Special issue: Research report

Neural correlates of observing joint actions with shared intentions

Terry Eskenazi [a,b,*], *Shirley-Ann Rueschemeyer* [c], *Floris P. de Lange* [a], *Günther Knoblich* [a,d] *and Natalie Sebanz* [d]

[a] *Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen, The Netherlands*
[b] *LNC, Ecole Normale Supérieure, Paris, France*
[c] *University of York, York, UK*
[d] *Central European University, Budapest, Hungary*

ABSTRACT

Studies on the neural bases of action perception have largely focused on the perception of individual actions. Little is known about perception of joint actions where two or more individuals coordinate their actions based on a shared intention. In this fMRI study we asked whether observing situations where two individuals act on a shared intention elicits a different neural response than observing situations where individuals act on their independent parallel intentions. We compared the neural response to perceptually identical yet intentionally ambiguous actions observed in varying contexts. A dialog between two individuals conveyed either a shared intention or two independent parallel intentions. The dialogs were followed by an identical video clip where the two individuals performed certain actions. In one task condition participants tracked the intentions of the actors, in the other, they monitored moving colored dots placed on the same videos. We found that in the intention task versus the color task, observing joint actions based on shared intentions activated the temporal poles, precuneus, and the ventral striatum compared to observing interactions based on parallel intentions. Precuneus and the temporal poles are thought to support mental state reasoning, the latter with a more specific role in retrieving memories associated with social scripts. Activation in the ventral striatum, an area involved in reward processing, likely indicates a hedonistic response to observed shared intentional relations similarly to those experienced when personally sharing mental states with others.

© 2015 Elsevier Ltd. All rights reserved.

To navigate the social world we need to make sense of individual actions as well as of actions performed jointly by others. To date, research has largely focused on the processes underlying the perception of others' individual actions.

However, from a couple taking a walk to a chamber orchestra performing a musical piece, social life abounds with examples of joint actions. Much is still unknown about the exact neural processes underlying the perception of joint actions.

* Corresponding author. Laboratoire de Neurosciences Cognitives U940 INSERM, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France.
E-mail address: terry.eskenazi@ens.fr (T. Eskenazi).

Broadly defined, joint actions are social interactions where two or more individuals spatially and temporally coordinate their actions to bring about a change in the environment (Sebanz, Bekkering & Knoblich, 2006). There is evidence suggesting that we are attentionally and perceptually tuned in to the joint actions we encounter. For instance, we can detect contingencies between movement trajectories of different agents, or arbitrary objects (for a review Scholl & Tremoulet, 2000), and distinguish joint actions from situations where multiple people act independently even in perceptually impoverished contexts (Centelles, Assaiante, Nazarian, Anton, & Schmitz, 2011; Manera Becchio, Schouten, Bara & Verfaille, 2011). Also, observing two people sharing attention engages our own attention more so than when people are not in any apparent interaction (Boeckler, Knoblich & Sebanz, 2011). This enhanced level of attention in turn can facilitate learning, for example, of new words (Akhtar, 2005; O'Doherty et al., 2011) or of new actions (Herold & Akhtar, 2008; Nielsen, Moore & Mohhammedaly, 2012).

Further, our propensity to detect contingencies between actions of individuals can elicit a tendency to ascribe mental states to the agents that relate them to each other (Ip, Chiu & Wan, 2006; Lakens, 2010; Lakens & Stel, 2011). For instance, closer rapport (Miles, Nind & Macrae, 2009) and a sense of unity ('entitativity', Lakens & Stel, 2011) is attributed to individuals who act in synchrony. In particular, we tend to think that if agents move together, they must be doing so intentionally (Ip et al., 2006; Lakens, 2010). We seem to be able to recognize, as early as at 14 months, when individuals engage in a complex intentional interaction and work together to achieve a shared goal (Henderson & Woodward, 2011). Once interacting individuals are perceived as a unity, they may sometimes be attributed a group-mind, allowing the observer to explain the behavior of the group in terms of group-level mental states (Waytz & Young, 2012). In addition, once a social situation has been identified as an instance of joint action, the relational mental states the observer ascribes to the observed individuals allow for successful predictions regarding how the interaction will unfold (Manera et al., 2011; Sartori, Becchio & Castiello, 2011). These studies suggest that we are not only perceptually sensitive to social interactions we observe, but we also have a particular way of reasoning about the mental states of those we observe interacting.

To corroborate these findings, recent neuroimaging studies have shown that the brain network subserving the ability to attribute mental states to individuals also supports observation of social interactions. This so-called mentalizing, or theory-of-mind, network—comprising the medial prefrontal cortex (mPFC), the temporo-parietal junction (TPJ), the temporal poles and the posterior cingulate (PCC) and/or the precuneus—is sensitive to processing mental states of individuals across a variety of situations, including reading vignettes or static cartoons, watching dynamic displays of characters (for reviews, see Frith & Frith, 2010; Gallagher & Frith, 2003). Robust responses in these regions are also noted when individuals are immersed in online interactions, such as playing online interactive games with another human compared to a computer (Gallagher, Jack, Roepstorff & Frith, 2002; McCabe, Houser, Ryan & Trouard, 2011; Rilling et al., 2002; Rilling, Sanfey, Aronson, Nystrom & Cohen, 2004).

The mentalizing network has also been implicated in tasks that involve passive observation of others' social interactions. Not only watching dynamic displays of contingent actions performed by arbitrary shapes (Castelli, Happe, Frith & Frith, 2000), but also attending to complex social scenes that depict multiple agents (Kujala, Carlson, & Hari, 2012; Pierno, Becchio, Turella, Tubaldi & Castiello, 2008; Sinke, Sorger, Goebel & de Gelder, 2010; Spiers & Maguire, 2006; Wagner, Kelley & Heatherton, 2011) recruits the mentalizing network. Particularly, compared to observing a single individual's actions observing an interaction (Becchio et al., 2012; Iacoboni, Lieberman & Knowlton, 2004) and distinguishing point-light displays of communicative exchanges between individuals from displays of two agents acting independently also elicits activation in these regions (Centelles et al., 2011). Together these findings suggest that the mentalizing network serves perception of social interactions by supporting the observer in ascribing mental states to the observed agents, and in generating inferences about the nature of their interaction.

While behavioral cues can facilitate predictions and inferences about observed actions, identical actions may be based on different prior intentions (Jacob & Jeannerod, 2005). That is, one and the same action — whether an individual or a jointly performed one — may reflect various different underlying intentions. Many accounts of both individual and joint action therefore emphasize the role of intention in distinguishing between different kinds of actions. In fact most philosophers, although disagreeing on the precise nature of shared intentions, believe that shared intentionality is what characterizes joint actions (Bratman, 1992; Searle, 1990).

Accordingly, when individuals act together with a shared intention, this constitutes a different instance than when individuals act in parallel with independent individual intentions. To illustrate the distinction Searle offers the following paradigm case (1990). Imagine a scenario where people are scattered across a park sunbathing. Suddenly rain starts pouring down and people run towards the nearest shelter. Here each individual acts on an intention that is independent from the intentions of others. Imagine a second scenario where the same park is used as a film set and the same people are actually actors. Fake rain starts pouring down and the actors run for the nearest shelter, in the same way as in the first scenario. However, here the individuals have a shared intention (i.e., to get the scene shot), which may involve the interlocking of their individual intentions (Bratman, 1992). The actions performed in these two scenarios are ostensibly identical, however, the context in which they occur distinguishes them in terms of the intentional relations between those involved.

In the present study we asked which neural regions are sensitive to observing individuals acting on shared intentions compared to observing individuals acting on independent intentions that run in parallel. To that end we showed participants perceptually identical video clips depicting two people engaged in action. These clips were ambiguous in terms of the intentionality of the people observed. We embedded the clips in varying contexts that were established by dialogs presented prior to the action videos. In one case the dialog conveyed a shared intention where the actors agreed on doing something together (i.e., shared intention). In the

other case, actors expressed independent individual intentions (i.e., parallel intentions). We compared the neural response to the action videos in the Shared Intention condition to the neural response to the Parallel Intentions condition. In the following we outline three specific aims of our study.

Our first aim was to test whether observing actions performed with a shared intention, where the observed individuals' intentions are 'interlocked' (Bratman, 1992), differentially engages the mentalizing network compared to observing the same actions performed with independent intentions. Neuroimaging evidence sketched above demonstrates that the theory of mind network is sensitive to observing social interactions, when the task concerns attending or identifying the nature of the interaction taking place. This implies something about tracking intentional relations between the observed agents (Centelles et al., 2011; Iacoboni et al., 2004; Kujala et al., 2012; Sinke et al., 2010). Particularly the mPFC node of this network seems to be sensitive to inferring social intentions, that is when an individual's intention references another person, either when it's directed at the observer himself (Grezes, Frith, & Passingham, 2004) or at a third individual (Centelles et al., 2011; Ciaramidaro et al., 2007; Iacoboni et al., 2004; Kujala et al., 2012; Sinke et al., 2010; Walter et al., 2004). Accordingly, if processing intentional relations involves mentalizing related computations, then observing individuals act with a shared intention should yield increased activation in the mPFC compared to when they act independently.

Our second aim was to determine whether similar neural processes are involved when observers explicitly and implicitly keep track of intentional relations behind observed actions. There is behavioral evidence to suggest that mental or perceptual states of others are represented implicitly even when it is not required by the task (Kovacs, Teglass & Endress, 2010; Samson, Apperly, Braithwaite, Andrews & Bodley-Scott, 2010; van der Wel, Sebanz, & Knoblich, 2014). Converging evidence come from neuroimaging studies that demonstrated the involvement of the mentalizing network when participants spontaneously attributed mental sates to others (Schneider, Slaughter, Becker & Dux, 2014). A few studies that investigated unattended or spontaneous observation of social situations also implicated these cortical regions (Sinke et al., 2010; Wagner et al., 2011). We asked if implicit tracking of intentional relations in the context of joint action also recruits the mentalizing network. To address this question we included two different tasks using the same stimuli (i.e., Intention task and Color task). In the Intention task we asked participants to detect mismatches between the intentions the actors had expressed in the dialog events and what they subsequently did. This creates an explicit demand on monitoring the intentions underlying the actions observed. In the Color task participants were asked to detect a color change in dots superimposed on the videos. This only required monitoring visual features of the observed actions. The aim of this manipulation was not to make it impossible to explicitly track mental states in the color task, but to test whether the mentalizing network would be recruited to the same or to a lesser extent when explicit mental state tracking was not demanded by the task. If intentional relations are implicitly monitored

even when this is irrelevant for task performance, then we should find the same neural areas to be recruited in both the Intention and the Color tasks.

Finally, the third aim of the study was to determine whether observing joint actions based on shared intentions elicits a reward response in the observer, which would be reflected in activations in the reward-relevant neural network, particularly in the ventral striatum. Numerous studies have implicated this area when people are actively engaged in social interactions with others, for example when acting in synchrony (Miles, et al., 2009), engaging in direct eye gaze (Pfeiffer et al., 2014; Redcay et al., 2010; Schilbach et al., 2010) or when cooperating with others (Rilling et al., 2002). There has been much debate as to whether being engaged in social interaction is fundamentally different from observing others engaging in social interaction (de Jaegher, di Paolo, & Gallagher, 2010), or whether others' intentional relations to objects and other people can be easily mapped onto our own experiences (Barresi & Moore, 1996). The latter view rests on the assumption that humans have evolved the ability to integrate first and third person information about intentional actions (Barresi & Moore, 1996). In support of this view, it has been shown that observing others share eye gaze has similar effects on gaze following as being directly looked at oneself (Boeckler et al., 2011). Accordingly, it could be that observing others acting on the basis of shared intentions elicits similar affective responses as being directly engaged in social interactions that rely on shared intentions. This would be reflected in increased activation of the ventral striatum in the Shared Intention condition when compared with the Parallel Intentions condition.

1. Method

1.1. Participants

21 native speakers of Dutch (9 females; M age = 22, SD = 2.93) participated in this study and received monetary compensation. Out of those 21 participants the data from five were excluded from the analysis due to excessive head movements (>5 mm) or other technical difficulties at the time of data collection. All participants were right handed and had normal or corrected sight. The study was approved by the local ethics committee and complied with the Declaration of Helsinki.

1.2. Stimuli and design

Video clips of two people performing actions were embedded in varying contexts that framed the way the videos could be interpreted. Context was established by means of a dialog that took place between the actors, which was presented prior to the action video. In the Shared Intentions (SI) condition, two people were heard to agree on doing something together (e.g., A: Shall we set the table? B: Yes good idea!). This was followed by the action video where two actors were engaged in action. In the Parallel Intentions (PI) condition the identical action video was preceded by a different dialog, where actors express independent intentions (e.g., A: Shall we set the table? B: No I think I'll clean the cutlery). We added a third condition,

Individual Intention (*II*) condition, in which the dialog portrayed only one actor expressing an intention (e.g., A: What are you going to do? B: I think I'll clean the cutlery). Here, in the action video that followed only one of the actors engaged in an action and was passively observed by the other. We included two tasks in the paradigm that were blocked. In both tasks the participants were instructed to count 'mismatch trials' within each block. In the mismatch trials of the Intention Task the content of the action video did not match the content of the dialog video. In the Color Task, two colored dots that followed the actions of the actors were superimposed on the action videos that were presented in the Intention task. Here the task was to detect the trials where one of the dots changed color.

The stimuli consisted of two types of videos: dialog videos and action videos. We recorded the videos in an actual apartment for natural scenery. In action videos, two different mixed-gender pairs acted out five different scenarios: cleaning the living room, setting the dining table, repairing the TV, folding the laundry, and making coffee. For example, in the cleaning scenario one of the actors sprayed cleaning liquid on a dining table and wiped it with a paper towel. The other actor, standing at the other end of the table, removed some books from a bookshelf located behind the actor and placed them on the table. The actors did not engage in eye contact during the videos, and no verbal or non-verbal communication took place. In the dialog videos that were presented prior to the action videos, a still photograph of the room set-up was seen while the dialog between the actors was overheard. In the Shared Intentions dialog (SD), the actors agreed on cleaning the living room together. In the *PI* dialog (PD) one actor expressed the intention to clean the living room, while the other expressed the intention to search for a book. Additional videos were recorded for mismatch trials of the Intention task. For example in the cleaning scenario, following the dialog, the actors were seen playing cards in the same setting. We piloted the stimuli prior to the scanning session. In a two choice task participants were asked after each trial (dialog video + action video) whether the trial was a match or a mismatch. The difference in accuracy scores between mismatch and match trials was not significant (mismatch: M = 96%, SD = 3.6, match: M = 86%, SD = 15.7, $p > .05$).

The same videos were further processed to create the stimuli used for the Color task. The action videos were superimposed with two colored dots, one placed on each actor. These dots followed the actions of the actors. For the mismatch trials in the Color task one of the two dots changed color at a particular time point. The average time the participants required to categorize each video as a match or a mismatch in the pilot study determined the time point for the color change of the dot in the Color task. All videos were processed using Adobe Premiere Pro CS4. The sound files in the dialog videos were equalized using EQ Filtering 2.0 (www.sens.com) software to correct for the distortion created by MRI signal.

1.3. Procedure

Having provided informed consent, participants received instructions and were familiarized with the task in a short practice run before the scanning session. The practice task included different videos than the ones included in the experiment. The experiment which included the within-subjects factors Intention and Task, consisted of 126 trials in total, equally distributed across the three Intention conditions (Shared Intention, Parallel Intentions, Individual Intention). The Color task comprised 60 trials, whereas the Intention task comprised 66. The extra six trials in the Intention task were the mismatches, where the action videos had a different content than any of the five scenarios. The trial types (five scenarios, two actor pairs, and two actor locations) for each of the two tasks (Intention and Color) were randomized across the experiment. In one such trial type Actor A started the dialog and B responded, which was followed by the action video where A was seen on the left part of the screen and B on the right. In another trial type B started the dialog and was seen on the left part of the screen in the subsequent action video. Tasks were blocked, and blocks were presented in random order. Each block consisted of 10 (Color task) or 11 trials (Intention task). Each trial (Fig. 1) commenced with a fixation cross which was succeeded by the dialog video that lasted for approximately 3 sec. A jittered delay of 5–7 sec separated the dialog event from the action event. The
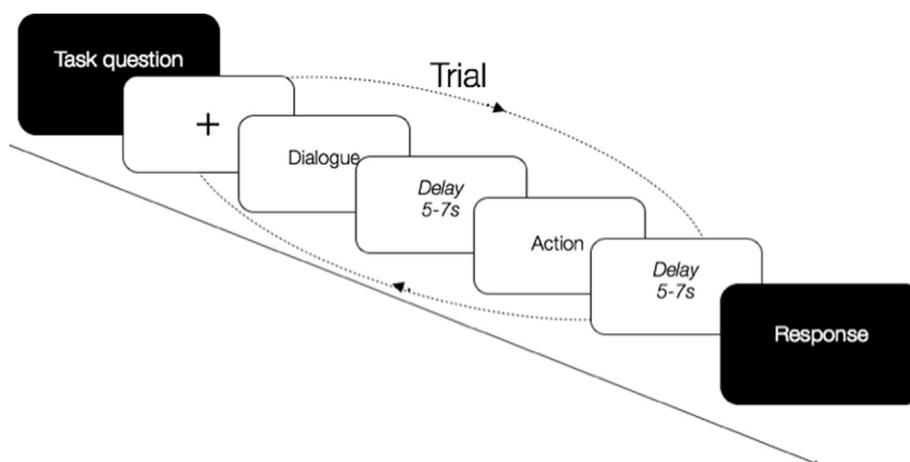

Fig. 1 – Schematic illustration of an example trial.

duration of the action events varied between 11 and 13 sec depending on the scenario. This was followed by a jittered inter-trial interval (ITI; 5−7 sec). The task question was presented in the beginning of each block. For the Intention task the question was *"In how many of the following videos did the actors not do what they had agreed to do?"*, and for the Color task the question was *"In how many of the following trials does one of the moving dots change color?"*. At the end of the block participants were prompted to provide a response by choosing one of the given 4 options (0, 1, 2 or 3) via the button-box placed under their right hands. Stimuli were presented with Presentation® software (www.neurobs.com).

1.4. Neuroimaging data acquisition

Participants lay in the scanner in supine position. Visual stimuli were presented on a computer screen outside of the scanner, which participants could see via mirror-glasses. Participants' heads rested within a 32-channel head array coil and was secured with pads and masking tape. T2-weighted echo-planar imaging blood-oxygenation-level-dependent (EPI-BOLD) fMRI measurements were acquired in 34 axial slices in ascending ordering (3 mm thickness, .5 mm slice gap) with a Siemens Trio 3T system using a 5 pulse multi-echo sequence (TE durations of 6.9,16.2, 25, 35, 44 msec, TR = 2.19 sec, flip angle = 80deg, isotropic voxel size = 3.5 mm, FoV = 224 mm). Following functional imaging T1-weighted images were obtained (TR = 2.3 sec, TE = 3.03 msec).

1.5. Neuroimaging data analysis

The fMRI data were analyzed in an event-related manner within the general linear model, using SPM8 software (Welcome Department of Imaging Neuroscience, London, UK) and Matlab R2007b (The MathWorks Inc., Natick, MA, USA). The 5 pulse multi-echo images were combined with an ad hoc Matlab script. Standard spatial pre-processing (spatial realignment, slice-time correction, normalization to MNI space and spatial smoothing with a 6 mm Gaussian kernel (Friston et al., 1995) was performed. For the statistical analysis, the dialog and action events were modeled as the events of interest for each condition. A jittered null event (5−7 sec) placed between the dialog and action events allowed the two events to be analyzed separately. 12 regressors were created for the six conditions (three action conditions × two tasks) with two events in each condition (dialog and action). Mismatch trials were modeled as separate regressors and were not included in further analyses. All regressors were convolved with a canonical hemodynamic response function (HRF). To correct for motion-related artifacts, we modeled subject-specific realignment parameters as covariates of no interest. To account for the inter-subject variability in the group analysis, linear contrasts of regression coefficients were computed at the individual subject level and then taken to a group-level random-effects analysis. The group analysis consisted of one-sample *t*-tests of the contrast images of all participants that indicated whether observed differences between conditions were significantly distinct from zero, or of paired *t*-tests of contrast images to analyze the statistical difference between. To protect against false-positive

activations a double threshold was applied, by which only regions with a z-score exceeding 2.94 ($p < .005$, uncorrected, and to $p < .05$, corrected) and a volume exceeding 60 contiguous voxels (corresponding to $p < .05$, corrected) fell above the threshold. This was determined in a Monte Carlo simulation using a Matlab script provided by Scott Slotnick (http://www2.bc.edu/-slotnics/scripts.htm).

2. Results

2.1. Behavioural results

Due to a technical problem we were not able to acquire behavioral data from four out of 16 participants. The task was to count the number of mismatch trials within each block of trials. Participants correctly counted 71 of the 74 mismatches in the Color task (one sample *t*-test, $t(11) = .737$, $p > .05$; $\alpha = .05$; $d = 2.08$). In the Intention task they overestimated the occurrence of mismatches (mean occurrence = 74, mean reported count = 150, one sample *t*-test, $t(11) = -5.175$, $p < .001$; $\alpha = .05$; $d = 1.4$). Performance was thus better for the Color task (paired samples *t*-test, $t(11) = 7.65$, $p < .05$; $d = .8$).

2.2. Neuroimaging results

The primary question of this study was whether any neural regions would be more sensitive to SI in observed social interactions than to multiple II when attending to intentions relative to when attending to perceptual events. To answer this question we first set out to investigate the main effects of the factors Intention and Task. When contrasting the Intention task against the Color task ($[SI_{intention} + PI_{intention}] > [SI_{colour} + PI_{colour}]$) we noted activations in the postcentral gyrus bilaterally, and the temporal pole in the right hemisphere (Table 1). In the reverse contrast, activity in the bilateral middle occipital gyri, right fusiform gyrus, left superior parietal lobe, right middle frontal and right precentral gyri were noted (Table 1).

To explore the interaction between the factors Intention (Shared *vs* Parallel) and Task (intention *vs* color) we carried out the following paired *t*-test: ($[SI_{intention} > PI_{intention}] > [SI_{colour} > PI_{colour}]$) An interaction effect was associated with activation in the bilateral middle frontal gyri, left precentral and superior frontal gyri (SFG), posterior cingulate/precuneus, right temporal pole, nucleus accumbens bilaterally, right Heschl's gyrus, thalamus/caudate and the cerebellum (Table 1, Fig. 2).

To further scrutinize the interaction between the Intention and Task factors we carried out the following simple contrasts. The contrast between SI and PI in the Intention task ($SI_{intention} > PI_{intention}$) revealed activations in the rostral part of the anterior cingulate cortex (rACC) bilaterally, bilateral SFG including pre/postcentral gyrus of the left hemisphere, middle cingulate gyrus and the precuneus, bilateral superior temporal gyri and medial parts of the temporal poles. Additional activations were found in midline thalamus, right caudate nucleus and right nucleus accumbens, and finally in the cerebellum (Table 1). Conversely, when PI were contrasted with SI ($PI_{intention} > SI_{intention}$) no suprathreshold activations were

Table 1 − Differential neural activations noted in (A). the paired-test [SI$_{intention}$ > PI$_{intention}$] > [SI$_{colour}$ > PI$_{colour}$] (B). Main effects, and (C). Simple contrasts. (voxel level $p < .005$, and cluster corrected for multiple comparisons, $p < .05$, unless otherwise indicated). Anatomical regions defined by the SPM Anatomy toolbox.

Anatomical region	MNI coordinates			Cluster size	Z
	x	y	z		
A. Interaction effect					
[SI$_{intention}$ > PI$_{intention}$] > [SI$_{colour}$ > PI$_{colour}$]					
Middle frontal gyrus	24	16	46	2869	4.78
Precentral sulcus	−32	−8	42	149	3.83
Superior frontal gyrus	−20	8	60	416	3.68
Middle frontal gyrus/ Frontal Pole	−26	54	26	983	3.64
Precuneus/PCC	6	−38	46	723	4.12
Temporal pole	34	20	−36	189	3.96
Heschyl's gyrus	50	−16	4	75	3.17
Thalamus/caudate	8	−4	4	1010	3.64
Thalamus/caudate	−16	−24	16	109	3.06
Nucleus accumbens	−18	10	−12	240	3.61
Nucleus accumbens/ putamen	28	−8	4	<60	3.27
Cerebellum	0	−50	6	971	3.58
	16	−32	−38	109	3.06
	46	−52	−28	239	3.19
	−10	−34	−40	992	4.01
B. Main effects					
Intention (SI > PI)					
Middle temporal gyrus	−52	−22	0	70	3.78
Superior occipital gyrus	24	−102	8	124	3.19
Middle occipital gyrus	−34	−80	2	65	3.17
Intention (PI > SI)					
None					
Task (Intention > Color)					
Postcentral gyrus	48	−30	52	399	3.64
	−50	−24	46	71	2.95
Middle temporal pole	56	10	−34	61	3.17
Task (Color > Intention)					
Middle occipital gyrus	−38	−72	14	1852	3.91
	30	78	18	118	4
Fusiform gyrus	28	−74	−8	93	3.16
Superior parietal lobule	−20	−58	52	173	3.26
Inferior frontal gyrus	48	18	6	502	3.7
Precentral gyrus	36	−2	48	320	3.58
Middle frontal gyrus	36	34	24	140	3.28
C. Simple contrasts					
SI$_{intention}$ > PI$_{intention}$					
rACC	6	36	8	173	3.56
	−4	40	12	<60	3.01
Precuneus/PCC	−12	−38	36	288	3.36
Middle frontal gyrus	32	30	42	593	4.04
	−24	6	68	760	3.83
Frontal pole	−16	62	16	331	3.93
Postcentral gyrus	−32	−32	56	70	3.42
Temporal pole	−18	0	−34	152	3.41
	−20	−28	−14	72	3.32
	−18	0	−34	96	3.36
Superior temporal gyrus	−50	−20	8	902	3.88
	52	−16	4	91	3.48
Putamen	26	10	−12	84	3.12
Caudate	24	2	16	376	3.61
Cerebellum	48	−52	−28	206	3.75
	22	−32	−28	472	3.71
	−14	−34	−32	627	3.89

Table 1 − (continued)

Anatomical region	MNI coordinates			Cluster size	Z
	x	y	z		
PI$_{intention}$ > SI$_{intention}$					
None					
SI$_{colour}$ > PI$_{colour}$					
None					

observed. The contrast between SI and *PI* events in the Color task (SI$_{colour}$ > PI$_{colour}$) yielded no significant activations either.

As for the Individual Intention condition, the contrast between SI and the Individual Intention in the Intention task (SI$_{intention}$ > II$_{intention}$) revealed significant activations (Table 2) in bilateral occipital cortices, bilateral superior temporal gyri, overlapping with auditory cortices and flowing posteriorly towards the TPJ, bilateral pre/postcentral sulci as well as the left frontal eye field, left frontal pole, precuneus and the left superior parietal lobule. PI when contrasted with Individual Intention in the Intention task (PI$_{intention}$ > II$_{intention}$) recruited the occipital cortices and the superior temporal gyri bilaterally.

3. Discussion

We set out to investigate whether observing social interactions with shared intentions recruits different neural regions than observing social interactions with multiple independent intentions, when either attending to the intentions of those involved in the interactions or when attending to visual features of the stimuli. To that end, we compared the BOLD response evoked by observing perceptually identical yet intentionally ambiguous joint actions embedded in two different intentional contexts (i.e., Shared Intentions *vs* Parallel Intentions) and in two different task settings (i.e., attend to intention *vs* attend to color). Our main contrast of interest was between the Shared Intention condition and the *PI* condition in the Intention task versus the Color task ([SI$_{intention}$ > PI$_{intention}$] > [SI$_{colour}$ > PI$_{colour}$]). This interaction test revealed activations in the right temporal pole, precuneus/PCC (posterior cingulate), and bilaterally in the ventral striatum, as well as in the bilateral superior frontal gyri (SFG) and the left frontal pole. We focus on these results in the following.

3.1. Mentalizing

Our first aim was to investigate whether processing intentional relations where the observed individuals' intentions are 'interlocked' (Bratman, 1992), involves mentalizing related computations. If shared intentionality in observed joint actions called upon processes involved in representing mental states of others' more than observing actions that appear to be guided by parallel intentions, this should manifest itself in increased activations in the mentalizing network in the Shared Intention condition compared to the Parallel Intentions condition. Both the interaction analysis of the factors
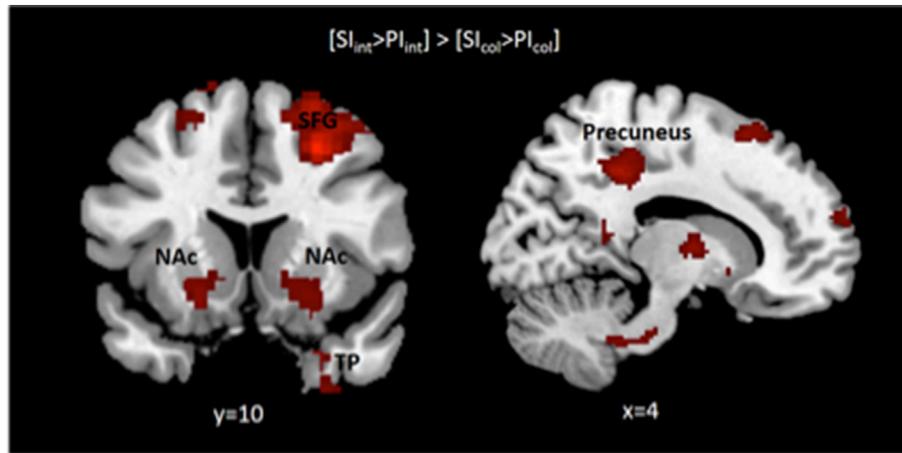
Fig. 2 — Significant activations revealed by the paired *t*-test ([SI$_{intention}$ > PI$_{intention}$] > [SI$_{colour}$ > PI$_{colour}$]) overlaid on a single subject T1 (voxel *p* < .005, cluster *p* < .05).

Intention and Task, and the simple contrast between Shared Intentions and Parallel Intentions conditions in the Intention task revealed activation in two areas of the mentalizing network, namely, the bilateral temporal poles and the precuneus/posterior cingulate (PCC). Other than participating in theory of mind reasoning, the precuneus/PCC seems to take on a more general role in social cognition (Amodio & Frith, 2006). It is involved in self-awareness and self-related processing as well as when thinking about intentions of one's own (Den Ouden, Frith, Frith, & Blakemore, 2005; Sommer

Table 2 — Differential neural activations noted in simple contrasts (A). Shared Intentions and Individual Intention in the Intention task, and (B). Parallel Intentions and Individual Intention in the Intention task (voxel level *p* < .005, and cluster corrected for multiple comparisons, *p* < .05).

Contrast	Coordinates (x. y, z)			Cluster size	Z
A. SI$_{intention}$ > II$_{intention}$					
Heschl's gyrus	58	−20	6	2096	4.92
	−50	−20	6	2442	4.73
Middle frontal gyrus	−36	16	52	546	3.88
	−26	56	8	259	3.25
	42	20	48	92	3.16
	38	60	10	64	3.22
Postcentral gyrus	−28	−40	62	130	3.32
	56	−24	56	239	3.29
	8	−44	68	209	3.15
	−30	−56	60	66	2.79
Superior fontal gyrus	−10	28	60	134	3.19
Precuneus	10	−44	12	222	3.24
Cuneus	10	−94	14	7904	4.72
Cerebellum	48	−60	−26	220	3.26
B. PI$_{intention}$ > II$_{intention}$					
Cuneus	20	−96	18	1890	4.7
	−30	−76	−2	1718	4.54
Heschl's gyrus	58	−20	6	203	3.31
	−50	−22	4	161	3.13
Middle frontal gyrus	−38	14	48	100	3.06
Cerebellum	−44	−74	−20	82	3.04

et al., 2007), and intentions of others' (Abraham, Werning, Rakoczy, von Cramon, & Schubotz, 2008). Activity in these regions is observed also when the task demands comprehension of cooperation or intentional deception as aspects of different intentional relations (Lissek et al., 2008). Importantly, precuneus has been found to support observing interactions taking place between others (Iacoboni et al., 2004; Kujala et al., 2012; Petrini, Piwek, Crabbe, Pollick, & Garrod, 2014; Sinke et al., 2010).

The temporal poles have traditionally been implicated in the storage and retrieval of semantic information, and are thought to play a supporting role in mental state reasoning (Gallagher & Frith, 2003). They are found active both in complex offline (Abraham et al., 2008; Calarge, Andreasen & O'Leary, 2003; Funnell, 2001; Ross & Olson, 2010), as well as in online mentalizing tasks (Assaf et al., 2009; Gallagher et al., 2002; Rilling et al., 2002) especially when the task involves a social story. Be it in the form of an online interaction, a written narrative or a comic strip, these tasks include a social script that captures conventions of conduct in particular situations. Social scripts are thought to aid processing of social situations as they outline the behavior of the interacting individuals in compliance with socially accepted practices. For example a script for dining at a restaurant entails the waiter seating the customers, bringing their order and ends with the customer paying the bill at the table. Patients with lesions of the temporal poles have been noted to struggle in acting according to these scripts (Funnell, 2001). Recent neuroimaging investigations of third-person viewings of social interactions also implicate the temporal poles when the task involves attending to the nature of the interaction taking place between two people (Kujala et al., 2012; Sinke et al., 2010; Spiers & Maguire, 2006; Wagner et al., 2011). Regarding the current study, we believe that the dialogs signaled an upcoming shared intentional interaction, which activated corresponding social scripts, reflected in the temporal pole activation. Social scripts are particularly relevant for social interactions where individuals act on a shared intention, as the interaction needs to follow certain rules of social conduct stipulated by the scripts (Gilbert, 2009). This is not so much the

case when people act independently, which could explain the weaker response in the Parallel Intentions condition.

The interaction analysis did not result in differential signals in the other two regions of the theory-of-mind network. These areas are the temporo-parietal junction (TPJ) and the medial prefrontal cortex (mPFC). The respective roles of these two areas mental state reasoning are still subject to debate in the field. Although the mPFC is commonly found in mentalizing tasks, including those that involve processing social interactions, it has been argued that the function of this area in mental state reasoning is not restricted to representation of mental states, but is more generalizable to processing social or emotional information of others (Saxe & Powell, 2006). Mental state representation, this account contends, is associated with TPJ activation. A study by Jenkins and Mitchell (2009) sought to differentiate two confounding aspects of commonly used ToM tasks: inferring intentions of others when the context is ambiguous, and representing different types of mental states (e.g., beliefs, intentions etc.). They found that while the latter process engages the TPJ, making sense of others' behavior when the situation is not sufficiently clear recruits the mPFC. Indeed, the tasks employed in the studies that implicate this region in ascribing mental states to single individuals required participants to resolve ambiguities in the stimuli and found mPFC to be differentially involved when comparing social to non-social conditions (e.g., De Lange, Spronk, Willems, Toni, & Bekkering, 2008). The studies investigating observation of interactions that implicated the mPFC also employed tasks that particularly required drawing mental state inferences in order to disambiguate and categorize the said social interaction (Centelles et al., 2011; Iacoboni et al., 2004; Kujala et al., 2012; Sinke et al., 2010; Spiers & Maguire, 2006; Wagner et al., 2011). This was not the case in the current study. In the current study the information pertaining to the intentions of the observed individuals was made available to the participants via the dialogs before they observed the interactions. This may explain the lack of mPFC activation in our results.

Whether processing mental states selectively recruits the TPJ (Jenkins & Mitchell, 2009; Saxe & Wexler, 2005), or the mPFC (Gallagher & Frith, 2003; Walter et al., 2004), the lack of activation in either of these areas in our results suggests that the brain is impartial to shared intentions and parallel intentions in observed social interactions with respect to the computations underpinning mental state representations. It has been proposed that the role of this computation is to decouple others' mental states from one's own (Leslie, German & Polizzi, 2005). Whether individuals act with a shared intention or with independent parallel intentions, both cases comprise two agents with two respective intentions, which remain decoupled from one another. In both cases processing the intentional relation involves second-order theory of mind reasoning (Frith, 2007) to comparable extents in the two conditions. As Bratman argues (1992) what separates SI from independent intentions is not their constitution, it is the way they are organized. Accordingly, shared intentions are ordinary intentions that are interlocked, processing which does not seem to additionally burden neural areas associated with representing mental states.

In addition, the simple contrast between Shared Intentions and Parallel Intentions in the Intention task

($SI_{int} > PI_{int}$) revealed a very similar pattern of activation as observed in the interaction analysis reported above. However, no suprathreshold activity was detected in the opposite contrast ($PI_{int} > SI_{int}$). In all, this unidirectional difference in the neural response between Shared Intentions and Parallel Intentions conditions when the task was to attend to intentions, suggests that observing social interactions with shared intentions has increased processing demands compared to observing individuals who act independently. This relative demand seems to be related to the (explicit) tracking of interlocking intentions, which involves social cognitive processes.

3.2. Expectation of interlocked intentional actions

Are there other neural regions involved in computing interlocking of intentions in observed shared intentional actions? Two other regions that were found significant in the interaction analysis are possible candidates; the bilateral superior and middle frontal gyri and the frontal poles (Fig. 2). Frontal poles have been implicated in integrating two or more separate cognitive operations in pursuit of higher goals (Ramnani & Owen, 2004), for example in multi-task coordination (Gilbert, Spengler, Simons, Steele, Lawrie et al., 2006), or in cooperative social interactions (McCabe et al., 2011). Superior frontal gyrus and the frontal poles have been associated with organization of sequential movements towards a desired final goal (Majdanzic et al., 2007; Van Schie & Bekkering, 2007). This suggests when perceiving joint actions, single actions of the partaking individuals are perceived as interlocking intermediary steps towards the attainment of the final goal. The activations we found in the SFG and the frontal pole in this contrast suggest that on the basis of the prior information provided with the dialogs people expect the upcoming observed action to be coordinated and interlocked in the Shared Intention condition. This may have lead them to monitor these actions as if they were indeed interlocked, even though they were perceptually identical to the actions in the Parallel Intentions condition. This implies a difference in processing of the otherwise identical stimuli. It is possible that the prior contextual information provided with the dialogs set the system to anticipate a certain contingency between the actions of the observed individuals. The expected contingency could explain the activation also noted in the cerebellum, an area known for its involvement in monitoring temporal aspects of movements (Schubotz, Friederici & von Cramon, 2000).

3.3. Spontaneous monitoring of intentional relations

Recent behavioral evidence suggests that mental states of others can be computed implicitly (Kovacs et al., 2010; Samson et al., 2010), with associated activity in the mentalizing network (Scheider et al., 2014). Our second aim was to test whether intentional relations in observed interactions are also processed implicitly. We addressed this question by including the Color task, which was a perceptual task requiring participants to attend to the visual features of the stimuli. Monitoring the intentions behind the observed interactions was not relevant for this task. Therefore any

mentalizing related activity found in this task (SI_{colour} > PI_{colour}) would have indicated an implicit processing of intentional relations in the observed actions. However, this contrast did not reveal any significant activation in the mentalizing network, or elsewhere. This indicates that participants did not automatically track the observed agents' intentions in the Color task.

3.4.   Reward processing

A further aim of the study was to determine if observing actions performed with shared intentionality triggers neural activity in reward areas in the onlookers. Indeed we found activation in the ventral striatum (nucleus accumbens), bilaterally. These areas (Izuma, Saito & Sadato, 2008) are involved in evaluation of social rewards (e.g., Delgado, Stenger & Fiez, 2004), such as money or reputation (Elliott, Friston & Dolan, 2000; Izuma et al., 2008). Just as in primary rewards, the reward circuitry mediates learning of contingencies between actions and their social rewards. Of particular interest, a number of neuroimaging studies detected activity in these areas when people were engaged in online social interactions where there is mutual contingency between the interacting individuals, compared to when individuals' actions are not contingent on one another (Behrens, Hunt, Woolrich & Rushworth, 2008; Walter, Abler, Ciaramidaro & Erk, 2005). From minimal cases such as interpersonal gaze (Kuzmanovic et al., 2009; Pfeiffer et al., 2014; Redcay et al., 2010; Schilbach et al., 2010; Williams, Waiter, Perra, Perrett, & Whiten, 2005) or simple interactions such as playing ball with avatars (David et al., 2006), to more complex ones that require reciprocal cooperative behavior, for example in neuroeconomic trust games (Rilling et al., 2002), social interactions of a range of complexity call on these areas. This suggests that interacting with others where there is mutual contingency between individuals' behavior generates a rewarding experience. What our findings suggest is that a similar response occurs when observing others in interaction, which possibly reflects a resonance with one's past experiences.

What could be the reward value in observing interactions? We know that participating in social interactions is rewarding and facilitates subsequent cooperative and prosocial behaviors (Tomasello, 2014; Tomasello, Carpenter, Call, Behne & Moll, 2005). When observing others in interaction, the pleasant rewarding experience we attribute to those we observe might be mapped onto our own experiences that we previously acquired during our interactions (Barresi & Moore, 1996). Such a response might further reinforce our motivation to engage in shared experiences with others (Tomasello, 2014; Tomasello et al., 2005). It may also serve as a means to learn information that is not available when observing others in isolation, thereby facilitating cultural transmission of joint practices. To our knowledge, this is the first study to demonstrate that performing actions together based on shared intentions creates a rewarding experience not only in those who are engaged in the interaction, but also in those who are passively observing it. Future studies may be able to dissociate motivational and affective/hedonic components of reward processing.

REFERENCES

Abraham, A., Werning, M., Rakoczy, R., von Cramon, D. Y., & Schubotz, R. I. (2008). Minds, persons, and space: an fMRI investigation into the relational complexity of higher-order intentionality. *Consciousness and Cognition, 17*(2), 438–450.

Akhtar, N. (2005). Is joint attention necessary for early language learning? In B. D. Homer, & C. S. Tamis-LeMonda (Eds.), *The development of social cognition and communication* (pp. 165–179) Mahwah, NJ: Erlbaum.

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

Assaf, M., Kahn, I., Pearlson, G. D., Johnson, M. R., Yeshurun, Y., Calhoun, V. D., et al. (2009). Brain activity dissociates mentalization from otivation during an interpersonal competitive game. *Brain Imaging and Behaviour, 3*, 24–37.

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioural and Brain Sciences, 19*, 107–154.

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

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2008). The computation of social behavior. *Science, 324*(5931), 1160–1164.

Boeckler, A., Knoblich, G., & Sebanz, N. (2011). Observing shared attention modulates gaze following. *Cognition, 120*, 292–298.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review, 101*(2), 327–341.

Calarge, C., Andreasen, N. C., & O'Leary, D. S. (2003). Visualizing how one brain understands another: a PET study of theory of mind. *American Journal of Psychiatry, 160*(11), 1954–1964.

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12*, 314–325.

Centelles, L., Assaiante, C., Nazarian, B., Anton, J.-A., & 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*(10), 1–8.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., et al. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychologia, 45*(13), 3105–3113.

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., et al. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *Journal of Cognitive Neuroscience, 28*(6), 898–910.

de Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10), 441–447.

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

Delgado, M. R., Stenger, V. A., & Fiez, J. A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cerebral Cortex, 14*, 1022–1030.

Den Ouden, H. E. M., Frith, U., Frith, C., & Blakemore, S.-J. (2005). Thinking about intentions. *NeuroImage, 28*, 787—796.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *The Journal of Neuroscience, 20*(16), 6159—6165.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 2*, 165—189.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society: Biology, 362*, 671—678.

Frith, C. D., & Frith, U. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society: Biology, 365*, 165—175.

Funnell, E. (2001). Evidence for scripts in semantic dementia: Implications for theories of semantic memory. *Cognitive Neuropsychology, 18*(4), 323—341.

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

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional in a competitive game. *NeuroImage, 16*, 814—821.

Gilbert, M. (2009). Shared intention and personal intentions. *Philosophical Studies, 144*, 168—187.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *Journal of Cognitive Neuroscience, 18*(6), 932—948.

Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *NeuroImage, 21*, 744—750.

Henderson, A. M. E., & Woodward, A. L. (2011). "Let's work together": what do infants understand about collaborative goals? *Cognition, 121*(1), 12—21.

Herold, K. H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: relations with self-recognition and perspective taking. *Journal of Experimental Child Psychology, 101*, 114—123.

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., & Fiske, A. P. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage, 21*, 1167—1173.

Ip, G. W., Chiu, C., & Wan, C. (2006). Birds of a feather and birds flocking together: physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology, 90*(3), 368—381.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Cell, 58*(2), 284—294.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences, 9*(1), 21—25.

Jenkins, A. D., & Mitchell, J. P. (2009). Mentalizing under uncertainty: dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex, 20*(2), 404—410.

Kovacs, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830—1834.

Kujala, M. V., Carlson, S., & Hari, R. (2012). Engagement of amygdala in third-person view of face-to-face interaction. *Human Brain Mapping, 33*, 1753—1762.

Kuzmanovic, B., Georgescu, A. L., Eickoff, S. B., Shah, N. J., Bente, G., Fink, G. R., et al. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage, 46*, 1154—1163.

Lakens, D. (2010). Movement synchrony and perceived entitativity. *Journal of Experimental Social Psychology, 46*(5), 701—708.

Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: movement synchrony leads to attributions of rapport and entitativity. *Social Cognition, 29*(1), 1—14.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45—85.

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., et al. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS One, 3*(4), e2023.

Majdanzic, J., Grol, M. J., van Schie, H. T., Verhagen, L., Toni, I., & Bekkering, H. (2007). The role of immediate and final goals in actin planning: an fMRI study. *NeuroImage, 37*, 589—598.

Manera, V., Becchio, C., Schouten, B., Bara, B. G., & Verfaille, K. (2011). Communicative interactions improve visual detection of biological motion. *PLoS One, 6*(1), 1—6.

McCabe, K., Houser, D., Ryan, L., & Trouard, T. (2011). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, 98*(20), 11832—11835.

Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of Experimental Social Psychology, 45*(3), 585—589.

Nielsen, M., Moore, C., & Mohamedally, J. (2012). Young children overimitate in third-party contexts. *Journal of Experimental Child Psychology, 112*(1), 73—83.

O'Doherty, K., Troseth, G. L., Shimpi, P. M., Goldenberg, E., Akhtar, N., & Saylor, M. M. (2011). Third-party social interaction and word learning from video. *Child Development, 82*(3), 902—915.

Petrini, K., Piwek, L., Crabbe, F., Pollick, F. E., & Garrod, S. (2014). Look at those two!: the precuneus role in unattended third-person perspective of social interactions. *Human Brain Mapping.* http://dx.doi.org/10.1002/hbm.22543.

Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., et al. (2014). Why we interact: on the functional role of the striatum in the subjective experience of social interaction. *NeuroImage, 101*, 124—137.

Pierno, A. C., Becchio, C., Turella, L., Tubaldi, F., & Castiello, U. (2008). Observing social interactions: the effect of gaze. *Social Neuroscience, 3*(1), 51—59.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience, 5*(3), 184—194.

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., et al. (2010). Live facte-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *NeuroImage, 50*(4), 1639—1647.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron, 35*(2), 395—405.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage, 22*(4), 1694—1703.

Ross, L. A., & Olson, I. R. (2010). Social cognition and the anterior temporal lobes. *NeuroImage, 49*(4), 3452—3462.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley-Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1255—1266.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: the role of movement information. *Cognition, 119*, 242—252.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science, 17*, 692—699.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia, 43*(10), 1391–1399.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., et al. (2010). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience, 22*(12), 2702–2715.

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

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*(8), 299–309.

Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing: a common cortical and subcortical revealed by fMRI. *NeuroImage, 11*(1), 1–12.

Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. L. Morgan, & M. E. Pollack (Eds.), *Intention and communication* (pp. 401–416). Massachussetts: MIT.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70–76.

Sinke, C. B. A., Sorger, B., Goebel, R., & de Gelder, B. (2010). Tease or threat? Judging social interactions from bodily expressions. *NeuroImage, 49*(2), 1717–1727.

Sommer, M., Doehnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *NeuroImage, 35*(3), 1378–1384.

Spiers, H. J., & Maguire, E. A. (2006). Spontaneous mentalizing during interactive real world task: an fMRI study. *Neuropsychologia, 44*, 1674–1682.

Tomasello, M. (2014). *A natural history of human thinking.* Cambridge, Massachusetts: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioural Brain Sciences, 28*(5), 675–690.

Van Schie, H. T., & Bekkering, H. (2007). Neural mechanisms underlying immediate and final action goals in object use reflected by slow wave brain potentials. *Brain Research, 1148*, 183–197.

Wagner, D. D., Kelley, W. M., & Heatherton, T. F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex, 21*, 2788–2796.

Walter, H., Abler, B., Ciaramidaro, A., & Erk, S. (2005). Motivating forces of human actions: neuroimgaing reward and social interaction. *Brain Research Bulletin, 67*(5), 368–381.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., & Bara, B. B. (2004). Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience, 16*(10), 1854–1863.

Waytz, A., & Young, L. (2012). The group-member mind trade-off: attributing mind to groups versus group members. *Psychological Science, 23*(1), 77–85.

van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition, 130*(1), 128–133.

Williams, J. H. G., Waiter, G. D., Perra, O., Perrett, D., & Whiten, A. (2005). An fMRI study of joint attention experience. *NeuroImage, 25*(1), 133–140.