INTRODUCTION

We tend to think of human communication as basically involving the coding-decoding of a conventional symbol system, which could be supported by parts of the human motor system (i.e., the “mirror neurons system”). However, this view does not explain how these conventions could develop in the first place. Here we target the neglected but crucial issue of how people organize their non-verbal behavior to communicate a given intention without pre-established conventions. We have measured behavioral and brain responses in pairs of subjects during communicative exchanges occurring in a real, interactive, on-line social context. In two fMRI studies, we found robust evidence that planning new communicative actions (by a sender) and recognizing the communicative intention of the same actions (by a receiver) relied on spatially overlapping portions of their brains (the right posterior superior temporal sulcus). The response of this region was lateralized to the right hemisphere, modulated by the ambiguity in meaning of the communicative acts, but not by their sensorimotor complexity. These results indicate that the sender of a communicative signal uses his own intention recognition system to make a prediction of the intention recognition performed by the receiver. This finding supports the notion that our communicative abilities are distinct from both sensorimotor processes and language abilities.

Keywords: social neuroscience, interactive game, fMRI, superior temporal sulcus

Human communication has been described as involving the coding-decoding of a conventional symbol system, which could be supported by parts of the human motor system (i.e., the “mirror neurons system”). However, this view does not explain how these conventions could develop in the first place. Here we target the neglected but crucial issue of how people organize their non-verbal behavior to communicate a given intention without pre-established conventions. We have measured behavioral and brain responses in pairs of subjects during communicative exchanges occurring in a real, interactive, on-line social context. In two fMRI studies, we found robust evidence that planning new communicative actions (by a sender) and recognizing the communicative intention of the same actions (by a receiver) relied on spatially overlapping portions of their brains (the right posterior superior temporal sulcus). The response of this region was lateralized to the right hemisphere, modulated by the ambiguity in meaning of the communicative acts, but not by their sensorimotor complexity. These results indicate that the sender of a communicative signal uses his own intention recognition system to make a prediction of the intention recognition performed by the receiver. This finding supports the notion that our communicative abilities are distinct from both sensorimotor processes and language abilities.

Keywords: social neuroscience, interactive game, fMRI, superior temporal sulcus

Human communication has been described as involving the coding-decoding of a conventional symbol system, which could be supported by parts of the human motor system (i.e., the “mirror neurons system”). However, this view does not explain how these conventions could develop in the first place. Here we target the neglected but crucial issue of how people organize their non-verbal behavior to communicate a given intention without pre-established conventions. We have measured behavioral and brain responses in pairs of subjects during communicative exchanges occurring in a real, interactive, on-line social context. In two fMRI studies, we found robust evidence that planning new communicative actions (by a sender) and recognizing the communicative intention of the same actions (by a receiver) relied on spatially overlapping portions of their brains (the right posterior superior temporal sulcus). The response of this region was lateralized to the right hemisphere, modulated by the ambiguity in meaning of the communicative acts, but not by their sensorimotor complexity. These results indicate that the sender of a communicative signal uses his own intention recognition system to make a prediction of the intention recognition performed by the receiver. This finding supports the notion that our communicative abilities are distinct from both sensorimotor processes and language abilities.

Keywords: social neuroscience, interactive game, fMRI, superior temporal sulcus

INTRODUCTION

We tend to think of human communication as basically involving the coding-decoding of a conventional symbol system, but framing human communication in terms of shared codes neglects its inferential nature (Levinson, 2000; Sperber and Wilson, 2001). Human communication rides on a large background of pragmatic inference—otherwise ironies, sarcasms, hints, and indirections would pass us by. Nor are we troubled by the vagueness or multiple ambiguities and semantic generalities in every utterance. The same system that resolves the coded messages probably lies behind our ability to communicate without any pre-existing conventions at all, as in the gestures one might use behind the boss’ back, or to signal to others out of earshot. A number of converging paths of evidence suggest that this faculty is distinct from our language abilities, and is ontogenetically and phylogenetically primitive to language (Levinson, 2006), yet at the same time constitutes the foundation for effective language use. This paper investigates the cognitive and cerebral bases of this faculty.

Given the pervasive ambiguity of communicative signals (Levinson, 1995; Sperber and Wilson, 2001), effective communication requires heuristics for selecting and interpreting the communicative intention of an observable behavior from a potentially infinite search-space. A recent and influential suggestion assumes that communication could occur “without any cognitive mediation”by means of an automatic sensorimotor resonance between the sender of a message and its receiver (Rizzolatti and Craighero, 2004). This framework, grounded in the discovery of ‘mirror neurons’ responding to both execution and observation of a given behavior (di Pellegrino et al., 1992), postulates that the intention conveyed by an observed behavior can be understood by means of a sensorimotor simulation (Gallese et al., 2004). However, communicative actions cannot be exclusively guided by predictions of the sensory consequences of motor commands acting on one’s own body (Wolpert et al., 2003), since they need to be selected by taking into account the receiver’s knowledge (Clark and Carlson, 1982) and they are designed to trigger a mental state, not an observable sensory event.

Here, we address the generation of human communicative actions, testing the hypothesis that effective communicative behavior relies on a predictive mechanism constrained by conceptual knowledge (Goldman, 2006; Nichols and Stich, 2003), rather than by sensorimotor routines. The problem facing a sender is how to select a communicative action appropriate to convey a specific intention to a receiver. A sender could solve this problem by predicting the intention that a receiver would attribute to the sender’s action (Levinson, 2006). Crucially, we hypothesize that this prediction relies on the sender’s intention recognition system, taking knowledge and beliefs of the receiver into account. This hypothesis implies a computational overlap between selection and recognition of a communicative behavior in senders and receivers, respectively. A stringent test of this cognitive scenario is that the same cerebral structures support the planning of communicative acts (in the sender) and the recognition of the intentions conveyed by those acts (in the receiver), and that these cerebral activities are
modulated by the ambiguity in meaning of the communicative acts, rather than by their sensorimotor complexity.

Most human communication is a compound of coded conventional symbolic meaning (as in language), and inferences about communicator intent and recipients’ abilities to infer it. In order to focus on the latter system, we tested the hypothesis of shared computational overlap in communicators and recipients in the context of a controlled and unfamiliar communication system that prevented the participants from using pre-established linguistic conventions, forcing them to generate and interpret new communicative visuomotor behaviors. Using a novel interactive game set in a real, on-line social context, we could study the mechanisms that create new communicative conventions, rather than the utilization of such conventions, while manipulating the communicative ambiguity of different experimental trials. Furthermore, by using a controlled communicative setting, it becomes possible to design control trials devoid of communicative purposes, but matched in sensory and motor features with the communicative trials. We have called this new experimental protocol the Tacit Communication Game (TCG).

Pairs of participants [labeled as sender (male) and receiver (female), each controlling one token on a common game board (Figure 1)] were asked to jointly reproduce the spatial configuration of two target tokens (goals). When the goals were shown to the Sender only (communicative trials), solving the game required him to communicate to the receiver her own goal. Therefore, the Sender had to communicate to the Receiver the position and the orientation of her token. The Sender could achieve this only by moving his own token over the game board (Figure 1, phase 4). The receiver could then move to her own goal position, as inferred from the Sender’s movements.

By using time-resolved event-related functional Magnetic Resonance Imaging (fMRI), we could measure neurophysiological correlates of planning a communicative action (in senders) and recognizing its communicative intention (in receivers), comparing these effects to the activity evoked during planning and observation of non-communicative actions, and distinguishing these effects from sensory and motor events occurring during the same trials. Given that sensorimotor- and conceptually-based accounts of mind-reading make opposite predictions on the involvement of the motor system during communicative behavior, it becomes possible to distinguish between these general frameworks by examining the sensory/motor characteristics of the cerebral activity evoked during the selection of a communicative action.

MATERIALS AND METHODS

PARTICIPANTS

We recruited 56 right-handed participants, aged between 18 and 26 years, with normal or corrected-to-normal vision. Participants gave informed consent according to institutional guidelines of the local ethics committee (CMO region Arnhem-Nijmegen, Netherlands), and were either offered a financial payment or given credits towards completing a course requirement. Twenty-four male–female pairs participated in the first experiment. Eight participants participated in the second experiment (six males). In the second experiment, the Sender (male, 31 years old) was an accomplice. For ease of explanation, in the following sections we consider a male sender and a female receiver.

It is known that cognitive abilities, stress responses, and corresponding regional patterns of cerebral activity are heavily influenced by the menstrual cycle (see for instance Fernandez et al., 2003). In order to control this source of variability, we have chosen to scan almost exclusively males. Although this approach limits the scope of our inferences, it goes beyond the scope of this study to extensively assess the influence of gender differences in communicative abilities and their cerebral correlates.

MATERIALS

The tacit communication game (TCG) involves two players, a sender and a receiver, moving a token on a game board displayed on a monitor. Participants were trained in front of two 19-inch computer monitors, playing the TCG with hand-held controllers (Figure 1). The spatial lay-out of the buttons on the hand-held controller allowed for unique mappings between finger and token movements: four face buttons moved the token to the left, right, up and down; two shoulder buttons rotated the token clockwise and counter-clockwise; a third shoulder button was used as a start button (see below). Players sat on opposing sides of a long table, each facing their own computer monitor, wearing sound-proof head sets and ear plugs (to minimize the influence of sounds in the environment and incidental noises produced by the players). The game was programmed using Presentation version 9.2 and was run on a Windows XP personal computer.

During scanning, one participant lay supine in the bore of the magnetic resonance (MR) scanner, playing the TCG with an MR-compatible hand-held controller. The other participant played the game from another room while wearing a sound-proof head set. Experiment 1 lasted about two and a half hours (30 min training, 45 min first fMRI session, 10 min rest, 10 min training, 45 min second fMRI, and 10 min anatomical scan). Experiment 2 lasted about 1 h and 30 min (30 min training, 20 min first fMRI session, 10 min rest, 20 min second fMRI session, 10 min anatomical scan).

PROCEDURES

Procedures – Experiment 1

The fMRI experiment consisted of two sessions: In one session 40 communicative trials were presented, and in the other session 40 non-communicative trials were presented. The exact same stimuli (including tokens and goal configurations) were used in the communicative and the non-communicative sessions. As described above (see also Figure 1), the communicative trials required the sender to move his token over the game board to indicate to the receiver where she needed to go. During the non-communicative trials it was made clear to the sender that the receiver also saw the goal configuration. Hence, there was no need for the sender to communicate to the receiver the position and orientation of her token. Further instructions ensured that the actions of the sender were similar during the communicative and non-communicative trials. Namely, during non-communicative trials, the sender was instructed to first move his token to the target position of the receiver, match the rotation of the target token as closely as possible, and then move to his own position. This procedure ensured that, during both types of trials, senders planned similar actions with their tokens. In contrast, during communicative trials, senders were meant to plan actions with others...
communicative value (for the receiver). The order of the two sessions was counterbalanced over participants.

To allow for comparisons between communicative and non-communicative trials over fMRI sessions both sessions also presented 40 basic control trials, which were the same in each session. This construction allowed for the comparison between different session events through the comparison of a shared control event. For a control trial it was made clear to the sender that the receiver could see the goal configuration (similar to a non-communicative trial). The control trials were simplified by asking senders to directly move their token to the correct location, and thereby to completely ignore the token of the receiver (in contrast to the non-communicative and communicative trials where the position and orientation of the token of the receiver played a crucial role).

Two further aspects were considered in the design: motoric complexity and communicative ambiguity. Motoric complexity varied naturally as some trials required planning of more moves (i.e. more button presses) and these could take longer to execute. Communicative ambiguity was varied by subdividing the communicative trials in 30 easy communicative trials and 10 difficult communicative trials. The difference between these two types of trials was that for easy trials senders did not face “orientation” problems. An example of an orientation problem is depicted in Figure 1: The sender has to indicate the rectangle needs to rotate with a token (a circle) that does not have a visible rotation. Pilot studies had shown us that without the orientation problem sender–receiver pairs quickly build up a set of successful unambiguous communicative actions. In contrast, with the orientation problem the actions of the sender stay ambiguous to the receiver for a longer time.

The experimental design was originally conceived to optimize the contrast between communicative and non-communicative conditions. As will become clear in the “Results” section, the most relevant comparison turned out to lay in the contrast between difficult communicative and easy communicative trials. While the numerical disparity between Difficult and Easy trials might lead to imbalances in the estimation of the signal associated with each of these two conditions, the effects we report are relative to the reliability of their estimates.

Participants were trained extensively before playing the TCG. In a first training session, the participants were individually familiarized with the procedure of translating and rotating their token.
around the game board. There were three different tokens: circles, triangles, and rectangles. In the second training session (10 trials) participants learned that the token placed below the game board was controlled by the sender, whereas the token placed above the game board was controlled by the receiver. During this training both sender and receiver could see two more (target) tokens inside the game board. These target tokens indicated the position and orientation that sender’s and receiver’s (playing) tokens should have at the end of the trial. The goal of this training was that each player positioned their own token in the designated position.

In the third training session, the participants were familiarized (15 trials) with condition-specific procedures of the TCG. There were two separate training sessions for communicative and non-communicative trials, counterbalanced according to the order used for the ensuing fMRI sessions. The sender was informed about the trial type (communicative/non-communicative or control) by the hue of the receiver’s token. A bright receiver’s token was used during control trials, a semi-transparent token was used during communicative/non-communicative trials.

**Procedures – Experiment 2**

In Experiment 2, we were interested in localizing brain regions involved in the observation/interpretation of communicative actions, in the context of the TCG. Accordingly, in this experiment we measured BOLD fMRI signals evoked in the receiver. The fMRI experiment consisted of two sessions: In one session 40 communicative trials were presented, and in the other session 40 non-communicative trials were presented. The communicative trials were the same in Experiments 1 and 2 (labelled as communicativeReceiver trials). Furthermore, we introduced a new type of non-communicative trial (labelled as non-communicativeReceiver trials). These trials ensured that receivers were monitoring the actions of the sender, but without attaching any communicative value to the movements. This was achieved by asking the receiver to move her token to the position in the game board where the sender last moved (i.e. translated or rotated) his token at least twice, and then rotate her token twice. We also informed the receiver that the sender moved according to our specific instructions, rather than trying to communicate to the receiver her target configuration. In fact, unknowingly to the receiver, we played back the sender’s moves of his actual movements performed during the communicative trials scanning session. This procedure ensured that the receiver was presented with identical visual input (from the sender) and followed the actions of the sender in both scanning sessions. Most importantly, receivers only observed actions with a communicative value (from the sender) in a communicativeReceiver trial.

Differently from Experiment 1, here we compared communicativeReceiver and non-communicativeReceiver Trials over fMRI sessions. We achieved this by relying on shared events in both sessions to serve as a baseline event (in this case the execution phase of the receiver).

Participants were trained extensively on the game procedures before playing the TCG during fMRI acquisition. The first two training sessions were identical to those of Experiment 1. The third training session consisted of 20 communicativeReceiver trials, during which the receiver could not see the target configuration (Figure 1). Following this scanning session, the receiver was informed that he needed to solve a different problem (the above mentioned visual following task). The receiver practiced this new task for 20 trials. During the ensuing scanning session, the receiver was in the MR scanner, and performed 40 non-communicativeReceiver trials. There were 30 difficult communicativeReceiver trials, and 10 easy communicativeReceiver trials. We changed the ratio for easy and difficult trials from Experiments 1 to 2, because we had learned from Experiment 1 (as will be shown below) that the difficult trials were most relevant.

**Behavioral Data Analysis**

For Experiment 1 we calculated mean planning times (senders; time between the onset of the goal configuration and the moment the sender pressed start), mean number of moves (senders), and mean accuracy scores (sender–receiver pairs). These dependent variables were analyzed using paired t-tests (threshold, p < 0.05) for communicative vs. non-communicative trials, and easy communicative vs. difficult communicative trials. We considered a highly conservative level of chance performance by taking into account only the position (and not the orientation) of the target token of the receiver. If the receiver randomly places a token in the game board there is a chance of one out of eight (12.5%) that it is correct (in terms of position), given that the sender knows where to position his token, the receiver cannot position her token on top of the token of the sender, and there are nine initial positions in the gameboard. In reality, tokens often had to be rotated adding additional options for the receiver, but this would further lower the chance level and the conservative estimate of 12.5% sufficed to show that sender–receiver pairs scored far above chance (see below).

During Experiment 2, mean planning times (receivers, time between the end of the visible movements of the sender and the moment the receiver pressed start), mean number of moves (receivers), and mean accuracy scores (sender–receiver pairs) were analyzed using paired t-tests (threshold, p < 0.05) for communicativeReceiver vs. non-communicativeReceiver trials and easy communicativeReceiver vs. difficult communicativeReceiver trials. Only the trials in which the receiver moved to the correct position were taken into account for calculating planning times and number of moves.

**Image Acquisition**

Images were acquired using a 3-Tesla Trio scanner (Siemens, Erlangen, Germany). Blood oxygenation level dependent (BOLD) sensitive functional images were acquired using a single shot gradient echo planar imaging (EPI) sequence (TR/TE 2.50 s/40 ms, 34 transversal slices, interleaved acquisition, voxel size 3.5 × 3.5 × 3.5 mm). At the end of the scanning session, structural images were acquired using a MP-RAGE sequence (TR/TE/TI 2300 ms/3.9 ms/1100 ms, voxel size 1 × 1 × 1 mm).

**Image Analysis**

Functional data were pre-processed and analyzed with SPM2 (Statistical Parametric Mapping)\(^1\). The first four volumes of each participant’s timeseries were discarded to allow for T1 equilibration. The image timeseries were spatially realigned using a sinc interpolation algorithm that estimates rigid body transformations (translations, rotations) by minimizing head-movements between

\(^1\)www.fil.ion.ucl.ac.uk/spm
each image and the reference image (Friston et al., 1995). The
timeseries for each voxel was realigned temporally to acquisition
of the middle slice. Subsequently, images were normalized onto a
custom Montreal Neurological Institute (MNI)-aligned EPI tem-
plate (based on 28 male brains acquired on the Siemens Trio at
the Donders Centre) using both linear and nonlinear transforma-
tions and resampled at an isotropic voxel size of 2 mm. Finally,
the normalized images were spatially smoothed using an isotropic
8 mm full-width-at-half-maximum Gaussian kernel. Each partici-
pant’s structural image was spatially coregistered to the mean of
the functional images (Ashburner and Friston, 1997) and spatially
normalized by using the same transformation matrix applied to
the functional images. The fMRI timeseries were analyzed using
an event-related approach in the context of the General Linear
Model (GLM).

STATISTICAL MODEL AND INFERENCE – EXPERIMENT 1

We considered 10 event types¹ for each scanning session (see
Table 1), following the sequence of events described in Figure 1.

During the non-communicative session, we defined the same 10
event types, replacing communicative with non-communicative
trials. In both sessions, each event timeseries was convolved with
a canonical hemodynamic response function and used as a regressor
in the SPM multiple regression analysis. In addition, we considered
the modulatory effects of two further parameters, adding four
further effects to the statistical model (for each session). First,
we considered the effects of planning movements with different
number of moves on the planning-related activities of the sender.
This was modelled as a parametric modulation of the number of
moves that the sender executed on a given trial on each of the three
planning periods of the sender (events 1, 2, 3 in Table 1). Second,
we considered the effects of executing movements of different
duration on the execution-related activity of the sender. This was
modelled as a parametric modulation of the movement time of the
sender on the execution effect (event 4 in Table 1). We assumed
a linear relation between number of sender moves and BOLD
signal, as well as between the duration of sender movement phases
and BOLD signal. The corresponding regressors were introduced
in the GLM on a subject-by-subject basis. Finally, the statistical
model also considered separate covariates describing the head-
related movements (as estimated by the spatial realignment pro-
cedure) and their first and second derivatives over time. Several
studies have included these derivatives of realignment parameters
to improve the sensitivity of their statistical analyses (e.g. Lund
et al., 2005; Salek-Haddadi et al., 2003; Verhagen et al., 2008). Data
were high-pass filtered (cut-off 128 s) to remove low frequency
confounds, such as scanner drifts. Temporal autocorrelation was
modelled as an AR(1) process.

The main focus of interest in this experiment was related to
the difference (in the sender) between planning communicative
and non-communicative actions. This difference was isolated
by testing planning-related activity for communicative versus
non-communicative trials (having subtracted out control
trials from each condition). This contrast can be expressed in
terms of events in the fMRI design (see Table 1): The activity
that was greater for events 1 and 2 (versus event 3) in the com-
municative sessions than for events 1 and 2 (versus event 3) in the
non-communicative sessions. We also tested the reverse contrast
to examine possible de-activations for communicative trials.
Finally, we also considered a more generic effect, related to the
difference between planning actions that required the sender to
take into account the target configuration of the receiver (i.e.,
communicative and non-communicative trials) and planning
action that were independent from the receiver (i.e., control
trials). This difference was isolated by testing for a difference
in planning-related activity between communicative and non-
communicative trials, as compared to the control trials.

Session-specific parameter estimates were calculated at each
voxel for each subject, and contrasts of the parameter estimates
were calculated for the effects of sender planning communicative
actions, sender planning non-communicative actions, and
sender planning control actions. These contrasts were entered

---

¹During a pilot phase, and after the collection of the imaging data, we tested whether
there were high correlations between relevant regressors in the design matrix
described above. The experimental design and the statistical model described above
ensured that the maximum correlation between planning-related and execution-
related regressors was <30%. Previous experience with these types of experimental
designs have shown the validity of this approach, and its ability to effectively dis-
associate planning- and execution-related effects (Tholenissen et al., 2002; Toni et al.,
1999, 2002).

<table>
<thead>
<tr>
<th>Event</th>
<th>Trial period</th>
<th>Trial type</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sender planning</td>
<td>COMMUNICATIVE EASY</td>
<td>Variable (from presentation of target configuration until Sender pressed the start button)</td>
</tr>
<tr>
<td>2</td>
<td>Sender planning</td>
<td>COMMUNICATIVE DIFFICULT</td>
<td>As above</td>
</tr>
<tr>
<td>3</td>
<td>Sender planning</td>
<td>CONTROL</td>
<td>As above</td>
</tr>
<tr>
<td>4</td>
<td>Sender moving</td>
<td>COMMUNICATIVE DIFFICULT CONTROL</td>
<td>Fixed – 5 s (from Sender pressing the start button)</td>
</tr>
<tr>
<td>5</td>
<td>Receiver planning</td>
<td>COMMUNICATIVE EASY COMMUNICATIVE DIFFICULT</td>
<td>Variable (from presentation of the Receiver’s shape until the Receiver pressed the start button)</td>
</tr>
<tr>
<td>6</td>
<td>Receiver planning</td>
<td>CONTROL</td>
<td>As above</td>
</tr>
<tr>
<td>7</td>
<td>Receiver moving</td>
<td>COMMUNICATIVE</td>
<td>Fixed – 5 s (from Receiver pressing the start button)</td>
</tr>
<tr>
<td>8</td>
<td>Receiver moving</td>
<td>CONTROL</td>
<td>As above</td>
</tr>
<tr>
<td>9</td>
<td>Feedback</td>
<td>Correct</td>
<td>Fixed – 1.5 s (from presentation of Feedback stimulus)</td>
</tr>
<tr>
<td>10</td>
<td>Feedback</td>
<td>Incorrect</td>
<td>As above</td>
</tr>
</tbody>
</table>
into a one-way, repeated measures analysis of variance (ANOVA),
treating subjects as a random variable. The degrees of freedom
were corrected for nonsphericity at each voxel.

We report the results of a random effects analysis, with
inferences drawn at the cluster level, corrected for multiple
comparisons over the whole brain using family-wise error cor-
rection ($p < 0.05$) (Friston et al., 1996). Furthermore, to improve
the sensitivity of the crucial test of the hypotheses described
above, we have assessed the results of the first two contrasts
on the basis of independent anatomical information (Friston,
1997), i.e. published stereotactical coordinates of areas related to
conceptually-based accounts of mind-reading (as studied with
Theory of Mind tasks – first cerebral network) or to sensori-
motor-based accounts (as implemented in the Mirror Neuron
System – second cerebral network). Whenever necessary, we
converted the published coordinates into MNI space. This pro-
cedure constrained our search space to a series of volumes of
interests (VOI)3, ensuring increased and matched sensitivity for
each cerebral network. We defined the first cerebral network on
the basis of (Saxe and Wexler, 2005; Saxe et al., 2004), position-
ing six VOIs along the left and right temporo-parietal junction
(TPJ) (−48, −69, 21; 54, −54, 24), the left and right posterior
superior temporal sulcuc (pSTS) (−54, −42, 9; 54, −42, 9), the
medial prefrontal cortex (mPFC) (0, 60, 12), and the posterior
cingulate (3, −60, 24). We defined the second cerebral network
on the basis of (Iacoboni et al., 1999), positioning six VOIs along
the left and right inferior frontal gyrus (−51, 12, 14; 51, 12, 14),
the left and right parietal operculum (−59, −26, 33; 59, −26, 33),
and the left and right intraparietal sulcus (−37, −44, 60; 37, −44,
60). We determined our VOIs on the basis of those particular
reports given that, in our judgement, they provided landmark
references for defining the cerebral correlates of Theory of Mind
and Mirror Neuron responses. Those studies were also instru-
mental in defining tasks most commonly associated with Theory
of Mind (i.e. false belief stories, recognizing intentions from
human actions) and the Mirror Neuron System (action obser-
vation and execution), and they have been highly influential in
determining the theoretical positions of these two accounts of
intention understanding.

3The radius of each VOI was set at 10 mm, using the WFU PickAtlas software tool-
box (Maldjian et al., 2003, 2004), correcting for multiple comparisons over the
joint volume spanned by the individual VOIs. This resulted in a $t$-value of 3 (cor-
responding to $p = 0.002$) for identification of supra-threshold clusters. Note that
this threshold is only used to define clusters, and does not denote the threshold for
significance of activations.

### STATISTICAL MODEL AND INFERENCE – EXPERIMENT 2

We considered five event types for each scanning session (see
Table 2), following the sequence of events described in Figure 1.

During the non-communicative session, we defined the same
five event types, replacing $\text{COMMUNICATIVE}_{\text{Receiver}}$ with $\text{NON-COMMUNICATIVE}_{\text{Receiver}}$ trials. In both sessions, each event timeseries
was convolved with a canonical hemodynamic response function
and used as a regressor in the SPM multiple regression analysis.
We also included separate covariates describing the head-related
movements (as estimated by the spatial realignment procedure) and
their first and second derivatives over time. We left out the parametric
modulations related to movements for Experiment 2 because the
moves for the receiver, unlike the moves for the sender, were only
related to simply moving their token to the target position. Data
was high-pass filtered (cut-off 128 s), and temporal autocorrelation
was modelled as an AR(1) process.

The main focus of interest in this experiment was related to
the difference (in the receiver) between observing $\text{COMMUNICATIVE}$
and $\text{NON-COMMUNICATIVE}$ actions. This difference was iso-
lated by testing planning related activity for $\text{COMMUNICATIVE}$
versus $\text{NON-COMMUNICATIVE}$ trials (each having subtracted
out common movement-related activity). This contrast can be
expressed in terms of events in the fMRI design (see Table 2):
The activity that was greater for event $2_R$ (versus event $4_R$)
in the $\text{COMMUNICATIVE}_{\text{Receiver}}$ Sessions than for event $2_\phi$
(versus event $4_\phi$) in the $\text{NON-COMMUNICATIVE}_{\text{Receiver}}$
Sessions.

Session-specific parameter estimates were calculated at each
voxel for each subject, and contrasts of the parameter estimates
were calculated as described above. These subjects-specific con-
trasts were entered into a non-parametric test, treating subjects
as a random variable. We report the results of a random effects
analysis, corrected for multiple comparisons over the whole brain
using family-wise error correction ($p < 0.05$). We employed the
non-parametric variant of SPM ($\text{SnPM}$; Nichols and Holmes,
2002). We used a locally pooled variance estimate (pseudo-$t$), with
a Gaussian kernel of 8 mm FWHM (Nichols and Holmes, 2002). To
optimize statistical sensitivity for both spatially extended clusters
and high intensity signals, we used a combined threshold on the
basis of voxel-intensity and cluster size (Hayasaka and Nichols,
2004), using a pseudo-$t$ value of 3 (corresponding to $p = 0.002$)
for identification of supra-threshold clusters.

### RESULTS

#### fMRI DATA

Table 3 provides an overview of the results obtained from the main
contrasts in the random effects analysis. We found that planning

<table>
<thead>
<tr>
<th>Event</th>
<th>Trial period</th>
<th>Trial type</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1_n$</td>
<td>Sender planning</td>
<td>$\text{COMMUNICATIVE}_{\text{Receiver}}$</td>
<td>Variable (from presentation of target configuration until Sender pressed the start button)</td>
</tr>
<tr>
<td>$2_n$</td>
<td>Sender moving Receiver observing</td>
<td>$\text{COMMUNICATIVE}_{\text{Receiver}}$</td>
<td>Fixed – 5 s (from Sender pressing the start button)</td>
</tr>
<tr>
<td>$3_n$</td>
<td>Receiver planning</td>
<td>$\text{COMMUNICATIVE}_{\text{Receiver}}$</td>
<td>Variable (from presentation of the Receiver’s shape until the Receiver pressed the start button)</td>
</tr>
<tr>
<td>$4_n$</td>
<td>Receiver moving</td>
<td>$\text{COMMUNICATIVE}_{\text{Receiver}}$</td>
<td>Fixed – 5 s (from Receiver pressing the start button)</td>
</tr>
<tr>
<td>$5_n$</td>
<td>Feedback</td>
<td>$\text{COMMUNICATIVE}_{\text{Receiver}}$</td>
<td>Fixed – 1.5 s (from presentation of Feedback stimulus)</td>
</tr>
</tbody>
</table>
novel communicative acts and understanding the communicative intention of these acts relied on the same cerebral tissue, namely the posterior part of the superior temporal sulcus (pSTS) of the right hemisphere (Figures 2A–C,E). During the planning phase of the TCG there was no change in sensory input or motor output (phase 2 in Figure 2). Therefore, this differential planning-related activity cannot be driven by visual motion or hand movements. Yet, it is possible that this particular pSTS cluster is sensitive to sensory stimuli or to motor responses, having been implicated in the perception of biological motion (Peelen et al., 2006) and in receiving reafferent motor-related activity (Iacoboni et al., 2001). We tested this possibility by assessing the BOLD activity measured during the execution of the communicative movements by the Sender (phase 4 in Figure 2). During this period the Sender moved his fingers over the controller and perceived his token moving over the game board. There was no reliable activity during this phase (Figure 2D, “Sender moves” bar). Taken together, these responses indicate that this portion of the right pSTS is not responsive to visual motion or to hand movements per se. Rather, this cluster appears to be involved in processing visual motion when this becomes relevant for inferring the communicative intentions of the agent (i.e. for the Receiver).

We reasoned that if the right pSTS is specifically involved in planning communicative intentions, then the response of this region should be modulated by communicative ambiguity, and not by the motoric complexity of the action used to convey the relevant information to the Receiver. Therefore, we tested whether the pSTS activity (measured during the planning phase) was sensitive to (1) the different communicative complexity of easy and difficult communicative trials; (2) the number of moves performed in the execution phase of the communicative trials; or (3) the time spent moving during the execution phase of the same trials. There was no reliable linear relationship between pSTS activity and motoric complexity (Figure 2D, “Number of moves” and “Movement Time” bars), but there was a strong effect of communicative complexity (Figure 3).

The pSTS activity was co-activated with the medial prefrontal cortex (mPFC) (Figure 4A) previously associated with conceptually-based accounts of the human ability to make inferences about mental states of other agents (Frith and Frith, 2006b; Saxe et al., 2004). The mirror-system regions hypothesized to provide efferent copies of motor commands to the pSTS (Iacoboni et al., 2001) showed increased de-activations during the planning of communicative actions (Figure 4B). The difference between communicative and non-communicative trials was present in the right pSTS, but not the left pSTS (Figure 4C).

In receivers, the same right pSTS activity was evoked during the recognition of the communicative intentions of the senders (Figures 2B,E). Similar to the sender, this activity was modulated by communicative ambiguity (Figure 5) and not present during motor execution (Figure 2F). The anatomical overlap between the activity evoked in the sender during the planning of communicative actions and the activity evoked in the receiver during the observation of the communicative actions (Figure 6) was unlikely (p = 0.017) to have occurred by chance, even at the relatively coarse spatial scale of group fMRI studies4.

**Behavioral data**

Senders had longer planning times, made more errors, and made more moves on communicative than on non-communicative trials, t(23) = 4.9, p < 0.001, t(23) = 7.9, p < 0.001, and t(23) = 4.7, p = 0.01, respectively (Figure 7A). Receivers made more moves on non-communicative trials than communicative trials, t(7) = 5.9, p = 0.001 (Figure 7B). There were no differences for the planning times and the accuracy scores obtained in these two types of trials (both t(7) < 1). These results suggest that the receivers were able to infer the communicative intentions of the sender, and that the motoric demands of the non-communicative trials were actually larger than those of the communicative trials.

Senders had longer planning times, made more errors, and made more moves on difficult communicative than on easy communicative trials, t(23) = 6.8, p < 0.001, t(23) = 13.4, p < 0.001, and t(23) = 3.6, p = 0.01, respectively (Figure 8A). Crucially, performance remained well above chance level also in the difficult communicative trials [t(23) = 7.2, p < 0.001, given the most conservative estimate of chance for every trial (12.5%, see Materials and Methods)], despite showing a significant increase in error rate as compared to the other trial types. Receivers made more moves on difficult communicative than on easy communicative trials, t(7) = 21.5, p < 0.001 (Figure 8B). There were no significant differences for planning times and accuracy scores, t(7) = 1.5, p = 0.19, and t(7) = 2, p = 0.09, respectively. It is obvious that the behavioral

---

4A conservative estimate of this probability was obtained by using a coarse but automated parcellation of the human brain in 116 unique structures (Tzourio-Mazoyer et al., 2002). Given that we performed independent experiments, on different groups of subjects, the probability of finding activity in the same region for the receiver as for the sender is given by the number of regions found for the receiver [2] divided by the number of regions in which activity might have occurred [116] (2/116 = 0.017).

---

**Table 3 | Results of the random effects analysis related to Communicative (COM), Non-Communicative (NON-COM) and Control trials.**

<table>
<thead>
<tr>
<th>Area</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>T value</th>
<th>No. of voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sender (planning)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM &gt; NON-COM</td>
<td>Right pSTS</td>
<td>50</td>
<td>–42</td>
<td>14</td>
<td>3.6</td>
</tr>
<tr>
<td>(COM and NON-COM) &gt; Control</td>
<td>Medial prefrontal cortex</td>
<td>4</td>
<td>54</td>
<td>18</td>
<td>4.9</td>
</tr>
<tr>
<td>Control &gt; (Com and NON-COM)</td>
<td>Left parietal operculum</td>
<td>–52</td>
<td>–24</td>
<td>44</td>
<td>8.6</td>
</tr>
<tr>
<td></td>
<td>Intraparietal sulcus</td>
<td>–38</td>
<td>–34</td>
<td>64</td>
<td>8.0</td>
</tr>
<tr>
<td><strong>Receiver (observing)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM &gt; NON-COM</td>
<td>Right pSTS</td>
<td>50</td>
<td>–38</td>
<td>6</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>Posterior paracingulate cortex</td>
<td>0</td>
<td>–38</td>
<td>20</td>
<td>6.3</td>
</tr>
</tbody>
</table>
Noordzij et al. Brain mechanisms underlying human communication

FIGURE 2 | Sequence of events in a COMMUNICATIVE trial of the TCG (all conventions as in Figure 1), and cerebral activity evoked in the sender and in the receiver during relevant trial epochs. (A) Planning communicative actions increased metabolic activity in the posterior part of pSTS of the sender’s brain (in red, MNI coordinates: 50, −42, 14, p < 0.05 corrected for multiple comparisons), as compared to planning similar movements during NON-COMMUNICATIVE trials (C); effect size: parameter estimates of a multiple regression analysis in standard error (SE) units), i.e. trials in which both players could see the goal configuration. The execution of the movements evoked no significant changes in the right pSTS activity of the sender’s brain (D). The observation of the same communicative actions increased right pSTS activity in the receiver’s brain (B), in red, MNI: coordinates: 56, −38, 6, p < 0.05 corrected for multiple comparisons), as compared to observing the same movements during non-communicative trials (E). The execution of the movements evoked no significant changes in the pSTS activity of the receiver’s brain (F).

FIGURE 3 | Planning-related activity in the right pSTS of the sender’s brain was modulated by the difficulty of selecting an adequate communicative behaviour. (A) When the sender could indicate the goal position of the receiver’s token (blue rectangle) by aligning his token (yellow rectangle) with her goal position, there was no significant right pSTS activity (in magenta). (B) When the sender could not match the goal orientation of the receiver’s token (blue rectangle) with his token (yellow circle), there was robust right pSTS activity (in orange), both for correct and incorrect trials. Other conventions as in Figures 1 and 2.

difference between easy and difficult was not so pronounced for the receiver (actually only significantly so for the number of moves). For the receiver the difference between trials was less pronounced than for the sender because in both trial types, the participant needed to move her token to a single target position on the board, and in this experiment the sender was a confederate, ensuring a reliable and homogeneous communicative behavior across trial types. The number of moves used by the receiver to provide the response...
Noordzij et al. Brain mechanisms underlying human communication

FIGURE 5 | Activity evoked in the receiver during EASY and DIFFICULT COMMUNICATIVE trials in the right pSTS (MNI coordinates: 56, −38, 6). The plots show effect sizes (in standard error units) of cerebral activity evoked in the receiver during the observation epochs of EASY and DIFFICULT COMMUNICATIVE trials. The effect sizes were larger for DIFFICULT than EASY COMMUNICATIVE trials.

DISCUSSION

We draw two main conclusions from these results. First, and most important, the findings indicate that the same cerebral region, the pSTS, is involved in recognizing communicative intentions and in planning communicative actions, supporting the hypothesis that both types of cognition involve similar conceptual inferences. These findings fit with known properties of this region, namely the involvement of the pSTS in inferring the intentionality of observed actions, both in humans and macaques (Barraclough et al., 2005; Jellema et al., 2000; Pelphrey and Morris, 2006). We extend the scope of those findings, showing that the contribution of the pSTS to intention recognition is not bound to the processing of biological cues, for example gaze or eye-hand joint movements (Barraclough et al., 2005; Puce et al., 1998), or point-light displays (Puce and Perrett, 2003). Crucially, here we show that the contributions of the pSTS extend beyond perceiving social cues, and include the generation of communicative actions. We suggest that this generative component is similar to what happens when people predict what type of action another person will do next (Aichhorn et al., 2006; Castelli et al., 2000; Jellema et al., 2000, 2004; Saxe et al., 2004). However, instead of making a prediction concerning the goals and intentions behind actions of others, a sender of a communicative action predicts the intentions that might be attributed by another person (i.e. the receiver) to that particular communicative action. In other words, pSTS is involved with the prediction of a forthcoming intention attribution, possibly on the basis of previous experience with how people have interpreted one's actions (Frith and Frith, 2006a; Schultz et al., 2005; Zilbovicius et al., 2006). Such a mechanism of predicting forthcoming intention recognition processes could explain the involvement of this region in different phenomena, from the recognition of biological motion (Allison et al., 2000), to the attribution of intentions (Castelli et al., 2000; Saxe et al., 2004), or the parsing of observed behaviors into conceptually relevant units (Zacks et al., 2001). Accordingly, it might not be surprising that developmental alterations in these basic perceptual mechanisms, as found in autism-spectrum disorder patients showed a significant difference between the two trial types, but this difference is trivially accounted by the additional rotations of the receiver's token required when solving the difficult trials.
(Dakin and Frith, 2005; Zilbovicius et al., 2006), have serious consequences for human social behavior. Further research will be needed to determine how biological motion and nonverbal communication are related. Perhaps, the same mechanisms putatively involved in the processing of biological motion, i.e. the extraction of posture sequences (Giese and Poggio, 2003), could also support the parsing of a string of object motions into communicative segments. In addition, although the present fMRI data support our apriori hypothesis of a computational overlap between recognizing communicative intentions and planning communicative actions, we cannot exclude the possibility that different parts of the pSTS support qualitatively different functions. In this respect, future single-subject electrophysiological studies might be able to test whether the same neuronal populations are involved in planning communicative acts and in recognizing the intentions conveyed by those acts, avoiding the anatomical scatter associated with the residual anatomical variability of fMRI comparisons (Petersson et al., 1999).

A second basic conclusion is that these findings do not support the view that the mirror-system provides the foundations for human communication (Rizzolatti and Craighero, 2004). The pSTS activity was indifferent to sensory input and motor output, but it was sensitive to the ambiguity in meaning of the communicative acts (Figure 3). In addition, the mirror-system regions hypothesized to provide efferent copies of motor commands to the pSTS (Iacoboni et al., 2001) showed increased de-activations during the planning of communicative actions (Figure 4B), an indication that generating intentional communicative behavior actually reduce metabolic activity (Shmuel et al., 2006) in the mirror-system (Rizzolatti and Craighero, 2004). These observations are not consistent with the idea that sensorimotor simulations can account for human communicative abilities. Nevertheless, it is theoretically conceivable that, in the pSTS, there is a non-linear relationship between number of moves/moving time and BOLD signal. In this case, our current approach would not capture this relationship, i.e. we cannot exclude the presence of higher order curvilinear relationships between motoric complexity and pSTS activity. However, given that these higher-order relationships would be devoid of any linear trend (since that would have been captured by our current analysis), it is not immediately obvious how they could be functionally interpreted.

Other interpretations of the findings seem ruled out by details of the experiment. For example, the pSTS response cannot simply reflect visual imagery of the movements, since it did not occur in non-communicative trials. Similarly, the pSTS response...
cannot be a consequence of its putative role in matching efferent copies of motor commands with visual inputs (Iacoboni, 2005; Iacoboni et al., 2001; Keysers and Perrett, 2004). That hypothesis predicts that the pSTS should be metabolically active during both action execution and action observation (Keysers and Perrett, 2004), but in this study the pSTS responses were strong during movement observation and absent during movement execution (Figures 2D,F).

It might be argued that the pSTS response reflects subjects’ verbalizations, but the right-hemispheric lateralization of the effect (Figures 2A and 4C) is not consistent with the left-hemispheric dominance for phonological and syntactic processing (Frost et al., 1999). Actually, the present findings confirm the crucial role of the right pSTS for processing pragmatic aspects of linguistic material (Jung-Beeman et al., 2004; Mashal et al., 2007), an instance of the right-hemispheric dominance for inferring the communicative intentions of a conversational partner (Sabbagh, 1999). It might be argued that the differential pSTS response during easy and difficult communicative trials (Figures 3 and 5) is driven by differential communicative success (Figure 8). However, this interpretation is not consistent with the finding that both correct and incorrect outcomes evoke significantly positive and equivalent responses in the right pSTS of senders during difficult communicative trials (Figure 3B). In fact, analysis of the senders’ movements during communicative trials suggests a different possibility. In the easy communicative trials, senders generated the same type of communicative action (namely, move to the target position of the receiver, pause, and then move to his own target position). This behavior was consistent across trials and across participants, and the senders developed it during the task familiarization period prior to the scanning session. In contrast, during the difficult trials, the sender generated different communicative movements according to the geometry of his token, of the receiver’s token, and the past history of communicative trials. Accordingly, we suggest that the right pSTS distinguishes between trials requiring the formation of new semiotic conventions (difficult communicative trials), and trials exploiting a recently established communicative behaviour (easy communicative trials). Finally, it should be emphasized that our findings pertain to the generation of non-verbal communicative behaviors that are independent from shared conventional codes. It could be argued that this experimentally controlled scenario is artificial, failing to capture relevant aspects of human linguistic
communication. However, we all start out as infants without access to the local communication conventions. In that respect, by addressing the human ability to quickly build new semiotic conventions, we deal with a crucial pre-condition for using gestures and language as a communicative tool (Galantucci, 2005; Levinson, 2006). Accordingly, the current findings are in line with ancillary evidence that points to distinct origins, in both ontogeny and phylogeny, of the foundational mechanisms for communication on the one hand, and language on the other (Levinson, 2006; Tomasello and Carpenter, 2005).

It has been argued that communicative actions are special (Frith, 2007). Their planning is special, since the immediate goal of the movement does not overlap with its actual communicative purpose. Their understanding is special, since they rely on the ability to recognize their communicative value, such that sender and receiver can share an intention (Tomasello et al., 2005). The present findings provide a cognitive and cerebral ground for this special human faculty, supporting the notion that our communicative abilities are distinct both from sensorimotor processes (with distinct areas of activation) and language abilities (with their largely left-hemisphere localization).

**Acknowledgements**

The present study was supported by the EU-Project “Joint Action Science and Technology” (IST-FP6-003747). We would like to thank Roger Newman-Norlund, Bram Daams, and Paul Gaalman for technical advice and assistance.

**References**


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.