doi:10.1093/scan/nsu111 SCAN (2015) 10, 729–734

Neural coding of assessing another person's knowledge based on nonverbal cues

Anna K. Kuhlen,[1,2,3,*] Carsten Bogler,[2,3,*] Marc Swerts,[4] and John-Dylan Haynes[1,2,3]

[1]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10099 Berlin, Germany, [2]Bernstein Center for Computational Neuroscience Berlin and Charité—Universitätsmedizin Berlin, 10115 Berlin, Germany, [3]Berlin Center for Advanced Neuroimaging, Charité – Universitätsmedizin Berlin, 10119 Berlin, Germany, and [4]School of Humanities, TiCC Research Center, Tilburg University, 5000 LE, Tilburg, The Netherlands

For successful communication, conversational partners need to estimate each other's current knowledge state. Nonverbal facial and bodily cues can reveal relevant information about how confident a speaker is about what they are saying. Using functional magnetic resonance imaging, we aimed to identify brain regions that encode how confident a speaker is perceived to be. Participants viewed videos of people answering general knowledge questions and judged each respondent's confidence in their answer. Our results suggest a distinct role of two neural networks known to support social inferences, the so-called mentalizing and the mirroring network. While activation in both networks underlies the processing of nonverbal cues, only activity in the mentalizing network, most notably the medial prefrontal cortex and the bilateral temporoparietal junction, is modulated by how confident the respondent is judged to be. Our results support an integrative account of the mirroring and mentalizing network, in which the two systems support each other in aiding pragmatic processing.

Keywords: feeling of another's knowing; nonverbal; mentalizing; mirroring; fMRI

Face-to-face communication is not only about 'what' is said but also 'how' it is said. Along with the message content, meta-communicative information informs the pragmatic interpretation of the message (so-called 'track 2' signals; Clark, 1996). One important type of meta-communicative information concerns the trustworthiness of the communicated information. Often we have an intuition about how confident our conversational partner is about what they are saying. This has consequences for how an interaction develops: if we think our partner is not confident, we may take the information they provide with a grain of salt, and seek for clarification or confirmation that what they say is correct. Using functional magnetic resonance imaging (fMRI), we investigated the neural mechanisms underlying inferences about a speaker's knowledge. Rather than focusing on how the brain processes what is said, we were interested in how the brain represents the perceived confidence of the speaker.

Inferences about a speaker's knowledge have been investigated in the psycholinguistic literature under the concept 'Feeling of Another's Knowing' (abbreviated: FOAK; Brennan and Williams, 1995). Although termed a feeling (as reference to the well-known 'feeling of knowing' first described by Hart in 1965), it refers to a metacognitive process by which an observer makes an inference about how confident another person appears to be of their knowledge.

Although a person's confidence in their knowledge is an internal mental state, it is often displayed outwardly in their behavior (Smith and Clark, 1993; Brennan and Williams, 1995; Swerts and Krahmer, 2005). For example, imagine a person responding to the question 'How do I get to Alexanderplatz?' with 'Take tram line 2'. While responding, the respondent may look straight at the person posing the question and smile in confidence that their answer is correct. But if less confident, the respondent may avert their gaze, or make a face. Nonverbal cues like these can inform a perceiver's inference about another person's knowledge (Swerts and Krahmer, 2005).

Two different types of cognitive processes may support inferences about another person's knowledge based on nonverbal cues: (i) Mirroring: inferences about another person's actions or mental states (we call them 'social inferences' from hereon) have been proposed to be based on an imitation process that relies on shared motor representations for performing actions and for observing similar actions performed by others (see e.g. Gallese and Goldman, 1998; Keysers and Gazzola, 2007). According to this view, observers may interpret the knowledge state of another by mapping perceived nonverbal behavior onto their own action repertoire (and associated mental states). (ii) Mentalizing: others have proposed that social inferences may rely on a reflective higher-level cognitive process comparable with the process known as attribution in social psychology (see e.g. Gallagher and Frith, 2003; Amodio and Frith, 2006; Frith and Frith, 2012). According to this view, observers may use perceived nonverbal behavior to form a more abstract judgment about the other person's knowledge.

Each of these cognitive processes, one imitative and the other reflective, has been associated with a different brain circuit (see Van Overwalle and Baetens, 2009 for a review). Mirroring is typically associated with the posterior superior temporal sulcus (pSTS), the premotor cortex (PMC) and the anterior intraparietal sulcus (aIPS) (Van Overwalle and Baetens, 2009). In particular, social inferences based on nonverbal behavior have been associated with areas of the so-called mirroring network, namely, the right fronto-parietal cortex and the right pSTS (Zaki *et al.*, 2010; Kuzmanovic *et al.*, 2012). Mentalizing, on the other hand, is commonly reported to involve the medial prefrontal cortex (mPFC), the bilateral temporoparietal junction (TPJ) and the precuneus (PC) (Van Overwalle and Baetens, 2009). In particular, several studies report the involvement of the so-called mentalizing network when people are asked to explicitly reflect on the mental states of another person (e.g. Brass *et al.*, 2007; De Lange *et al.*, 2008).

Yet, whether, and how, these two networks divide up the labor of forming social inferences is disputed (for discussion see e.g. de Keysers and Gazzola, 2007; De Lange *et al.*, 2008; Van Overwalle and Baetens,

Received 20 February 2014; Revised 21 July 2014; Accepted 14 August 2014
Advance Access publication 19 August 2014

*The contributions of the first and the second author to this paper are equal.

This work was funded by the Bernstein Computational Neuroscience Program of the German Federal Ministry of Education and Research BMBF (Grants 01GQ1001C, 01GQ0851), and by the Berlin School of Mind and Brain, Humboldt-Universität zu Berlin (German Research Foundation, Excellence Initiative GSC 86/3).

Correspondence should be addressed to Anna Kuhlen, Bernstein Center for Computational Neuroscience, Charité – Universitätsmedizin Berlin, Philippstraße 13, Haus 6, 10115 Berlin, Germany. E-mail: anna.kuhlen@bccn-berlin.de

2009; Brennan *et al.*, 2010; Zaki *et al.*, 2010; Waytz and Mitchell, 2011; Spunt and Lieberman, 2012). On the one hand, a meta-analysis of more than 200 fMRI studies (Van Overwalle and Baetens, 2009) concludes that the two networks are rarely concurrently active, suggesting that they serve independent functions (see also Zaki *et al.*, 2010; Waytz and Mitchell, 2011). On the other hand, some recent studies report the mutual involvement of mirroring and mentalizing areas during social inference making (e.g. Schippers *et al.*, 2009; Centelles *et al.*, 2011; Spunt *et al.*, 2011; Becchio *et al.*, 2012; Ciaramidaro *et al.*, 2013; Mainieri *et al.*, 2013), supporting a more integrative account of the two networks (e.g. Keysers and Gazzola, 2007; Spunt and Lieberman, 2012).

With the current study, we aim to elucidate the contribution of these two networks, mirroring and mentalizing, when judging how confident another person appears of their knowledge. Specifically, our study investigates whether the 'outcome' of a social inference process (i.e. judging a person as more or less confident) modulates the observed pattern of neural activity. Participants in our study viewed short video clips of people answering general knowledge questions and were instructed to judge on a scale from 1 to 7 how confident the people in the video appeared of their answers. Videos were played mute so that participants would form their judgments on the basis of visual nonverbal information only. A parametric analysis approach identifies those brain areas indicative for the outcome of the social inference process by investigating how neural activity correlates with participants' individual judgments.

METHODS

Participants

Twenty-three right-handed participants between the ages of 18 and 35 years (mean age 27.01) took part in the fMRI study. They were native speakers of German, had normal or corrected-to-normal vision and no reported history of neurological or psychiatric disorder. Three participants had to be excluded owing to motion artifacts, and two further participants were excluded owing to technical difficulties. In total, 18 participants entered the final analysis (11 female). All participants gave informed consent and were compensated with €10 for every hour they participated in the experiment. The study was approved by the local ethics committee of the Psychology Department of the Humboldt University of Berlin.

Stimuli

A corpus of naturalistic video recordings collected by Swerts and Krahmer (2005) was used for selecting our stimuli. This corpus consists of 20 Dutch persons who had been videotaped while responding to 40 general knowledge questions (e.g. 'What is the capital of Switzerland?'). In their answer, respondents spontaneously exhibited different degrees of confidence in their knowledge. From this corpus, we chose 14 individuals (eight females) who each contributed three video sequences in which they appeared highly confident of their knowledge (high FOAK), and three video sequences in which they appeared little confident of their knowledge (low FOAK). Thus, in total, 84 videos were used in the fMRI experiment, with 42 videos associated with high and 42 videos associated with low levels of FOAK.

The selection of video sequences was based on a behavioral pretest in which 20 German participants (10 females) rated 175 videos from the original corpus on a scale of 1 to 7 with respect to 'how confident the respondent appeared to be that their answer was correct' (7 = very confident). Video sequences that, on average, elicited ratings lower than the median split (ratings < 4) were associated with judgments of low FOAK, and videos that elicited ratings higher than the median split (ratings > 4) were associated with high FOAK. Videos that elicited judgments equaling the median, as well as videos that elicited inconsistent judgments (ratings with standard deviations higher than two standard deviations of the average deviation across all videos), were excluded. Please note that our primary goal was to select videos that would reliably elicit judgments of high or low FOAK in our participants. To that end our selection criterion did not require a close correspondence between the respondent's own confidence and observer-rated confidence. In other words, even if participants misjudged the respondent's confidence, this would not be problematic, as our study is concerned with the subjective experience of inferring another person's confidence. Following this rationale, an influence of possible cultural differences between our Dutch respondents and our German participants in expressing or interpreting nonverbal cues is avoided. Participants who served in the pretest did not participate in the main fMRI study.

Each video sequence captured only the answer phase of the respondent, and was edited to be of 2.5 s length, with 15 video frames preceding the respondent's lexical response. Any information on the context of the question was excluded so it would not interfere with the evaluation of the answer (see procedure Brennan and Williams, 1995; Swerts and Krahmer, 2005). To allow only nonverbal information to influence participants' judgments, videos were presented without sound.

Design and procedure

The fMRI session was divided into six functional runs. Each run consisted of 14 trials during which one video sequence was presented and rated on a 7-point scale by the participant (see Figure 1). The order in which the stimuli were presented was randomized across all runs, with the following restrictions: in each run, half of the videos were associated with high, the other half with low, FOAK judgments. A maximum of three videos of one condition was presented in a row. Each character appeared once within each run. Stimuli were presented and responses recorded using MATLAB 7.1 (The MathWorks, Inc.) in combination with the Cogent toolbox (http://www.vislab.ucl.ac.uk/

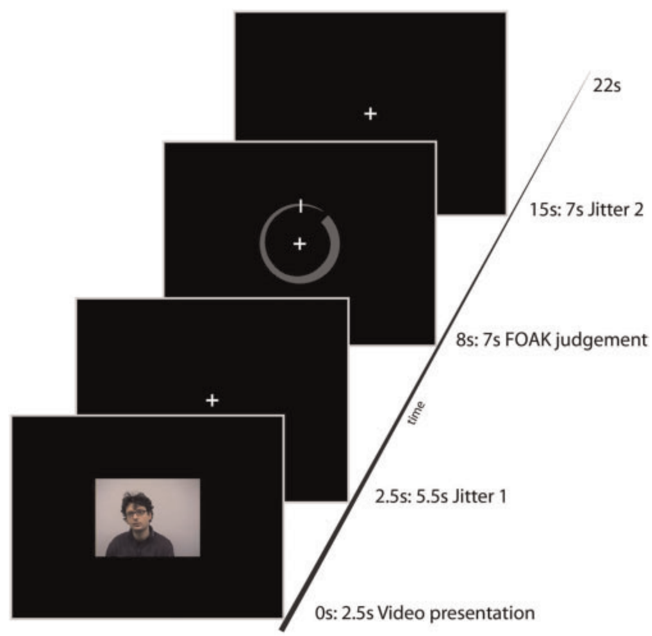

Fig. 1 A trial of the experiment. After the presentation of a video of 2.5 s length and a variable delay with a mean of 5.5 s (4.5, 5.5 or 6.5 s), participants had 7 s to perform the FOAK rating. Before the onset of the next trial, there was another variable delay with a mean duration of 7 s (8, 7 or 6 s). The length of the first and second delay in a trial always added up to 12.5 s. Taken together, a trial lasted 22 s.

Cogent). Videos were projected onto a screen (1024×768 pixel, 60 Hz) from the head-end of the scanner.

After a video was presented, participants rated how confident the respondent appeared to be of their answer (parallel to procedure in pretest). To avoid possible confounds due to covert motor response preparation, participants indicated their judgment on a circular rating scale with seven increments (Kahnt *et al.*, 2011). Instead of numbers, the line weight of the scale indicated how confident the respondent was perceived to be (very strong = very confident). The orientation and the starting position of the scale was randomized for each trial, making it impossible for participants to know in advance how they would need to adjust the scale to indicate their judgment. Participants gave their response with an MR-compatible response device. Two buttons rotated the rating scale clockwise or counterclockwise. The position of the rating scale after 7 s was taken as the final rating. Participants were trained in how to give their responses before scanning.

A jitter of varying length was inserted between video presentation and rating (4.5, 5.5 and 6.5 s; mean 5.5 s), and between rating and the beginning of the next block (8, 7 or 6 s) (see Figure 1). The length of the first and second delay in a trial always added up to 12.5 s. Taken together, one trial was 22 s long. The entire run was 314 s long ($14 \times 22\,s + 3$ volumes before the first trial started).

fMRI acquisition

Gradient-echo EPI functional MRI volumes were acquired with a Siemens TRIO 3 T scanner with standard head coil (33 slices, TR = 2000 ms, echo time TE = 30 ms, resolution $3 \times 3 \times 3$ mm^3 with 0.75 mm gap, Field of View (FoV) 192×192 mm). In each run, 157 images were acquired for each participant. The first three images were discarded to allow for magnetic saturation effects. For every subject, six runs of functional MRI were acquired. We also acquired structural MRI data (T1-weighted MPRAGE: 192 sagittal slices, TR = 1900 ms, TE =2.52 ms, flip angle = 9°, FoV = 256 mm).

fMRI preprocessing and analysis

Data were preprocessed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). The functional images were slice-time corrected with reference to the first recorded slice, motion corrected and then spatially smoothed with a Gaussian kernel of 6 mm FWHM. Data were highpass filtered with a cutoff period of 128 s.

The onsets of the videos and the individual FOAK ratings were used to model the fMRI data. We applied a general linear model (Friston *et al.*, 1994) with finite-impulse-response regressors (Henson, 2003) to the data of each run. Eight consecutive time bins were estimated for the onset of the videos and for the parametric modulation of the individual FOAK rating. Each time bin represents an MRI volume of 2 s and the eight time bins represent 16 s in total. Taken together, 16 regressors were estimated (8 time bins $\times$ 2 regressors; onset and parametric modulation). The confidence ratings were not included in the model. However, including the ratings resulted in extremely similar results for both the mean response and the parametric modulation.

The resultant contrast maps for the 16 regressors were normalized to a standard stereotaxic space (Montreal Neurological Institute) and resampled to an isotropic spatial resolution of $3 \times 3 \times 3$ mm^3. A random effects general linear model (Analysis of Variance (ANOVA)) was estimated for the 16 regressors across subjects, and final results were estimated using t-contrasts over those time bins during which the strongest response was expected given the hemodynamic delay (time bin 2, 3 and 4, i.e. between 2 and 8 s after stimulus onset for both the onset of the videos and the parametric modulation).

In our first analysis, we searched for brain regions that showed a mean response to the presentation of the videos that was larger compared with the implicit baseline condition. In a second analysis, we searched for brain regions that showed positive or negative parametric modulations corresponding to the individual FOAK ratings. Both analyses were performed as whole-brain analysis. To focus our results on the hypothesized involvement of the mirroring and mentalizing network, we present results for regions of interest (ROI). Please note that these ROIs merely serve as a mask to our results. All analyses were performed on the whole brain. All reported results were FWE corrected on the whole brain, unless otherwise stated.

ROI definition

ROIs were defined based on the Talairach coordinates of brain areas associated with the mentalizing and the mirroring network as reported in the meta-analysis by Van Overwalle and Baetens (2009). The mentalizing network included the TPJ (±50 −55 25), the mPFC (0 50 20) and the PC (0 −60 40); the mirroring network included the pSTS (±50 −55 10), the aIPS (±40 −40 45) and the PMC (±40 5 40). Around the reported center coordinates, a sphere with a radius of 9 mm was created using WFU PickAtlas for SPM8 (Maldjian *et al.*, 2003, 2004). WFU Pickatlas was also used to convert Talairach coordinates to MNI coordinates.

RESULTS

Behavioral results

As expected, based on the results of our pretest, participants rated respondents in videos associated with high FOAK as more confident of their knowledge ($M = 4.48$; SEM = 0.12) than respondents in videos associated with low FOAK ($M = 3.96$; SEM = 0.11), $t(17) = 9.21$, $P < .0001$.

Neuroimaging results

Brain regions activated by video sequences (mean response)

Brain regions activated by the presentation of the videos in general (independent of the outcome of the FOAK judgment) included the striate and extrastriate visual cortex, including the fusiform gyrus, the amygdala, the premotor cortex, bilateral pSTS, the bilateral TPJ and the mPFC, see Figure 2 (designated in blue) and Table 1.

Brain regions modulated by FOAK judgment

Cortical responses in the TPJ, left caudate nucleus, left inferior frontal gyrus and the mPFC were negatively modulated by participants' individual judgments about the respondents' perceived confidence (FOAK). These regions were thus more active when participants judged the respondents to be less confident of their knowledge (lower FOAK scores). Figure 2 (designated in red) visualizes the results of this univariate parametric analysis. The complete list of significantly activated clusters can be found in Table 2. There was no significant positive modulation.

ROI analysis

When we constrained the whole-brain fMRI analysis with the ROIs for the mentalizing and the mirroring network (see Figure 3), we found significantly activated voxels for the mean response in both the mirroring (pSTS and PMC) and the mentalizing network (PC, TPJ and mPFC). However, exclusively the mentalizing network (TPJ and PMC) showed a representation of the negative parametric modulation relative to the FOAK ratings (see also Table 3).

DISCUSSION

Inferring a conversational partner's knowledge state is essential for successful communication (Clark and Marshall, 1981; Krauss and

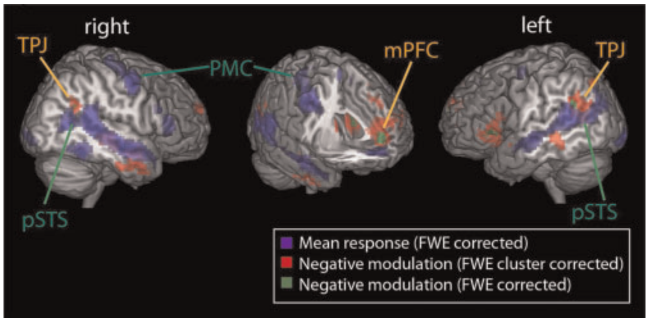

Fig. 2 Results of the fMRI analyses. Increased activity for the presentation of the videos compared with the implicit baseline was found in large networks, including temporal lobe (pSTS, middle temporal gyrus), TPJ, early visual cortex, fusiform gyrus, amygdala, mPFC and ACC (blue, $P < 0.05$ FWE corrected; see Table 1 for a complete list). A negative parametric modulation for the FOAK ratings was identified in TPJ, mPFC, middle temporal gyrus, caudate nucleus and left inferior frontal gyrus (green, $P < 0.05$ FWE corrected; see Table 2 for a complete list; red, $P < 0.05$ FWE corrected on the cluster level for illustrative purposes). Please note the more superior cluster of voxels in the bilateral TPJ that are modulated by the FOAK ratings compared with the more inferior cluster of voxels of the pSTS that only show an increased mean response to the presentation of the videos.

Table 1 Main effect for the presentation of the videos compared with the implicit baseline condition

Anatomical area	L/R	T-value	x	y	z
Fusiform gyrus	R	14.61	39	−55	−14
	L	15.94	−39	−46	−17
Amygdala	R	9.93	21	−7	−14
	L	8.17	−21	−10	−14
Middle temporal gyrus	R	18.29	54	−40	10
	L	14.51	−54	−46	13
	L	11.58	−54	−7	−11
Superior temporal gyrus	R	15.69	66	−40	16
	R	15.27	51	−22	−5
Early visual cortex	R	10.32	27	−94	−5
	L	14.52	−24	−97	−8
Inferior frontal gyrus	L	5.26	−39	29	1
	L	4.97	−27	26	1
	L	12.99	−42	14	22
	R	4.89	27	29	−11
	L	4.89	−36	32	−11
	R	15.91	42	17	22
	R	9.65	54	32	7
Superior frontal gyrus	L	5.51	−27	38	49
Supplementary motor area	R	9.42	6	5	64
	R	6.73	3	−25	61
	L	6.06	−9	8	55
Precentral gyrus	R	13.97	51	−1	46
	L	7.22	−48	−4	46
Rectal gyrus	R	11.91	3	41	−14
Middle orbital gyrus		5.85	0	62	1
Superior medial gyrus	R	8.6	6	53	25
Thalamus		6.94	3	−13	4
Postcentral gyrus	R	5.59	15	−40	76

The coordinates are given according to MNI space with their T-values. L = left hemisphere, R = right hemisphere; all regions FWE corrected $P < 0.05$.

Fussell, 1991). How confident the partner appears in what they are saying provides relevant information on how to interpret and react to their utterances. Our study investigates the neural processing of nonverbal cues for another person's perceived confidence. Our imaging data specify the role of two neural networks associated with social inference making, the mentalizing and the mirroring network: While areas of both networks are active when processing the nonverbal cues

Table 2 Negative parametric modulation relative to the individual FOAK ratings

Anatomical area	L/R	T-value	x	y	z
Inferior frontal gyrus	L	6.35	−42	26	1
	L	5.96	−33	20	−14
Anterior cingulate cortex	R	5.58	3	44	19
	L	6.11	−6	50	4
	R	4.85	6	38	7
Superior medial gyrus	R	5.35	6	62	22
	L	5.4	−9	62	22
Middle temporal gyrus	R	5.36	57	−19	−8
	L	5.03	−54	−31	−5
	L	5.01	−57	−61	16
	L	4.97	−69	−37	−8
Medial temporal pole	R	5.03	54	2	−26
Inferior temporal gyrus	L	5.04	−66	−46	−11
Temporoparietal junction	R	4.98	45	−55	25
	L	6.04	−60	−49	25
Caudate nucleus	L	5.67	−9	5	13

The coordinates are given according to MNI space with their T-values. L = left hemisphere, R = right hemisphere; all $P < 0.05$ FWE corrected.

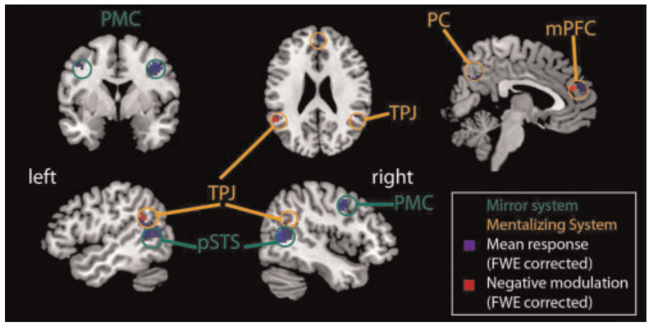

Fig. 3 Results of the fMRI analyses masked with region of interests for the mirroring (green: pSTS, PMC) and the mentalizing network (orange: TPJ, mPFC). Increased activity for the presentation of the videos was found in both the mirroring and the mentalizing network (blue). However, a negative parametric modulation for the FOAK ratings was identified explicitly in TPJ and mPFC (red) ($P < 0.05$ FWE corrected on the whole brain and masked with the ROIs).

of another person's confidence, only activity in the mentalizing network is modulated by the outcome of this judgment.

The less confident the respondent is judged to be the more active are core areas of the mentalizing network, the bilateral TPJ and the mPFC. This confirms the central role these two areas play in reflective mentalizing about another person's mind and higher order social cognition (e.g. Saxe and Kanwisher, 2003; Amodio and Frith, 2006). We observe no evidence for a modulation of the mirroring network in response to the participants' judgment. Yet, we do find activation of areas typically associated with the mirroring network, namely, the PMC and the bilateral pSTS, during the presentation of the nonverbal behavior (independent of how they were judged). This is consistent with studies that report activity in the mirroring network, in particular the right pSTS, when processing facial or bodily displays (compared with verbal information; Zaki *et al.*, 2010; Kuzmanovic *et al.*, 2012).

Our data support proposals for an integrative account of the mentalizing and mirroring network, in which the two networks work together, yet serve dissociable functions (Keysers and Gazzola, 2007; Lieberman, 2007; De Lange *et al.*, 2008; Olsson and Ochsner, 2008; Thioux *et al.*, 2008; Spunt *et al.*, 2011; Spunt and Lieberman, 2012): While mirroring assists the encoding of observable motor behavior, mentalizing assists the interpretation of this behavior with respect to underlying mental states. Along these lines, a study by Spunt and

Table 3 Main effect for the presentation of the videos compared with the implicit baseline condition, and negative parametric modulation relative to the individual FOAK ratings

Region of interest	L/R	Mean activation				Negative parametric modulation			
		T-value	x	y	z	T-value	x	y	z
pSTS	R	13.29	51	−64	10				
	L	13.61	−54	−49	13				
PMC	R	13.84	48	−1	46				
	L	7.13	−48	−1	49				
aIPS	R								
	L								
TPJ	R	10.78	51	−64	16	4.98	45	−55	25
	L	10.42	−54	−49	19	5.82	−57	−49	25
PC		6.34	0	−70	31				
mPFC		8.6	6	53	25	5.58	3	44	19

The results are masked by the ROIs of the mirroring and mentalizing network. The coordinates are given according to MNI space with their T-values. Corresponding rows for the ROIs in which no significantly activated voxels were found are empty. L = left hemisphere, R = right hemisphere; all $P < 0.05$ FWE corrected. Mirroring: pSTS = posterior superior temporal sulcus, PMC = promotor cortex, aIPS = anterior intraparietal sulcus; Mentalizing: TPJ = temporoparietal junction, PC = precuneus, mPFC = medial prefrontal cortex.

Lieberman (2012) reports the involvement of the mirroring network when observers identify the motor behavior of an affective facial display ('how' a person is showing their feelings); and the involvement of the mentalizing network when observers attribute meaning to this display ('why' the person is showing their feelings). The pattern of activity we find is in line with such a division of labor between the mirroring and mentalizing network.

We observed a differential activation of two brain regions that often lack a clear anatomical and functional distinction, the pSTS and the TPJ (for discussion see Saxe, 2006; Van Overwalle and Baetens, 2009). In our study, a brain region typically labeled as pSTS (our identification of both regions is based on the coordinates reported in the review by Van Overwalle and Baetens, 2009) was activated during the general presentation of the video, while a region typically labeled TPJ was associated with the outcome of the FOAK judgment. This differential pattern of activation is in line with a proposed functional distinction by Saxe and colleagues (Saxe and Kanwisher, 2003; Saxe, 2006), which associates activity in the pSTS with observing and interpreting body motion, and activity in the TPJ with reasoning in more abstract terms about the affective and cognitive states of others. As the pSTS has been associated with the mirroring network and the TPJ with the mentalizing network (Van Overwalle and Baetens, 2009; but see e.g. Amodio and Frith, 2006), this differential pattern of activation is in line with our interpretation that mirroring supports the general processing of nonverbal FOAK cues, while mentalizing supports judging them.

The increased activation of the mPFC and the TPJ when respondents were judged to be less confident may indicate that nonverbal cues for low confidence are marked compared with cues for high confidence. Indeed, cues for high confidence are often described by a lack of cues for low confidence (see e.g. Smith and Clark, 1993; Swerts and Krahmer, 2005). The increased recruitment of the mPFC and the TPJ, both areas associated with the mentalizing network, may suggest an increased use of mentalizing when processing displays of low confidence. Along these lines, mentalizing, and in particular activity in the mPFC and the TPJ, has been associated with observing actions that are faulty or unusual in a given context (Van Overwalle and Baetens, 2009). Possibly displaying low confidence when answering a question is more out of order than being confident, and hence, triggers more pronounced mentalizing in the observer.

High and low levels of confidence are associated with different types of nonverbal behaviors (Swerts and Krahmer, 2005). For example, cues for low confidence are diverted gaze, and, conversely, for high confidence presumably direct gaze. Imposing experimental control on these cues would take away precisely the type of information underlying FOAK judgments. Yet, systematic differences in visual cues may elicit distinct patterns of neural activation simply due to processing low-level perceptual features. In other words, neural activation that we assigned to making judgments of confidence may, in fact, be triggered by a characteristic type of visual cue.

Observing another person's gaze has, for example, been associated with activity in the pSTS (Kuzmanovic *et al.*, 2009; Ethofer *et al.*, 2011). It is therefore possible that some of the brain activity found when observing FOAK cues might reflect the processing of respondents' gaze displayed in the videos. Yet, the activation of the pSTS that we find is not related to the outcome of the FOAK judgment. Likewise, other types of stimulus-induced differences (e.g. potentially more general characteristics in motor movement) would elicit systematic activation patterns in early visual areas, which we do not find.

A second related concern may be that different cues may cause participants to attend systematically to different regions in the displayed video, thus resulting in distinct patterns of eye movement while watching the videos. Yet again we do not find a modulation of activity in areas associated with oculomotor processes or visuospatial attention (overt or covert)—for example, the precentral sulcus, intraparietal sulcus or lateral occipital cortex (Corbetta *et al.*, 1998; Beauchamp *et al.*, 2001)—that is related with the outcome of the judgment. A confounding influence of processing low-level visual properties of the stimulus videos on the activation pattern associated with judging another person's confidence is therefore unlikely.

Lastly, other higher-level features of our stimulus videos may have coincided with expressions of high or low confidence. Notably, the emotional expression of our respondents may have corresponded with their levels of confidence (e.g. happy expression if confident). To address this concern, we asked seven volunteers to assess the emotional valence respondents displayed in the videos using a 7-point scale version of the Self-Assessment Manikin (Lang, 1980). There was no correlation between perceived valence and perceived confidence on a group level (t-test on Fisher-Z normalized correlation coefficient $t(6) = 1.16$, $P = 0.29$), and also none on the level of individual participants ($r = 0.04$, $n = 84$, $P = 0.7$; $r = 0.05$, $n = 84$, $P = 0.64$; $r = 0.02$, $n = 84$, $P = 0.83$; $r = 0.19$, $n = 84$, $P = 0.08$; $r = -0.09$, $n = 84$, $P = 0.39$; $r = -0.03$, $n = 84$, $P = 0.81$; $r = 0.09$, $n = 84$, $P = 0.43$). The neural activation we have associated with judging a speaker's confidence is therefore unlikely to be confounded by participants judging the speaker's emotional expression.

Nonverbal cues make an important contribution to the pragmatic interpretation of utterances: the manner is which an utterance is produced can reveal how committed to, or confident, a speaker is about what they are saying (Smith and Clark, 1993; Brennan and Williams, 1995; Swerts and Krahmer, 2005). This study makes a first step toward investigating the neural circuitry underlying the metacognitive process of interpreting another person's displayed confidence. Our results show that the processing of these nonverbal cues recruits neural circuits associated with imitation, while the interpretation of the cues recruits neural circuits associated with reflective mentalizing. Our study highlights the important contribution social inferences about the conversational partner can make to pragmatic processing, and identifies core neural circuits that may underlie these processes. Specifically how the outcome of such social inferences informs language processing is an important question for future research.

Conflict of Interest

None declared.

REFERENCES

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277.

Beauchamp, M.S., Petit, L., Ellmore, T.M., Ingeholm, J., Haxby, J.V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *Neuroimage*, *14*(2), 310–321.

Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., Castiello, U. (2012). Social grasping: from mirroring to mentalizing. *Neuroimage*, *61*(1), 240–248.

Brennan, S.E., Galati, A., Kuhlen, A.K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In: Ross, B., editor. *The Psychology of Learning and Motivation*, Vol. 53, Burlington, VT: Academic Press, pp. 301–344.

Brennan, S.E., Williams, M. (1995). The feeling of Another's Knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*(3), 383–398.

Centelles, L., Assaiante, C., Nazarian, B., Anton, J.-L., Schmitz, C. (2011). Recruitment of both the mirror and the mentalizing networks when observing social interactions depicted by point-lights: a neuroimaging study. *PLoS One*, *6*(1), e15749.

Ciaramidaro, A., Becchio, C., Colle, L., Bara, B.G., Walter, H. (2013). Do you mean me? Communicative intentions recruit the mirror and the mentalizing system. *Social Cognitive and Affect Neuroscience*, *9*(7), 909–916.

Clark, H.H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H.H., Marshall, C.R. (1981). Definite reference and mutual knowledge. In: Joshi, A.K., Webber, B., Sag, I., editors. *Elements of Discourse Understanding*. Cambridge: Cambridge University Press, pp. 10–63.

Corbetta, M., Akbudak, E., Conturo, T.E., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773.

De Lange, F.P., Spronk, M., Willems, R.M., Toni, I., Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current Biology*, *18*(6), 454–457.

Ethofer, T., Gschwind, M., Vuilleumier, P. (2011). Processing social aspects of human gaze: a combined fMRI-DTI study. *Neuroimage*, *55*(1), 411–419.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, *2*(4), 189–210.

Frith, C.D., Frith, U. (2012). Mechanisms of social cognition. *Annual Review Psychology*, *63*, 287–313.

Gallagher, H.L., Frith, C.D. (2003). Functional imaging of "theory of mind.". *Trends in Cognitive Sciences*, *7*(2), 77–83.

Gallese, V., Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Hart, J.T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–16.

Henson, R.N.A. (2003). Analysis of fMRI time series. In: Frackowiak, R.S.J., Friston, K.J., Frith, C. et al., editors. *Human Brain Function,* 2nd edn. San Diego: Academic Press.

Kahnt, T., Heinzle, J., Park, S.Q., Haynes, J.-D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage*, *56*(2), 709–715.

Keysers, C., Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, *11*(5), 194–196.

Krauss, R.M., Fussell, S.R. (1991). Perspective-taking in communication: representations of others' knowledge in reference. *Social Cognition*, *9*(1), 2–24.

Kuzmanovic, B., Bente, G., von Cramon, D.Y., Schilbach, L., Tittgemeyer, M., Vogeley, K. (2012). Imaging first impressions: distinct neural processing of verbal and nonverbal social information. *Neuroimage*, *60*(1), 179–188.

Kuzmanovic, B., Georgescu, A.L., Eickhoff, S.B., et al. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, *46*(4), 1154–1163.

Lieberman, M.D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Reviews Psychology*, *58*, 259–289.

Mainieri, A.G., Heim, S., Straube, B., Binkofski, F., Kircher, T. (2013). Differential role of the Mentalizing and the Mirror Neuron system in the imitation of communicative gestures. *Neuroimage*, *81*, 294–305.

Maldjian, J.A., Laurienti, P.J., Burdette, J.H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage*, *21*(1), 450–455.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*(3), 1233–1239.

Olsson, A., Ochsner, K.N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*(2), 65–71.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion Neurobiology*, *16*(2), 235–239.

Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage*, *19*(4), 1835–1842.

Schippers, M.B., Gazzola, V., Goebel, R., Keysers, C. (2009). Playing charades in the fMRI: are mirror and/or mentalizing areas involved in gestural communication? *PLoS One*, *4*(8), e6801.

Smith, V.L., Clark, H.H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*(1), 25 – 38.

Spunt, R.P., Lieberman, M.D. (2012). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *Neuroimage*, *59*(3), 3050–3059.

Spunt, R.P., Satpute, A.B., Lieberman, M.D. (2011). Identifying the what, why, and how of an observed action: an fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, *23*(1), 63–74.

Swerts, M., Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, *53*(1), 81–94.

Thioux, M., Gazzola, V., Keysers, C. (2008). Action understanding: how, what and why. *Current Biology*, *18*(10), R431–R434.

Van Overwalle, F., Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, *48*(3), 564–584.

Waytz, A., Mitchell, J.P. (2011). Two mechanisms for simulating other minds dissociations between mirroring and self-projection. *Current Directions in Psychological Science*, *20*(3), 197–200.

Zaki, J., Hennigan, K., Weber, J., Ochsner, K.N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *Journal of Neuroscience*, *30*(25), 8481–8488.