ORIGINAL ARTICLE

Seeing the Unexpected: How Brains Read Communicative Intent through Kinematics

James P. Trujillo[1,2], Irina Simanova[1], Asli Özyürek[2,3] and Harold Bekkering[1]

[1]Donders Institute for Brain, Cognition and Behaviour, [2]Centre for Language Studies, Radboud University Nijmegen, 6500HD Nijmegen, the Netherlands and [3]Max Planck Institute for Psycholinguistics, 6525XD Nijmegen, the Netherlands

Address correspondence to James P. Trujillo, Radboud University, Montessorilaan 3, B.01.25, 6525GR Nijmegen, the Netherlands. Email: jptrujillo88@hotmail.com

Abstract

Social interaction requires us to recognize subtle cues in behavior, such as kinematic differences in actions and gestures produced with different social intentions. Neuroscientific studies indicate that the putative mirror neuron system (pMNS) in the premotor cortex and mentalizing system (MS) in the medial prefrontal cortex support inferences about contextually unusual actions. However, little is known regarding the brain dynamics of these systems when viewing communicatively exaggerated kinematics. In an event-related functional magnetic resonance imaging experiment, 28 participants viewed stick-light videos of pantomime gestures, recorded in a previous study, which contained varying degrees of communicative exaggeration. Participants made either social or nonsocial classifications of the videos. Using participant responses and pantomime kinematics, we modeled the probability of each video being classified as communicative. Interregion connectivity and activity were modulated by kinematic exaggeration, depending on the task. In the Social Task, communicativeness of the gesture increased activation of several pMNS and MS regions and modulated top-down coupling from the MS to the pMNS, but engagement of the pMNS and MS was not found in the nonsocial task. Our results suggest that expectation violations can be a key cue for inferring communicative intention, extending previous findings from wholly unexpected actions to more subtle social signaling.

Key words: communication, connectivity, fMRI, gesture, intention

Introduction

In order to successfully interact with others, it is important to understand their social and communicative intentions. The human brain is remarkable in its ability to attribute goals and intentions to actions, allowing us to interpret not only what a person is doing (i.e., the concrete intention) but also why they are doing it (i.e., the abstract intention) (Van Overwalle 2009). For example, as a customer lifts a glass, the waiter can predict whether the customer is going to drink from the glass or uses this act as a request to have another drink. In this example,

the social or communicative intention of the actor must be quickly read from their motor behavior (Blakemore et al. 2001). An interesting question is how the brain picks up on the subtle, socially relevant modulation of the motor act to accomplish this abstract intention reading.

Previous research suggests that humans modulate the kinematics of their movements based on high-level, abstract intentions (Becchio et al. 2012; Pezzulo et al. 2013). For example, when an object-directed action is produced with a communicative intention, the kinematic profile of the action is quantitatively different from when the same action is produced without or

with a different degree of communicative intention (Sartori et al. 2009; Campisi and Özyürek 2013). In a previous behavioral study, we quantified the differences in kinematics of motor acts produced in a more- compared with less-communicative context. We found that, in actions and gestures, both spatial and temporal kinematic features were modulated, becoming more exaggerated in the more-communicative context (Trujillo et al. 2018). Furthermore, we found that observers were able to read this communicative intent from the actors' movement kinematics (Trujillo et al. 2018). These results are well in line with previous suggestions that humans are able to use differences in kinematic profiles in order to infer an underlying intention (Becchio et al. 2012).

The ability to read intentions from movement kinematics has been shown both for concrete end-state intentions, for example, grasp to drink versus grasp to pour (Cavallo et al. 2016; Becchio et al. 2018), and for more abstract social intentions, for example, engaging in a social task (Manera et al. 2011; Trujillo et al. 2018). It has been suggested that the end-state intentions may be read by directly mapping the kinematics onto actions in our own motor repertoire (Blakemore et al. 2001; Rizzolatti et al. 2014; Cavallo et al. 2016). While direct mapping could work for concrete (action end-state) intentions, it is less clear how we read more abstract (i.e., high level) social intentions that may not have a direct mapping. Abstract intentions are more difficult due to the necessity of having a mapping of all potential socially modulated forms of every action.

A potential solution is to infer intentions based on whether the action follows a typical, expected kinematic pattern or not. This follows from literature describing how we ascribe high-level intentions to movements that are otherwise unusual or implausible, given the context, as a way to rationalize them (Gergely and Csibra 2003; Brass et al. 2007; Csibra and Gergely 2007). For example, when we see someone activating a light switch with their knee, we may rationalize this as being due to their hands being occupied by a heavy stack of books (Brass et al. 2007). In this way, we explain away the unusual movement as being due to the observable context. In the case of communicatively intended acts, the exaggerated kinematics would be inconsistent with how an observer expects the action to be produced according to previous experience, resulting in the observer attributing a more abstract intention to the actor. This is consistent with the theory of sensorimotor communication (Pezzulo et al. 2013), which suggests that movements can be made communicative by deviating from the most optimal way of performing the action. This also fits with previous results showing that kinematically inefficient movements are seen as unexpected (Hudson et al. 2018). This framework would predict that we do not understand by mapping the observed kinematics to our own motor system but rather by actively inferring a hidden intention that would explain the unusual movement.

In the brain, processing abstract intentions typically involves the mentalizing system (MS) (Kampe et al. 2003; Frith and Frith 2006; Spunt et al. 2011; Ciaramidaro et al. 2013). At the same time, a meta-analysis by Van Overwalle and Baetens suggests that the brain likely utilizes the motor system to understand what the observed action is together with the MS to process the intention (Van Overwalle and Baetens 2009). This is especially important when considering the case of communicative kinematic modulation. If we are to read the underlying intention from kinematic modulation alone, we must first recognize that the action is being performed in an unusual or exaggerated fashion. Recognizing the act as unusual likely involves the putative mirror neu-

ron system (pMNS) (Newman-Norlund et al. 2010) attempting to match the observed action with one already in the observer's motor repertoire (Kilner et al. 2007). The exaggerated kinematics would therefore elicit a breach of expectation, resulting in the recruitment of the MS to process the underlying intention that generated the unusual behavior (Brass et al. 2007; de Lange et al. 2008; Schiffer et al. 2014). The recruitment of the pMNS and MS in response to unusual movements and the reading of intentions has been shown previously, utilizing movements that are unusual given their end goal (e.g., using one's knee to activate a light switch) and context (e.g., whether one's hands are free). Distinctly unusual kinematics, specifically in terms of movement trajectory, have also been shown to recruit pMNS and MS regions (Marsh et al. 2011, 2014). This suggests that observers are sensitive to the rationality or efficiency of movement, and unexpected kinematics may lead to intention inferences. However, these studies did not explicitly test whether brain response scales with unexpectedness or inefficiency of the movement kinematics.

Here, we specifically investigate the question of whether a difference in the intention to communicate can be recognized from the kinematics provided. As kinematic modulation is a relatively subtle intentional signal based purely in movement, testing the recruitment of the pMNS and MS in recognizing abstract intention provides a direct test of this model of intention reading.

Processing of abstract intentions in the pMNS and MS is likely achieved via an interaction between the 2 systems. This is because the 2 systems are often not activated concurrently. Instead, studies of intention recognition often show activation of either the pMNS or the MS, but not both for the same task, suggesting that information likely flows from one to the other when both are needed. The results from Van Overwalle and Baetens (2009) seem to suggest that this process would be bottom-up, with the pMNS influencing the MS when breaches of movement expectation are encountered. In this framework, expectations originate in the premotor cortex (PMC), and the MS is recruited to resolve these breaches of expectation. An alternative account is the predictive coding framework (Kilner et al. 2007). This framework suggests that high-level expectations, originating in this case in the MS, might influence lower level expectations, such as movement expectations (Ondobaka et al. 2015). Although the theoretical framework of predictive coding computationally predicts bidirectional influence (i.e., top-down and bottom-up), experimental work seems to primarily find top-down modulation (Chennu et al. 2016; Chambon et al. 2017). This is particularly the case when participants are actively attending to the unexpected stimulus (Chennu et al. 2016). This would argue for a stronger top-down influence, with the MS primarily influencing the pMNS. This account is supported by findings from studies of perceptual breaches of expectation, where unexpected changes in auditory stimuli (Chennu et al. 2016) as well as the processing of more abstract intentions (Chambon et al. 2017) result in modulation of top-down connectivity strength. It is therefore necessary to investigate directional connectivity in order to understand how the 2 systems interact when reading abstract (e.g., communicative) intentions from movement.

An important aspect of previous studies on intention recognition is the role of context. For example, in the study by Brass et al. (2007), the unusual action of turning on a light switch was informed by the presence of a stack of folders that the actor was holding. The act itself was of course unusual due to the effector used (i.e., the knee, rather than the hand) to complete

the action. Similarly, intention may be largely inferred from the combination of action and object. For example, picking up an apple and extending it towards the viewer is likely to be seen as communicatively or socially intended, whereas picking up a book and opening it directly in front of one's self is seen as privately or personally intended (Ciaramidaro et al. 2007). In order to understand how kinematics can inform intention recognition, we must therefore disentangle subtle, communicatively intended kinematic modulation from other visual contextual cues.

Finally, it is important to address the effect of exogenous cues on intention recognition. While it is clear that observers can read even abstract intentions from movement kinematics, this inference on the underlying intention is not likely to be actively made under all circumstances (de Lange et al. 2008; Spunt and Lieberman 2013). Instead, intention inferences may only be made when it is task-relevant. However, it is possible that the brain responds in a similar way even when the intention is not being attended. Therefore, testing whether activation and connectivity changes are dependent on the presence of explicit task instructions would indicate whether the brain responds implicitly to communicative cues in movement kinematics.

Current Study

This study aims to determine the neural systems and mechanisms underlying the recognition of communicative intention at the level of movement kinematics. Particularly, we 1) test whether communicative kinematic modulation results in activation of the pMNS and MS and 2) determine whether there is evidence for a top-down or bottom-up interaction between the systems. We additionally will determine whether there is evidence for implicit processing of abstract intentions from kinematic modulation alone. We further build on previous studies by investigating whether this neural mechanism of intention inference also holds for more complex movement sequences such as representational gestures (i.e., movements that visually simulate a manual action).

We address these issues using 2 forced-choice gesture viewing tasks during functional magnetic resonance imaging (fMRI). In the 2 tasks, participants viewed stick-light figures created in a previous study where we measured the kinematics of more- and less-communicative gestures (Trujillo et al. 2018). In one task, the Social Task, participants were asked after each video if they believe the action being depicted in the video was intended for the actor or the viewer (representing more- and less-communicative intentions). In the other task, the (Nonsocial) Handedness Task, participants saw the same videos but were asked to decide whether the action being depicted was performed with the left hand or the right hand. Using participant responses, we calculated the average perceived communicativeness of the kinematic modulation in each of the videos. By correlating this value with fMRI blood oxygen level–dependent (BOLD) response, we calculated the extent to which brain activation increases with increasingly communicative kinematics. We therefore use kinematics to provide an extension of the abstract intention inference model beyond the perception of purely categorical, contextually embedded stimuli. We further specify the model by assessing whether communicative kinematic modulation affects top-down or bottom-up information flow between the systems (effective connectivity analysis). Finally, as a secondary analysis, we use the Handedness Task to determine whether the neural

response to communicative kinematics is dependent on task instruction (secondary task analysis).

Methods

Participants

Twenty-eight participants took part in this study, recruited from the Radboud University. Participants were recruited with the criteria of being between the ages of 18 and 35 years, being right-handed, with correct or corrected-to-normal vision, being native speakers of Dutch, and with no history of psychiatric or communication impairments. One participant was excluded due to an error in the projection of stimuli, resulting in a difference in size in the projection. One additional participant did not complete the first task due to discomfort in the scanner. This led to a total sample size of 26 participants (11 male) with a mean age of 25.10 years. The procedure was approved by a local ethics committee.

Materials

Kinematic Feature Quantification
The current study used the same kinematic features quantified in Trujillo et al. (2018). We used a toolkit for markerless automatic analysis of kinematic features, developed earlier in our group (Trujillo et al. 2019). The following briefly describes the features of quantification procedure: All features were measured within the time frame between the beginning (hands start to move) and ending (hands no longer moving) of the gesture. This was the same method used by Trujillo et al. (2018), allowing us to more faithfully replicate behavioral findings and ensuring the kinematic features represent the movement in the entirety of the video. Motion-tracking data from the Kinect provided measures for our kinematic features: "Distance" was calculated as the total distance traveled by both hands in 3D space over the course of the item. "Vertical amplitude" was calculated on the basis of the highest space used by either hand in relation to the body. "Peak velocity" was calculated as the greatest velocity achieved with the dominant hand. "Hold time" was calculated as the total time, in seconds, counting as a hold. Holds were defined as an event in which both hands and arms are still for at least 0.3 s. "Submovements" were calculated as the number of individual ballistic movements made, per hand, throughout the item. Ballistic movements were calculated using a peak analysis, similar to the description of submovements given by Meyer et al. (1988). In line with the Trujillo et al. (2018) study, our peak analysis used a velocity threshold of 0.2 m/s, a between-peak distance of 8 frames, and minimum peak height and prominence of 0.2 m. To account for the inherent differences in the kinematics of the various items performed, z-scores were calculated for each feature/item combination across all actors including both conditions. This standardized score represents the modulation of that feature, as it quantifies how much greater or smaller the feature was when compared with the average of that feature across all of the actors. This means that high z-score values for a video indicate that the kinematics were significantly larger than what is typical for that action. For a more detailed description of these quantifications, see Trujillo et al. (2018).

Stimuli
We included 120 videos recorded in a previous study (Trujillo et al. 2018). In this previous study, 40 participants performed

31 different representational (pantomime) gestures. Twenty performed the gestures in a less-communicative context, while the other 20 performed them in a more-communicative context. Motion capture data of participants (henceforth actors) in this previous experiment were captured using Microsoft Kinect while the actors were seated at a table. The gestures were pantomime versions of object-directed actions, such as cutting paper with scissors or peeling a banana. For each act, actors began with their hands placed on designated starting points on the table, marked with tape. Target objects were placed on the table (e.g., scissors and a sheet of paper for "cutting paper with scissors"), but actors were instructed beforehand not to actually touch the objects. After placing the object(s) on the table, the experiment moved out of view and recorded instructions were played in Dutch (e.g., "knip het papier doormidden met de schaar" ["cut the paper with the scissors"]). Immediately following the instructions, a bell sound was played, indicating that the actor could start performing the gesture. Once the act was complete, the hands returned to the starting points, after which another bell sound indicated the end of the trial. The more-communicative context was elicited by introducing a confederate who sat in an adjacent room and was said to be watching through the video camera and learning from the participant. In this way, an implied communicative context was created. The same procedure was applied to the less-communicative context, except the confederate was said to be learning the experimental setup. The less-communicative context was therefore exactly matched, including the presence of an observer, but only differed in that there was no implied interaction.

In order to provide a representative sample of the videos, we first ranked all videos according to the overall kinematic modulation (z-scores derived from the kinematic features described in the Stimuli section) and the communicative context (more or less communicative). This placed all of the videos on a continuum from low kinematic modulation, as was typical of the less-communicative videos, up to high kinematic modulation, as seen in the more-communicative videos. We then selected 60 more-communicative videos, favoring high z-scores, and 60 less-communicative videos, favoring low z-scores, on the basis of keeping the 2 contexts matched in all raw kinematic (i.e., non-modulation) values as well as overall duration, while also keeping the modulation values of all kinematic features significantly different. This was done using standard t-tests on the raw and modulation values. Therefore, the more-communicative videos were primarily characterized by high positive z-scores, and less-communicative videos were characterized by high negative (e.g., slower, smaller than typical) z-scores. Once a suitable selection was made, the selected videos were transformed into stick-light figures based on the Kinect motion capture data (see Fig. 1 for still frames). This ensured that the visual information being processed while viewing the videos was identical besides the movements, or kinematics, of the act.

Physical Setup and Briefing
Participants were informed that they would be viewing short videos of actions being depicted by "stick figures," which were created from the motion capture data of real participants in a previous experiment. They were informed that half of the participants performed the actions for themselves, and the other half performed them explicitly for someone else. We informed the participants that, in their first task, they should try to guess if each action was performed for the actor or for the viewer and, in the second task, they should try to determine if the actions were performed more with the left hand or the right hand. The Social Task was always given first, followed by the Handedness Task. The ordering was fixed to ensure that the stimuli were novel during the Social Task.

Participants were positioned in the supine position in the scanner with an adjustable mirror attached to the head coil. Through the mirror, participants were able to see a projection screen outside the scanner. Participants were given an MRI-compatible response box, which they were instructed to operate using the index finger of their right hand to press a button on the right and the index finger of their left hand to press a button on the left. Button locations corresponded to response options given on the screen, which always include 2 options: one on the left of the screen and one on the right of the screen. The resolution of the projector was 1024×768 pixels, with a projection size of 454×340 mm and a 755-mm distance between the participant and the mirror. Video size on the projection was adjusted such that the stick figures in the videos were seen at a size of 60×60 pixels. This ensured that the entire figure fell on the fovea, reducing eye movements during image acquisition. Stimuli were presented using an in-house developed PsychoPy (Peirce et al. 2019) script.

Tasks
Social Task. The Social Task was designed to explicitly elicit intention recognition by attending to the movements. In this task, participants first saw a Dutch action verb that served as a linguistic prime for the upcoming video. This was provided to ensure participants understood the gesture that they were seeing. Next, there was a 3.5-s fixation cross, with a 1.5-s jitter. Participants were then presented with the stick-light gesture. The average duration for these videos was 6.34 s. After the video completed, participants were then visually presented with the question of whether the action was intended for the actor or the viewer. The 2 options were presented on random sides of the screen, and participants responded by pressing either the left or right button of the response box. No feedback was given regarding the accuracy of the response. The order of videos was randomized for each participant.

Handedness Task. The Handedness Task was designed so that participants would attend to the movements without any social or communicative implication, allowing us to test for evidence of automatic processing of intention. This task followed the same procedure, with a new randomized order of stimuli. However, in this task, participants were asked whether the action was performed with the left hand or the right hand. See Figure 2 for a schematic timeline of one trial.

Behavioral Data

Data Preparation and Implementation
Response time (RT) and intention classification were utilized for analyses. Data were first checked for outliers at the participant level in terms of RT, with outliers considered to be more than 2.5 standard deviations above the group mean. This led to a removal of 73 individual trials in the Social Task and a removal of 76 trials in the Handedness Task. All preparatory procedures and statistical tests were carried out separately for the Social and Handedness Tasks. All testing of behavioral data was performed using the R statistical program (R Core Team, 2017). Mixed-effects modeling utilized the lme4 package (Bates et al. 2014), and *P* values were estimated using the Satterthwaite approxi-
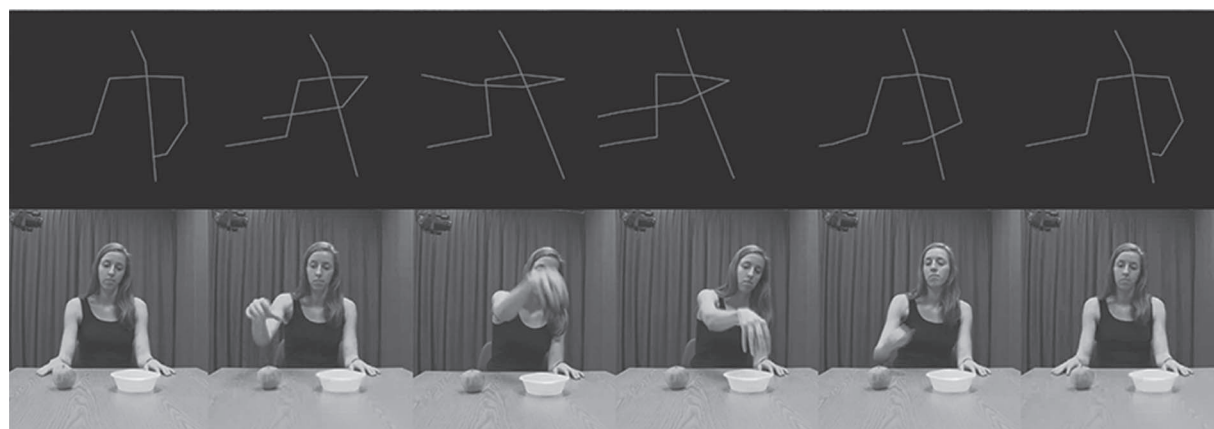
Figure 1. Still frames of a stick-light figure and a comparison with the corresponding video images. The lower panel depicts a series of still frames from one of the videos recorded in Trujillo et al. (2018) at various stages of action completion. The upper panel depicts the corresponding stick-light figure derived from the kinematics of this action. Note that the images in the upper panel represent what was seen by participants, who had no exposure to the video images. Figure was adapted with permission from Trujillo et al. (2019).
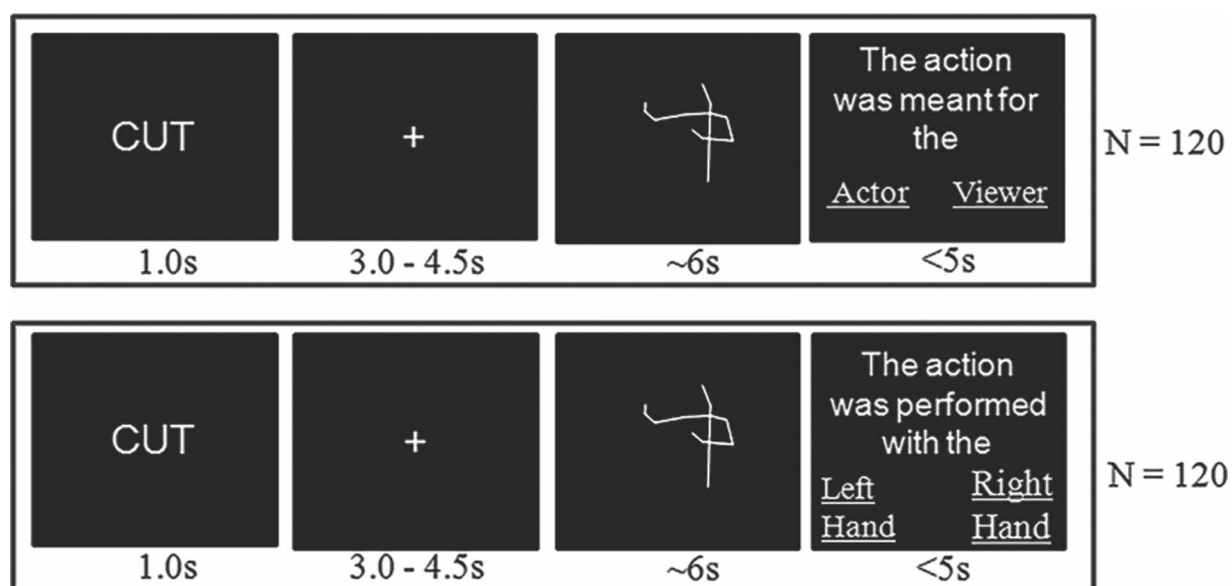

Figure 2. Overview of trial progression. The upper panel depicts the Social Task, while the lower panel depicts the Nonsocial Handedness Task. Participants first saw a single prime word, followed by a fixation cross of variable length, then the video, and, finally, the task-specific response screen.

mation of denominator degrees of freedom, as implemented in the lmerTest package (Kuznetsova 2016).

Statistical Analyses

Social Task. Statistical analyses were carried out in order to assess whether kinematic modulation was correlated with intention classification. Note that we did not test whether classification decisions matched the context labels from the previous study (Trujillo et al. 2018). This is because the primary interest of the study was the spectrum of kinematic modulation, rather than the initial categories that are also highly variable.

We used linear mixed-effects modeling to determine the correlation between kinematic features and intention classification. Kinematic modulation values were entered into the model as fixed effects with the classification decision (communicative, for the viewer, or noncommunicative, for

the actor) as the dependent variable. In the first model, participant was additionally included as a random intercept variable, allowing us to control for individual variation between participants. We used a χ^2 test to determine if this model better explained the data than a null model in which only participant variation was given as an explanatory (independent) variable. Next, we compared our initial model with a more complex model that additionally included actor and action as random intercepts. This model was again tested against the null and initial models to determine which provided the best explanation of the data using χ^2 tests. Only fixed effects results from the winning model are interpreted. To reduce the risk of Type I error, we used the Simple Interactive Statistical Analysis tool (http://www.quantitativeskills.com/sisa/calculations/bonfer.htm) to calculate an adjusted alpha threshold based on the mean correlation between all of the tested kinematic features, as

well as the number of tests (i.e., number of variables in the mixed model). Our 4 variables (vertical amplitude, peak velocity, submovements, and hold time) showed an average correlation of 0.063, leading to a Bonferroni-corrected alpha threshold of 0.013.

Handedness Task.   Statistical analyses were carried out in order to assess whether participants were attending to the movement kinematics. This ensures that our fMRI results reflect only a difference in the task, rather than the stimuli, which participants should be attending to similarly in both the Social and Handedness Tasks.

We used linear mixed-effects modeling following the same procedure described for the Social Task. The only difference was that we included peak velocity and submovements for the left hand and excluded vertical amplitude and hold time. This was done due to vertical amplitude and hold time being features that were quantified from both hands. Therefore, we included the single-hand features for both right and left in order to test the hypothesis that participants classified the handedness of the videos according to hand-specific features. In other words, we assume that right-handed classifications will be made based on submovements and/or peak velocity of the right hand if participants are attending to the kinematics.

We again calculated an adjusted alpha threshold based on the mean correlation of the tested kinematic features and the number of tests (again 4). The 4 variables in this model set (right peak velocity, right submovements, left peak velocity, and left submovements) showed a mean correlation of 0.138, leading to a Bonferroni-corrected alpha threshold of 0.015.

Calculation of "Communicativeness" Metric

In order to test our hypothesis that the communicative quality of movement kinematics would be correlated with hemodynamic response in the mirroring and mentalizing systems, we used the behavioral data to calculate a metric of how communicative each video was. In order to calculate this communicativeness value, we first calculated a new mixed-effects model with intent classification as the dependent variable; vertical amplitude, hold time, peak velocity, submovements, and RT as fixed effects predictors; and actor, action, and participant as random intercepts. RT was included in this model as a measure of certainty, allowing us to capture not only the effect of the kinematics on the final classification decision of the participants but also how quickly the participants made this decision. Finally, we used this model to calculate the mean predicted probability of judging each video as communicative. As the predicted probability serves as a measure of how likely a new participant would be to judge a video as communicative, this is taken to represent a quantification of video communicativeness. The process of calculating the predicted probability was carried out in a leave-one-out manner, where the values were calculated separately for each individual participant, based only on the rest of the participants' response data. For example, to calculate the communicative values that would be used to model participant 5's brain response, we used the response data from participants 1–4 and 6–26 to calculate a mean value for each video. Participant 5's data are thus not included in the calculation of her own fMRI regressors. This was repeated for each participant. This was done to prevent overfitting the data. In the end, each participant had a unique set of communicativeness values assigned to the videos, with one value per video. The *communicativeness* metric therefore provided a single value for each video that described, based on

participant responses and the underlying kinematic modulation values, the probability that the video would be classified as being communicatively intended when viewed by a new, naïve participant. These values were then used to model the fMRI data at the first (subject) level.

Brain Imaging

fMRI Data Acquisition
Anatomical and task-related MRI images were acquired on a 3-T Siemens Magnetom Skyra MR scanner with a 32-channel head coil at the Donders Institute for Brain, Cognition and Behaviour in Nijmegen, the Netherlands. Structural images (1 $\times$ 1 $\times$ 1 mm^3) were acquired using a T_1-weighted magnetization prepared rapid gradient echo sequence with time repetition (TR) = 2300 ms, time echo (TE) = 3.03 ms, flip angle = 8°, and field of view (FOV) = 256 $\times$ 256 $\times$ 192 mm^3. Two behavioral tasks (described below) were carried out by participants while T_2*-weighted dual-echo echo-planar imaging (EPI) BOLD-fMRI images were acquired using an interleaved ascending slice acquisition sequence (slides = 40, TR = 730 ms, TE = 37.8 ms, flip angle = 90°, voxel size = 3 $\times$ 3 $\times$ 3, slice gap = 0.34 mm, FOV = 212 $\times$ 212 mm^2).

fMRI Analysis—General Linear Model
All analyses were performed using SPM12 (Statistical Parametric Mapping; Wellcome Department, http://www.fil.ion.ucl.ac.uk/spm). All functional data were preprocessed following the same pipeline: Functional and structural images were realigned and coregistered, with spatial normalization with the Montreal Neurological Institute (MNI) template and spatial smoothing using an 8-mm full-width at half-maximum kernel. After preprocessing, we checked motion parameters in the task-related acquisitions to ensure that no participants moved more than 3° in rotation or 3 mm in translation.

We created an event-related design matrix for within-subject first-level analysis, wherein we modeled the video-viewing period, response, and fixation as separate regressors. Communicativeness of the videos was added as a parametric modulator, with the values convolved with the video-viewing events in a separate regressor. Finally, the 6 motion parameters were added as regressors of no interest. Our primary first-level contrast was communicativeness over baseline, which effectively modeled a linear correlation between the BOLD signal and the communicativeness score. The 2 tasks were modeled in separate design matrices, with no direct comparisons between the 2. This is because the Handedness Task was only used to test whether brain activation or connectivity is related to kinematic modulation when the task does not require a communicative intent decision.

Contrast images from the first-level analysis were used in the second (group) level analysis, using whole-brain voxel-wise *t*-tests. Contrast maps were thresholded at $P < 0.001$, uncorrected, with cluster threshold set as $k > 10$.

fMRI Analysis—Dynamic Causal Modeling
General overview.   We used dynamic causal modeling (DCM; Friston et al. 2003) in order to quantify how the mentalizing and mirroring systems interact during intention understanding. DCM allows the researcher to define a subset of brain regions and their connections and model how the activity of the regions or strength of the connections is dependent upon an experimental manipulation. After building and estimating a
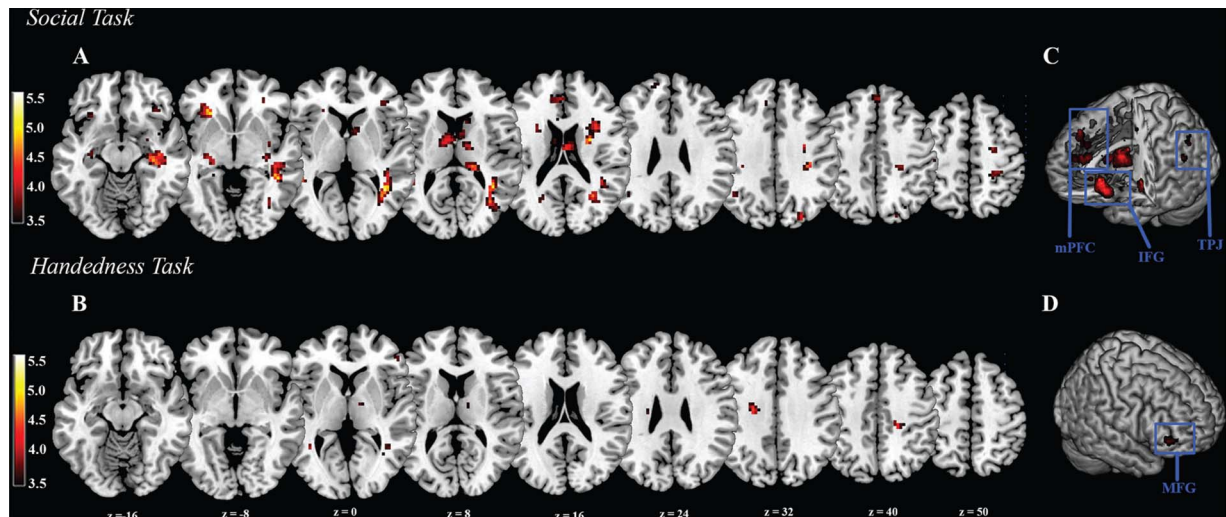
Figure 3. Overview of GLM results. The top panels (*A, C*) depict slices from the Social Task, while the bottom panels (*B, D*) depict the Handedness Task. Red areas indicate a significant (*P* < 0.001) correlation between BOLD response and video communicativeness. The red color bars show the corresponding *T* values. Panels *A* and *B* provide a slice-by-slice overview of the 2 tasks, while panels *C* and *D* provide a 3D rendering of the same data, with significant areas of interest highlighted (IFG = inferior frontal gyrus; MFG = middle frontal gyrus).

set of potential causal models, a model selection analysis is performed in order to find the model that represents the best fit to the data. In order to keep the models relatively simple and balanced, we opted to only model 2 regions: one from the MS and one from the mirroring system. We based our initial selection criteria on the meta-analysis of intention understanding by Van Overwalle and Baetens (2009), which lists the posterior superior temporal sulcus (pSTS), anterior inferior parietal sulcus (aIPS), and PMC as the primary mirroring system regions and the temporoparietal junction (TPJ) and medial prefrontal cortex (mPFC) as the primary MS regions. As the TPJ, aIPS, and pSTS show some degree of overlap, we chose not to use these regions and therefore selected the PMC as the representative mirroring region and the mPFC as the representative mentalizing region to contrast the 2 networks in a neuroanatomically optimal manner.

Regions of interest. We defined the location of these group-level regions of interest around the peak-voxel coordinates of our second-level communicativeness contrast from the Social Task. Functional regions were defined from the coordinates based on the definitions by Lacadie et al. (2008). Note that the same coordinates were used in our DCM analysis of the Handedness Task in order to ensure a direct comparison of the results and that this analysis is carried out regardless of general linear model (GLM) results of the Handedness Task as this was an a priori planned analysis in order to compare against the Social Task. The PMC was located at $x = 24$, $y = -10$, $z = 53$, while the mPFC was located at $x = -9$, $y = 38$, $z = 23$. The coordinates were used as starting points to locate subject-specific regions. This was done using SPM12's volume-of-interest utility, which takes a starting coordinate and moves it, per participant, to the nearest peak voxel within a 5-mm range. This method takes individual variation in functional neuroanatomy into account and increases sensitivity of subsequent analyses. Each newly assigned peak was manually checked to ensure that it still was in the designated region. Mean time courses were extracted from a 10-mm sphere surrounding the peak coordinate, using the

communicativeness contrast and a liberal threshold of $P < 0.100$ to ensure a robust estimate of the time series.

Model space. We created an initial model composed of the PMC and mPFC with bidirectional intrinsic connections. The video-viewing event (video onset, with length equal to video duration) was modeled as a possible direct, or driving, influence on regional activity, while the communicativeness regressor (as explained under the Calculation of "Communicativeness" Metric section) was defined as a possible modulating influence on the strength of interregion connections. By varying the presence of the driving and modulation influences on the 2 regions and connections, we created 14 models that included all possible combinations of these influences, including one fully parameterized model that had both driving influences and both modulations as well as one "null" model that had no influence from the task. See Supplementary Figure 1 for a schematic overview of all these models.

Model selection. Bayesian model selection was used to test the probability of our data given each of the models. As our participants are relatively homogeneous (i.e., no group-based inferences), we utilized a fixed effects approach. A posterior probability of >0.95 was taken to be a strong evidence in favor of a particular model.

Results

Behavioral Results—Social Task

For the Social Task, we tested whether higher kinematic modulation values predicted classification of an act as being communicative. In line with our hypothesis, our mixed-effects regression model containing the kinematic features as fixed effects predictors was a better fit to the data than the null model that did not contain kinematics, $\chi^2(4) = 51.629$, $P < 0.001$. Adding actor and action as random intercepts further improved model fit, $\chi^2(2) = 18.605$, $P < 0.001$. All results at the kinematic feature level are therefore based on the full model, including all kine-

matic modulation values as fixed effects as well as participant, actor, and action as random intercepts. In terms of kinematic features, we found that increased vertical amplitude ($z = 4.113$, $P < 0.001$) and hold time ($z = 3.243$, $P = 0.001$) were significantly predictive of classifying an act as communicative. An increased number of submovements showed a near-significant relation to intent classification ($z = 2.432$, $P = 0.015$), while peak velocity was not related to communicative intent classification ($z = 0.924$, $P = 0.356$). Results therefore confirm that intention classification was related to kinematic modulation.

Behavioral Results—Handedness Task

For the Handedness Task, we tested whether higher kinematic modulation values of a particular hand predicted classification of an act being performed more with that same hand. This was to ensure participants were attending to the kinematics in this task. We found that the model containing kinematic modulation values was a better fit to the data than the null model, $\chi^2(4) = 83.291$, $P < 0.001$. Adding actor and action to the model further improved model fit, $\chi^2(2) = 368.57$, $P < 0.001$. All results at the kinematic feature level are therefore based on the full model, including all kinematic modulation values as fixed effects as well as participant, actor, and action as random intercepts.

In terms of kinematic features, we found that submovements of the right hand were predictive of classifying an act as being more right-handed ($z = 5.143$, $P < 0.001$). We found no association between handedness classification and submovements of the left hand ($z = -1.676$, $P = 0.094$), peak velocity of the right hand ($z = 1.817$, $P = 0.069$), or peak velocity of the left hand ($z = 1.643$, $P = 0.100$). Results therefore confirm that participants attended to kinematic modulation also during the Handedness Task, while further suggesting that the right hand was attended to primarily.

Whole-Brain Results—Social Task

Whole-brain results reflect BOLD correlation with video communicativeness. Results of the whole-brain analysis of the Social Task show primarily regions associated with the pMNS, such as the right PMC and right inferior parietal lobe, as well as regions associated with the MS, such as the left mPFC and left TPJ. We additionally found activation in the left inferior frontal gyrus, left caudate nucleus, right hippocampus, and several areas of the cerebellum. See Figure 3A for a graphical overview of these results. Table 1 provides an overview of peak coordinates, given in MNI space, with statistics and cluster sizes. All regions were significant at $P < 0.001$.

Whole-Brain Results—Handedness Task

Results of the whole-brain analysis of the Handedness Task show only the middle frontal gyrus being correlated with communicativeness. See Figure 3B for a graphical overview of these results. See Table 1 for peak coordinates and statistics.

Connectivity Results—Social Task

In the Social Task, we found strong evidence (posterior probability = 1.00) for a model with no driving effects of video viewing on the PMC or mPFC but modulation of the top-down (mPFC → PMC) connection. See Figure 4 for a schematic overview of the winning model and the exceedance probability.

Connectivity Results—Handedness Task

In the Handed Task, we did not find evidence above our defined probability threshold. However, 2 models together showed a posterior probability of 1.00. The model with the highest evidence (posterior probability = 0.561) showed driving influence of video viewing on the PMC and modulation by communicativeness of the videos on the bottom-up (PMC → mPFC) connection. The second model (posterior probability = 0.439) showed no driving effects but modulation by communicativeness of the bottom-up connection. Together, this can be taken as strong evidence in support of modulation of the bottom-up connection, with weaker support for the driving effect on the PMC. See Figure 4 for a schematic overview of the 2 models and the exceedance probabilities associated with them.

Discussion
General Overview of Findings

This study set out to test the brain activation and connectivity during the recognition of communicative intentions from kinematic modulation. We found that 1) participants recognize communicative intent based on spatial and temporal kinematic features if explicitly asked to classify intentionality, 2) the perceived communicativeness of the videos correlates with activation of the mentalizing and mirroring systems when this is task-relevant, and 3) top-down connectivity between these systems is altered by communicativeness in the Social Task, while bottom-up connectivity is modulated in the Nonsocial Task.

Behavioral Results

Our behavioral results show that our participants were able to utilize kinematic modulation in their intention classifications. This result is a direct replication of earlier work from our group that showed that increased vertical amplitude was perceived as communicative (Trujillo et al. 2018). The current study replicated this finding while extending it in 2 important ways. First, we additionally found hold time to be predictive of communicative intent classification. Second, our use of stick-light figures, rather than real videos, shows that intention recognition can occur even from highly reduced stimuli. Together, these results support the hypothesis that communicative intent can be read purely from movement kinematics (Becchio et al. 2012; Cavallo et al. 2016) and that both spatial and temporal features are important signals of intention.

We found that the exaggeration of submovements of the right hand was associated with perceiving an act as right-handed. This finding indicates that participants also attended to kinematic modulation in the Handedness Task, although the specific features were different from the Social Task. Given this finding, we are able to compare brain activation and connectivity results between the 2 tasks, as the primary difference is whether participants were basing judgments of communicative intentionality or handedness on the perceived kinematic modulation.

Brain Activation in Response to Communicative Kinematics

In the Social Task, we found activation of areas associated with the MS, such as the mPFC and left TPJ, as well as several areas associated with the mirroring system such as the inferior parietal lobe and PMC. Our results largely replicate the meta-

Table 1 Significant activation correlated with communicativeness across tasks

L/R	BA	Region	T	Z	k	x	y	z
Social Task								
R		Hippocampus	6.02	4.69	474	30	−19	−10
L		Caudate nucleus	5.59	4.46	438	−9	−1	14
L	32	mPFC	5.26	4.28	362	−9	38	23
L	47	IFG	5.23	4.26	130	−24	29	−1
L		Hippocampus	5.06	4.16	55	−27	−16	−7
L	39	TPJ	4.49	3.81	23	−54	−49	29
R	46	IPL	4.31	3.69	36	39	35	5
R	7		4.12	3.57	18	27	−79	38
R	40		3.99	3.47	52	57	−34	38
R		Cerebellum	3.94	3.44	11	9	−28	−40
R	6	PMC	3.86	3.34	11	24	−10	53
R		Cerebellum	3.82	3.36	16	3	−76	41
L			3.78	3.33	18	−24	−76	−25
R	6	PMC	3.74	3.3	11	21	11	47
Handedness Task								
R	46	MFG	4.16	3.56	17	51	41	2

BA = Brodmann area; IFG = inferior frontal gyrus; IPL = inferior parietal lobe; *k* = cluster size; L = left; MFG = middle frontal gyrus; R = right.
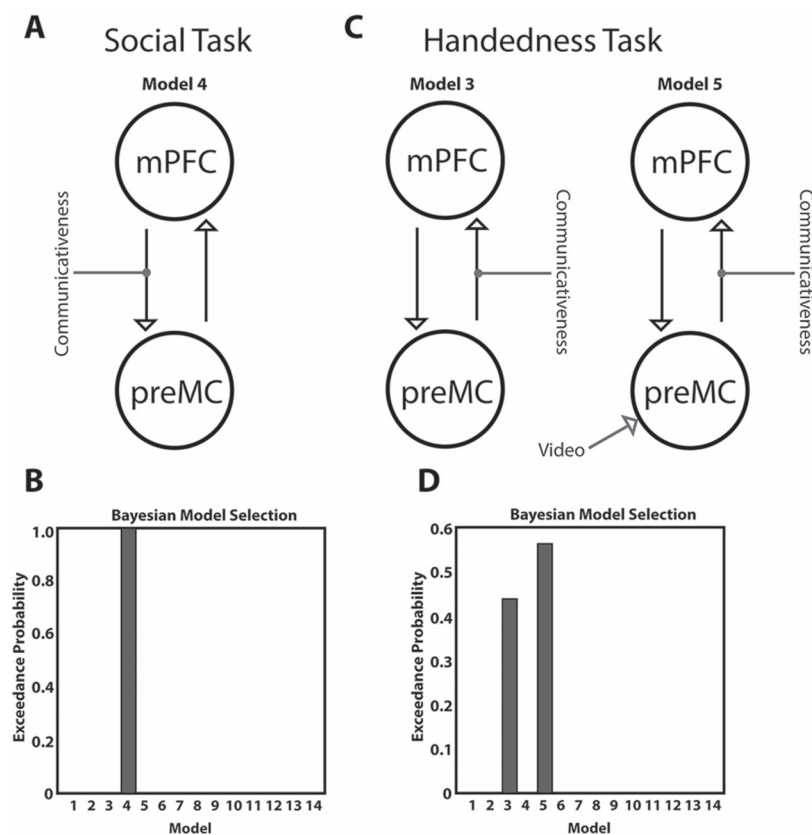

Figure 4. Overview of winning DCM models. (A) The winning model for the Social Task. (B) The exceedance probability. In all models, circles depict the individual regions, while arrows depict the intrinsic, directional coupling between them. Video viewing is modeled as a driving input to the regions, while communicativeness is modeled as a modulator of coupling strength. (C) The 2 high-probability models for the Handedness Task. (D) The exceedance probabilities for these models.

analytic findings by Van Overwalle and Baetens regarding brain activation while reading intentions from unusual or unexpected actions, experimental findings of brain activation in response to unexpected or unusual motions (Van Overwalle and Baetens 2009; Marsh et al. 2011, 2014), as well as implicit intention recog-

nition tasks using object-directed actions (Ciaramidaro et al. 2013). Similar to previous reports on violations of movement expectations, we found the right PMC (Manthey et al. 2003; Koelewijn et al. 2008; Van Overwalle and Baetens 2009), mPFC (Van Overwalle and Baetens 2009; Schiffer et al. 2014), and left

TPJ (Ciaramidaro et al. 2013) responding to increasingly communicative movements. One major distinction between our findings and those of the meta-analysis is that we found the left TPJ, whereas Van Overwalle and Baetens found the right TPJ. This can be explained by the left TPJ being primarily responsible for the processing of communicative intentions (Van Overwalle and Baetens 2009; Becchio et al. 2012; Ciaramidaro et al. 2013), whereas the right TPJ is involved in the processing of many other types of intentions as well (Van Overwalle and Baetens 2009; Ciaramidaro et al. 2013). These results are therefore directly in line with the idea that inferring abstract intentions is based on breaches of expectation originating in the MS, while expanding these previous findings by specifically showing that the brain responds similarly to subtle breaches at the kinematic level.

Besides the a priori predicted mentalizing and mirroring areas, we also found activation of the hippocampus and caudate nucleus to be correlated with communicative kinematics. Activation of both of these regions is directly in line with our theoretical framework. For example, previous work shows the caudate nucleus responding to expectation violations in a human movement observation paradigm (Schiffer and Schubotz 2011) as well as more generally in response to less familiar action sequences (Diersch et al. 2013). The hippocampus has similarly been linked to processing less familiar actions (Diersch et al. 2013) and is furthermore involved in signaling the presence of novel information (Lisman and Grace 2005) such as unfamiliar actions (Caligiore et al. 2013). These findings suggest that the caudate nucleus and hippocampus play an important role in processing unexpected movement kinematics in order to infer communicative intentions.

In the Handedness Task, we did not find any activation in our a priori defined regions of interest. This means that the regions found in the Social Task only respond when communicativeness is task-relevant. This finding is contrary to studies that used implicit viewing tasks and still found significant activation. However, a major difference in our study is that, while we used kinematic variations of the same overall action, previous studies typically use categorically different actions, such as lifting up an apple to take a bite compared with lifting it up to pass to the observer (Ciaramidaro et al. 2013). Thus, while the brain may respond robustly to categorically distinct socially intended actions, response to subtle kinematic differences may itself also be much more subtle in the absence of explicit attention to the underlying intention. On the other hand, we are not the first to report a task-dependent response to the intentionality of observed actions. Our finding is in agreement with an earlier study by de Lange and colleagues who similarly found activation of the MS in response to unusual actions, but only when explicitly attending to the intention (de Lange et al. 2008). de Lange et al. additionally found that an area of the mirroring system remained active in response to unusual actions even when not explicitly attending to the intention. Similarly, we found the middle frontal gyrus, which may also be involved in the pMNS (Molenberghs et al. 2011). Similarly, Spunt and Lieberman (2013) found that cognitive load, in the form of a competing memory task, extinguished activation of MS regions during abstract intention inference. Overall, we suggest that robust activation of the MS and pMNS in response to communicative kinematic modulation only occurs when the observer is actively attending to this aspect of the movement. Future studies will be needed in order to determine whether kinematic modulation will naturally draw attention in the absence of explicit task instructions, given that our control task may have inadvertently drawn attention away from this feature of the stimuli, rather than simply making it less task-relevant.

Effective Connectivity

In the experiment, participants had to infer intentionality of the observed actions, that is, decide if the action was performed "for the actor" or "for the viewer." The model-driven connectivity analysis showed that the kinematic modulation affected top-down coupling strength between mPFC and PMC and not vice versa. Our findings therefore provide evidence for a hierarchical system utilizing top-down expectations and bottom-up detection of kinematic deviations. This suggested mechanism allows us to draw a parallel with perceptual studies that empirically test the effect of unexpected stimuli on brain dynamics. Specifically, recent studies using DCM show that, while attending to auditory stimuli, unexpected omissions or mismatches of the stimulus result in changes to top-down connections between relevant brain regions (Auksztulewicz and Friston 2015; Chennu et al. 2016). More generally, these findings are also directly in line with models of top-down control in social cognition (Wang and Hamilton 2012; Hillebrandt et al. 2013).

Our finding fits well with experimental evidence of expectations shaping the dynamics of higher and lower level cognitive systems when processing concrete (i.e., end-goal) intentions. For example, in a recent study, Jacquet and colleagues measured corticospinal excitability to show that, when viewing and identifying the end goal of an action, changes to expectations regarding end-goal intentions result in tuning of the motor system (Jacquet et al. 2016). Interestingly, and in line with our study, these expectations could be based on observed kinematics and whether or not they were optimal for goal completion. While Jacquet et al. only looked at the motor system, a later study by Chambon and colleagues investigated the use of sensory evidence versus prior expectations to recognize concrete intentions while measuring whole-brain activation (Chambon et al. 2017). Chambon et al. found that top-down connections within the MS are modulated by an increasing reliance on prior expectations, which occurs when sensory evidence becomes less available or reliable (Chambon et al. 2017). Similarly, Ondobaka and colleagues found that the posterior cingulate cortex, another region of the MS, has a top-down effect on the action observation network during the processing of movement expectations of others (Ondobaka et al. 2015). While the specific regions in this study are different from our results, this may be due to the difference in the types of movement goals, or intentions, being processed. Ondobaka et al. conclude that their result shows support for a hierarchical account of action goal understanding with high-level midline (mentalizing) regions processing expected goals (or intentions) and lower level action observation, or mirroring, regions processing the movements. However, this study did not directly show changes in connectivity between higher and lower levels. Our results therefore provide an interesting extension to these previous findings, showing evidence for the importance of top-down connections when observing other's actions—including gesture.

In the Handedness Task, we see the pattern of connectivity modulation reversed. Increased communicativeness of the videos results in more modulation of the bottom-up coupling strength. This is in line with the study of coupling strength changes in response to unexpected auditory stimuli. In that study, top-down coupling changes were associated with an unexpected stimulus when this stimulus was the focus of

attention. When the stimulus was not the focus of attention, the top-down coupling effect was still present but paired with a bottom-up coupling change as well (Chennu et al. 2016). However, the DCM results from the Handedness Task should be interpreted with caution, as the GLM analysis of this task did not reveal significant activation of these regions at our specified threshold. Additionally, the fixed task order and different cognitive demands of the 2 tasks make it difficult to determine whether these connectivity differences are due to that lack of explicit attention to the communicative intent or to some other factor. We will therefore keep our discussion of these results to a minimum.

Overall, these results suggest that unexpected events result in top-down changes in connectivity at multiple levels of the brain. The detection of unexpected kinematics allows the recognition of communicative intentions.

Conclusions

In sum, we found that communicative intent can be read from isolated and subtle kinematic cues and that this recognition process is reflected in activation and (top-down) changes in connectivity of the mirroring and mentalizing systems. These results shine new light on how motor and social brain networks work together to process statistical irregularities in behavior to understand or "read" the complex dynamics of socially and communicatively relevant actions. Most directly, it highlights expectation violations as a key cue for inferring communicative intention, linking studies of movement, communication, and low-level perception. In particular, we show that even subtle kinematic differences in an otherwise typical motor act can be used to infer intention. This has theoretical implications for understanding the fundamental neurobiological mechanisms underlying perceptual inferences and communicative behavior as well as the evolutionary origins of communicative signaling. Practical implications extend to understanding human and human–machine interactions and providing a novel neuroscientific basis to investigate clinical conditions in which movement or social skills are impaired (e.g., autism spectrum disorder).

Funding

Notes

Conflict of Interest: None declared.

References

Auksztulewicz R, Friston K. 2015. Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb Cortex*. 25(11):4273–4283.

Bates D, Mächler M, Bolker B, Walker S. 2014. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 67:1–48.

Becchio C, Koul A, Ansuini C, Bertone C, Cavallo A. 2018. Seeing mental states: an experimental strategy for measuring the observability of other minds. *Phys Life Rev*. 24:67–80.

Becchio C, Manera V, Sartori L, Cavallo A, Castiello U. 2012. Grasping intentions: from thought experiments to empirical evidence. *Front Hum Neurosci*. 6:1–6.

Blakemore S, Decety J, Albert C. 2001. From the perception of action to the understanding of intention. *Nat Rev Neurosci*. 2(8):561–567.

Brass M, Schmitt RM, Spengler S, Gergely G. 2007. Investigating action understanding: inferential processes versus action simulation. *Curr Biol*. 17(24):2117–2121.

Caligiore D, Pezzulo G, Miall RC, Baldassarre G. 2013. The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neurosci Biobehav Rev*. 37(10):2504–2515.

Campisi E, Özyürek A. 2013. Iconicity as a communicative strategy: recipient design in multimodal demonstrations for adults and children. *J Pragmat*. 47(1):14–27.

Cavallo A, Koul A, Ansuini C, Capozzi F, Becchio C. 2016. Decoding intentions from movement kinematics. *Sci Rep*. 6(1).

Chambon V, Domenech P, Jacquet PO, Barbalat G, Bouton S, Pacherie E, Koechlin E, Farrer C. 2017. Neural coding of prior expectations in hierarchical intention inference. *Sci Rep*. 7:1–16.

Chennu S, Noreika V, Gueorguiev D, Shtyrov Y, Bekinschtein TA, Henson R. 2016. Silent expectations: dynamic causal modeling of cortical prediction and attention to sounds that weren't. *J Neurosci*. 36(32):8305–8316.

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

Ciaramidaro A, Becchio C, Colle L, Bara BG, Walter H. 2013. Do you mean me? Communicative intentions recruit the mirror and the mentalizing system. *Soc Cogn Affect Neurosci*. 65:461–468.

Core Team 2017. R: A language and environment for statistical computing. *R Foundation for Statistical Computing* Vienna, Austria. URL https://www.R-project.org/.

Csibra G, Gergely G. 2007. "Obsessed with goals": functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychol (Amst)*. 124(1):60–78.

de Lange FP, Spronk M, Willems RM, Toni I, Bekkering H. 2008. Complementary systems for understanding action intentions. *Current Biology*. 18(6):454–457.

Diersch N, Mueller K, Cross ES, Stadler W, Rieger M, Schütz-Bosbach S. 2013. Action prediction in younger versus older adults: neural correlates of motor familiarity. *PLoS One*. 8(5):e64195.

Friston KJ, Harrison L, Penny W. 2003. Dynamic causal modelling. *Neuroimage*. 19(4):1273–1302.

Frith CD, Frith U. 2006. The neural basis of mentalizing. *Neuron*. 50:531–534.

Gergely G, Csibra G. 2003. Teleological reasoning in infancy: the naïve theory of rational action. *Trends Cogn Sci*. 7(7):287–292.

Hillebrandt H, Blakemore J, Roiser JP. 2013. Dynamic causal modelling of effective connectivity during perspective taking in a communicative task. *Neuroimage*. 76:116–124.

Hudson M, McDonough KL, Edwards R, Bach P. 2018. Perceptual teleology: expectations of action efficiency bias social perception. *Proc R Soc B Biol Sci*. 285(1884):20180638.

Jacquet PO, Roy AC, Chambon V, Borghi AM, Salemme R, Farnè A, Reilly KT. 2016. Changing ideas about others' intentions: updating prior expectations tunes activity in the human motor system. *Sci Rep*. 6(1):26995.

Kampe KKW, Frith CD, Frith U. 2003. "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *J Neurosci*. 23(12):5258–5263.

Kilner JM, Friston KJ, Frith CD. 2007. Predictive coding: an account of the mirror neuron system. *Cogn Process.* 8:159–166.

Koelewijn T, van Schie HT, Bekkering H, Oostenveld R, Jensen O. 2008. Motor-cortical beta oscillations are modulated by correctness of observed action. *Neuroimage.* 40(2): 767–775.

Kuznetsova A. 2016. lmerTest package: tests in linear mixed effects models. *J Stat Softw.* 82:1.

Lacadie C, Fulbright R, Arora J, Constable R, Papademetris X. 2008. Brodmann areas defined in MNI space using a new tracing tool in BioImage suite. *Proceedings of the 14th Annual Meeting of the Organization for Human Brain Mapping*, June 14-19, 2008 Melbourne, Australia. 771 p.

Lisman JE, Grace AA. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron.* 46(5):703–713.

Manera V, Becchio C, Cavallo A, Sartori L, Castiello U. 2011. Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Exp Brain Res.* 211(3–4):547–556.

Manthey S, Schubotz RI, Von Cramon DY. 2003. Premotor cortex in observing erroneous action: an fMRI study. *Cogn Brain Res.* 15(3):296–307.

Marsh LE, De AF, Hamilton C. 2011. Dissociation of mirroring and mentalising systems in autism. *Neuroimage.* 56(3): 1511–1519.

Marsh LE, Mullett TL, Ropar D, Hamilton AF D C. 2014. Responses to irrational actions in action observation and mentalising networks of the human brain. *Neuroimage.* 103:81–90.

Meyer DE, R a A, Kornblum S, Wright CE, Smith JE. 1988. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychol Rev.* 95(3):340–370.

Molenberghs P, Mattingley J, Cunnington R, Mattingley JB. 2011. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci Biobehav Rev.* 36:341–349.

Newman-Norlund R, Van Schie HT, Van Hoek MEC, Cuijpers RH, Bekkering H. 2010. The role of inferior frontal and parietal areas in differentiating meaningful and meaningless object-directed actions. *Brain Res.* 1315:63–74.

Ondobaka S, De Lange FP, Wittmann M, Frith CD, Bekkering H. 2015. Interplay between conceptual expectations and movement predictions underlies action understanding. *Cereb Cortex.* 25(9):2566–2573.

Peirce J, Gray JR, Simpson S, Höchenberger R, Sogo H, Kastman E, Lindeløv JK. 2019. PsychoPy2: experiments in behavior made easy. *Behav Res Methods.* 51(1):195–203.

Pezzulo G, Donnarumma F, Dindo H. 2013. Human sensorimotor communication: a theory of signaling in online social interactions. *PLoS One.* 8(11):e79876.

Rizzolatti G, Cattaneo L, Fabbri-Destro M, Rozzi S. 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol Rev.* 94(2):655–706.

Sartori L, Becchio C, Bara BG, Castiello U. 2009. Does the intention to communicate affect action kinematics? *Conscious Cogn.* 18:766–772.

Schiffer A-M, Krause KH, Schubotz RI. 2014. Surprisingly correct: unexpectedness of observed actions activates the medial prefrontal cortex. *Hum Brain Mapp.* 35(4):1615–1629.

Schiffer A-M, Schubotz RI. 2011. Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Front Hum Neurosci.* 5:38.

Spunt RP, Lieberman MD. 2013. The busy social brain: evidence for automaticity and control in the neural systems supporting social cognition and action understanding. *Psychol Sci.* 24(1):80–86.

Spunt RP, Satpute AB, Lieberman MD. 2011. Identifying the what, why, and how of an observed action: an fMRI study of Mentalizing and mechanizing during action observation. *J Cogn Neurosci.* 23(1):63–74.

Trujillo JP, Simanova I, Bekkering H, Özyürek A. 2018. Communicative intent modulates production and comprehension of actions and gestures: a Kinect study. *Cognition.* 180: 38–51.

Trujillo JP, Vaitonyte J, Simanova I, Özyürek A. 2019. Toward the markerless and automatic analysis of kinematic features: a toolkit for gesture and movement research. *Behav Res Methods.* 51(2):769–777.

Van Overwalle F. 2009. Social cognition and the brain: a meta-analysis. *Hum Brain Mapp.* 30(3):829–858.

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

Wang Y, Hamilton AF de C. 2012. Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Front Hum Neurosci.* 6:153.