vs. Social speech (Fig. 2B). Participants with higher levels of autistic-like traits showed diminished neural specialization for live versus matched recorded speech. Additional analyses with the other three group-identified clusters revealed a negative trend in left dMPFC ($r = -.36$, $p = .06$) and a significant positive relation in left angular gyrus ($r = .37$, $p = .045$). There was no correlation between AQ scores and activation to Social vs. Standard speech in the right STG cluster defined by the corresponding group-level contrast ($r < .1$).

Discussion

This study used a novel paradigm to isolate the neural systems subserving live human communication. Specifically, this paradigm identified the neural mechanisms underlying the differential processing of otherwise well-matched stimuli in live and recorded contexts. Participants completed an fMRI task in which they listened to two types of prerecorded speech: speech that they believed was being spoken in real-time by the experimenter in another room over a live audio-feed (Live) and matched human speech that participants knew was recorded (Social). Analyses of the content- and prosody-matched story portion revealed that simply believing that a social partner was speaking in real-time more strongly activated social cognitive regions, specifically in each participant's independently identified mentalizing network. Further investigation revealed that this effect was not attributable to low-level differences in audio characteristics, speaker likeability, or attention, and was instead driven by perceptions of speaker liveness. These perceptions of liveness and neural responses to live interaction were related to Autism Quotient (AQ) scores; individuals with higher levels of autistic-like personality traits were less sensitive to the live context.

In the current paradigm, the analyzed speech segments did not contain any explicit mentalizing demands, and yet live speech still more strongly activated the mentalizing network. This finding indicates that mentalizing may be a critical component of real-time interaction. Live interaction did not increase activation in attention or reward networks, in contrast to some past findings examining social interaction (Pfeiffer et al., 2014; Schilbach et al., 2010). The current paradigm, however, made efforts to control contingency and attention between conditions, and no reaction time differences were found between conditions. Past interactive studies that have found reward system activation have examined interacting with humans versus computers (Pfeiffer et al., 2014) or the effect of directing another person's attention (Schilbach et al., 2010), neither of which was examined in the current paradigm. Importantly, the null finding of the current study does not contradict

evidence suggesting that humans are intrinsically motivated to engage with others (Chevallier et al., 2012; Tomasello, 2008) and that motivational brain networks play an important role in social interaction, but rather indicates that real-world, dynamic interaction is likely subserved by several systems, which may be targeted by different contexts.

Activation in regions identified by the mentalizing localizer provides strong evidence that the regions engaged by live stimuli in the current study are part of the network involved in explicit mental state inference. Specifically, although regions from the broader mentalizing network have also been implicated in more domain-general processes (e.g., attention), this functional localizer matched attentional and other cognitive demands across mentalizing and control items in order to identify, for each individual, the regions associated with mental state reasoning. Further, the regions identified by the localizer revealed no increased activation for the more friendly recorded condition versus the more standard recorded condition, despite the fact that these conditions differed on self-reported engagement and low-level audio characteristics. Thus, although it is possible that other processes support the complex phenomenon of live interaction, these convergent results suggest that the mentalizing network is altered during social interaction.

In the context of the current study, it is unknown how these regions involved in explicit mentalizing subserve interaction. Perhaps most straightforwardly, participants could be engaging in more explicit or even implicit (Schneider et al., 2014a, 2014b) reasoning about the mental state contents of the live versus recorded speaker within the context of experimental task demands. Three factors, however, argue against this interpretation: first, the analyzed story portion did not contain mentalizing language; second, the mental state content in the unanalyzed question portion of the trial was matched across conditions; and third, there were no behavioral differences between conditions on the questions after each story, which required mental state reasoning to answer.

Given the highly controlled experimental context, increased activation in the mentalizing network may result from tracking belief states beyond task demands. For example, individuals may reflect on the evaluative or interpersonal mental states of the live social partner (e.g., Does she think I am giving quick enough answers?) but not of the recorded speakers. This ongoing monitoring of the mental states of a social partner, perhaps particularly in regards to how that social partner perceives the interaction, may assist in the creation of a shared psychological state and ensure successful communication. Another possibility is that the mentalizing system is more active not because of ongoing mental state reasoning, but rather because live interaction inherently contains the possibility of suddenly needing to consider another's mental state. That is, the system is primed for activation. Both of these possibilities speak to facets of human interpersonal understanding that may separate live interaction from the sum of its component pieces.

Alternatively, rather than the mentalizing system supporting quantitatively "more" belief state reasoning during live interaction, it is possible that the mentalizing system is engaged in a qualitatively different way. Computationally, the live and recorded conditions are equally demanding in terms of understanding spoken language and making mental state inferences. Thus, perhaps increased activation in these regions during live speech is not driven by differences in representational content, but rather by this network's broader role in social interaction. The exact nature of this role is unknown, but may involve the rapid, ongoing representation of a social partner that underlies phenomena such as social resonance, synchrony, and coordination (Garrod and Pickering, 2004; Kopp, 2010). Interestingly, previous fMRI studies of speech comprehension have implicated regions also involved in mentalizing (e.g., aSTS; Scott et al., 2000, MPFC; Obleser et al., 2007, pSTS; review, Redcay, 2008, and TPJ; review, Mar, 2011), suggesting a coupling between social and speech processing within this network. Future research should continue to investigate the role of mentalizing systems in speech processing even outside of live contexts.

Due to the constrained and controlled nature of the current paradigm, increased activity in the mentalizing network was necessarily driven by primarily top-down information (i.e., participants were told stimuli were live) as opposed to more bottom-up cues (e.g., real-time contingency, nonverbal synchrony, "interactive alignment"; Pfeiffer et al., 2014; Garrod and Pickering, 2004). This top-down information about speaker liveness affected other judgments about the speaker: as compared to a control experiment where participants knew the live speaker was recorded, simply being told that the speaker was live significantly increased ratings of social engagement and likeability for that speaker. Ontogenetically, in contrast, the perception of live interaction unspools via bottom-up cues and is likely implicit, given that young infants are sensitive to contingent interaction (Murray and Trevarthen, 1985). How these bottom-up processes differ from the top-down experimental manipulation of liveness in the current study is unclear. Understanding factors influencing the perception of social interaction has implications for real-world pedagogical contexts, including listening to and learning from live versus recorded material, and is an important direction for future study.

Differential behavioral and neural response to live social partners has implications for autism. Research suggests that social disability in autism may result from a developmental cascade beginning with early disengagement from social stimuli (Klin et al., 2003; Pelphrey et al., 2011). Although the current study examined typical individuals, higher AQ scores were associated with diminished neural sensitivity to live versus recorded speech in dMPFC, a region linked to atypical ASD activation during live interaction (Redcay et al., 2013a), but not during offline, explicit mentalizing (Dufour et al., 2013). The mechanism driving this association in the current sample is unclear, but one possibility is found in the behavioral ratings of speaker liveness. Higher AQ scores were correlated with lower ratings of perceived liveness of the live speaker, but not the recorded speakers, and, further, increased perceived liveness of the live speaker was related to increased activation in dMPFC. Although the current sample size precludes a mediation analysis, the relation between AQ scores and sensitivity to live social partners is concordant with the hypothesis that individuals with autism feel less social resonance with their social partners (Izuma et al., 2011; Klin et al., 2003) and suggests that heightened sensitivity for live social partners may characterize typical social development. Interestingly, left angular gyrus specialization for live speech was associated with *increased* levels of autistic-like traits. This angular gyrus cluster did not overlap with regions identified by the mentalizing localizer, suggesting a nonlinear pattern where reduced sensitivity to live interaction within the mentalizing system, but increased sensitivity outside of this system, may be associated with atypical social processing. Given that the current study had a relatively small number of individuals with high levels of autistic-like traits, however, future work should examine larger typical samples and also test individuals with autism in similar paradigms.

For this initial within-subjects approach to understanding the effect of live interaction on communication, we needed to ensure a believable and salient interactive setup. We thus created an experiment that ensured participants believed the live illusion (i.e., meeting the live partner beforehand) and that reinforced to participants that they were engaging with a real social partner (i.e., "live" videos of the partner, hearing the partner use first-person language, hearing a different speaker for the recorded trials). Indeed, this manipulation had a high success rate as 29 out of 31 adults believed the live illusion. However, interpretations of the current paradigm are limited by the fact that the matched live speech was embedded in this social context. Although our additional analyses indicate that these confounds are not driving the results, future work could employ paradigms with sparser social context. For example, participants could listen to the same speaker, but be told for some trials that the speaker is live and for other trials that the speaker is recorded. If these two conditions do not reveal differences in neural activation, it is possible that with all other factors controlled, belief that speech is presented via a real-time audiofeed is not enough to alter neural processing. However, it is also possible that without a richer social context, participants would find the live stimuli less believable or

that, given the sparse, top-down instruction, the live nature of the stimuli would not be salient enough to affect speech processing (cf. Okita et al., 2007). Future studies could manipulate various aspects of social-interactive setups to help isolate the necessary and sufficient conditions to observe the effects found in the current study.

The current study offers important insight into understanding a fundamental human behavior: communication. Regions of the mentalizing network have been most frequently conceptualized as carrying out specific social cognitive computations (e.g., determining another's beliefs). The current study illustrates that, even when explicit social cognitive demands are held constant, the mentalizing network is sensitive to social interaction. Specifically, this paradigm demonstrates that merely believing that you are listening to a live speaker (versus a matched recording) changes the neural correlates of processing speech, specifically by increasing activation in traditionally defined mentalizing regions. Although the current paradigm only examined spoken language, research suggests that the communicative medium is less important than the creation of a shared psychological state (Noordzij et al., 2009; Stolk et al., 2014). Suggestively, the regions identified in the current study are similar to those identified in studies of interactive but nonverbal joint visual attention (Redcay et al., 2010, 2012). In addition to illuminating the neural mechanisms underlying interaction, the current findings attest to the necessity of studying communicative cues (e.g., speech) in a communicative context. Future research on disorders characterized by social disabilities that are more acute in live interaction, such as autism and social anxiety, should consider the role of communicative context; interactive paradigms may be more suited to identify these disorders' neural mechanisms. Ultimately, if scientists have the goal of understanding human behavior, such study must take into account human context.

Conflict of interest

None

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.11.041.

References

Barman, A., Richter, S., Soch, J., Deibele, A., Richter, A., Assmann, A., ... Schott, B.H., 2015. Gender-specific modulation of neural mechanisms underlying social reward processing by Autism Quotient. Soc. Cogn. Affect. Neurosci. 10, 1537–1547.
Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., 2001. The autism-spectrum quotient AQ: evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J. Autism Dev. Disord. 31, 5–17.
Bayliss, A.P., Tipper, S.P., 2005. Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. Br. J. Psychol. 96, 95–114.
Boersma, P., 2002. Praat, a system for doing phonetics by computer. Glot Int. 5, 341–345.
Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.
Chevallier, C., Kohls, G., Troiani, V., Brodkin, E.S., Schultz, R.T., 2012. The social motivation theory of autism. Trends Cogn. Sci. 164, 231–239.
Clark, H.H., 1996. Using Language. Cambridge University, Cambridge.
Coricelli, G., Nagel, R., 2009. Neural correlates of depth of strategic reasoning in medial prefrontal cortex. Proc. Natl. Acad. Sci. U. S. A. 106, 9163–9168.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance NeuroImages. Comput. Biomed. Res. 29, 162–173.
Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. Neuroscientist 13, 580–593.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., Saxe, R., 2011. fMRI item analysis in a theory of mind task. NeuroImage 55, 705–712.
Dufour, N., Redcay, E., Young, L., Mavros, P.L., Moran, J.M., Triantafyllou, C., Gabrielli, J.D.E., Saxe, R., 2013. Similar brain activation during false belief tasks in a large sample of adults with and without autism. PLoS One 8 (9), e75468.
Fox, J., Ahn, S.J., Janssen, J.H., Yeykelis, L., Segovia, K.Y., Bailenson, J.N., 2015. Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. Hum. Comput. Interact. 30, 401–432.
Freeth, M., Foulsham, T., Kingstone, A., 2013. What affects social attention? Social presence, eye contact and autistic traits. PLoS One 8 (1), e53286.
Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. Neuron 504, 531–534.
Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. NeuroImage 163, 814–821.
Gallotti, M., Frith, C.D., 2013. Social cognition in the we-mode. Trends Cogn. Sci. 174, 160–165.
Garrod, S., Pickering, M.J., 2004. Why is conversation so easy? Trends Cogn. Sci. 8, 8–11.
Goldstein, M.H., Schwade, J.A., 2008. Social feedback to infants' babbling facilitates rapid phonological learning. Psychol. Sci. 19, 515–523.
Hasegawa, N., Kitamura, H., Murakami, H., Kameyama, S., Sasagawa, M., Egawa, J., ... Someya, T., 2013. Neural activity in the posterior superior temporal region during eye contact perception correlates with autistic traits. Neurosci. Lett. 549, 45–50.
Holler, J., et al., 2015. Eye'm talking to you: speakers' gaze direction modulates co-speech gesture processing in the right MTG. Soc. Cogn. Affect. Neurosci. 10, 255–261.
Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.
Izuma, K., Matsumoto, K., Camerer, C.F., Adolphs, R., 2011. Insensitivity to social reputation in autism. Proc. Natl. Acad. Sci. U. S. A. 108, 17302–17307.
Kampe, K.K., Frith, C.D., Frith, U., 2003. "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing", regardless of modality. J. Neurosci. 23, 5258–5263.
Klin, A., Jones, W., Schultz, R., Volkmar, F., 2003. The enactive mind, or from actions to cognition: lessons from autism. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 345–360.
Kopp, S., 2010. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. Speech Comm. 52, 587–597.
Kuhl, P.K., 2007. Is speech learning 'gated' by the social brain? Dev. Sci. 10, 110–120.
Kuhl, P.K., 2011. Social mechanisms in early language acquisition: understanding integrated systems supporting language. In: Decety, J., Cacioppo, J.T. (Eds.), The Oxford Handbook of Social Neuroscience. Oxford University Press, Oxford, pp. 649–667.
Kuhl, P.K., Tsao, F.M., Liu, H.M., 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Proc. Natl. Acad. Sci. U. S. A. 100, 9096–9101.
Laidlaw, K.E., Foulsham, T., Kuhn, G., Kingstone, A., 2011. Potential social interactions are important to social attention. Proc. Natl. Acad. Sci. U. S. A. 108, 5548–5553.
Mar, R.A., 2011. The neural bases of social cognition and story comprehension. Annu. Rev. Psychol. 62, 103–134.
Mars, R.B., Neubert, F.X., Noonan, M.P., Sallet, J., Toni, I., Rushworth, M.F., 2012. On the relationship between the "default mode network" and the "social brain". Front. Hum. Neurosci. 6, 189.
McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. Proc. Natl. Acad. Sci. U. S. A. 98, 11832–11835.
Miller, L.E., Saygin, A.P., 2013. Individual differences in the perception of biological motion: links to social cognition and motor imagery. Cognition 128, 140–148.
Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. Cereb. Cortex 18, 262–271.
Mundy, P., Neal, R.A., 2000. Neural plasticity, joint attention, and a transactional social-orienting model of autism. Int. Rev. Res. Ment. Retard. 23, 139–168.
Murray, L., Trevarthen, C., 1985. Emotional regulation of interactions between two-month-olds and their mothers. In: Field, T.M., Fox, N.A. (Eds.), Social Perception in Infancy. Ablex, Norwood, NJ, pp. 177–198.
Noordzij, M.L., et al., 2009. Brain mechanisms underlying human communication. Front. Hum. Neurosci. 3, 14.
Nummenmaa, L., Engell, A.D., Von Dem Hagen, E., Henson, R.N., Calder, A.J., 2012. Autism spectrum traits predict the neural response to eye gaze in typical individuals. NeuroImage 59, 3356–3363.
Obleser, J., Wise, R.J., Dresner, M.A., Scott, S.K., 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. J. Neurosci. 27, 2283–2289.
Okita, S.Y., Bailenson, J., Schwartz, D.L., 2007. The mere belief of social interaction improves learning. In: McNamara, D.S., Trafton, J.G. (Eds.), The Proceedings of the 29th Meeting of the Cognitive Science Society, pp. 1355–1360.
Pelphrey, K.A., Shultz, S., Hudac, C.M., Vander Wyk, B.C., 2011. Research review: constraining heterogeneity: the social brain and its development in autism spectrum disorder. J. Child Psychol. Psychiatry 52, 631–644.
Pfeiffer, U.J., et al., 2014. Why we interact: on the functional role of the striatum in the subjective experience of social interaction. NeuroImage 101, 124–137.
Poljac, E., Poljac, E., Wagemans, J., 2013. Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits. Autism 17, 668–680.
Redcay, E., 2008. The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. Neurosci. Biobehav. Rev. 32, 123–142.
Redcay, E., et al., 2010. Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. NeuroImage 50, 1639–1647.
Redcay, E., Kleiner, M., Saxe, R., 2012. Look at this: the neural correlates of initiating and responding to bids for joint attention. Front. Hum. Neurosci. 6, 169.

Redcay, E., Dodell-Feder, D., Mavros, P.L., Kleiner, M., Pearrow, M.J., Triantafyllou, C., Gabrieli, J.D.E., Saxe, R., 2013a. Atypical brain activation patterns during a face-to-face joint attention game in adults with autism spectrum disorder. Hum. Brain Mapp. 34, 2511–2523.

Redcay, E., Rice, K., Saxe, R., 2013b. Interaction versus observation: a finer look at this distinction and its importance to autism. Behav. Brain Sci. 36, 435.

Risko, E.F., Laidlaw, K.E., Freeth, M., Foulsham, T., Kingstone, A., 2012. Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. Front. Hum. Neurosci. 6, 143.

Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., Vogeley, K., 2008. Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. Conscious. Cogn. 17, 457–467.

Schilbach, L., et al., 2010. Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. J. Cogn. Neurosci. 22, 2702–2715.

Schilbach, L., et al., 2013. Toward a second-person neuroscience. Behav. Brain Sci. 36, 393–414.

Schneider, D., Nott, Z.E., Dux, P.E., 2014a. Task instructions and implicit theory of mind. Cognition 133, 43–47.

Schneider, D., Slaughter, V.P., Becker, S.I., Dux, P.E., 2014b. Implicit false-belief processing in the human brain. NeuroImage 101, 268–275.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123, 2400–2406.

Sebanz, N., Bekkering, H., Knoblich, G., 2006. Joint action: bodies and minds moving together. Trends Cogn. Sci. 102, 70–76.

Senju, A., Southgate, V., White, S., Frith, U., 2009. Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. Science 325, 883–885.

Shintel, H., Keysar, B., 2009. Less is more: a minimalist account of joint action in communication. Top. Cogn. Sci. 12, 260–273.

Speer, L.L., Cook, A.E., McMahon, W.M., Clark, E., 2007. Face processing in children with autism effects of stimulus contents and type. Autism 113, 265–277.

Sperber, D., Wilson, D., 1996. Relevance: Communication and Cognition. second ed. Blackwell, Oxford.

Spreng, R.N., Mar, R.A., Kim, A.S., 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J. Cogn. Neurosci. 21, 489–510.

Stolk, A., et al., 2014. Cerebral coherence between communicators marks the emergence of meaning. Proc. Natl. Acad. Sci. U. S. A. 111, 18183–18188.

Teufel, C., et al., 2009. Social cognition modulates the sensory coding of observed gaze direction. Curr. Biol. 19, 1274–1277.

Tomasello, M., 2008. Origins of Human Communication. MIT Press, Cambridge.

Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., 2005. Understanding and sharing intentions: the origins of cultural cognition. Behav. Brain Sci. 28, 675–691.

Tottenham, N., et al., 2009. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res. 168, 242–249.

Young, L., Cushman, F., Hauser, M., Saxe, R., 2007. The neural basis of the interaction between theory of mind and moral judgment. Proc. Natl. Acad. Sci. U. S. A. 104, 8235–8240.