Action familiarity and the abstract representations of object-directed actions in the human brain:

demonstrated with magnetoancephalography

Margarita Krylova

Tufts University

Abstract

Action representation in the human brain has been localized to the Action Observation Network (AON). Magnetoencephalography (MEG) was used to study temporal modulation in seven regions of interest (ROI) per hemisphere, located within the AON, during viewing of realistic videos depicting familiar and less familiar goal-directed activities. A significant difference in activation was found in four out of the fourteen ROIs tested: the left dorsal premotor area (dPM), the left anterior intra-parietal sulcus (aIPS), the left ventral premotor area (vPM) and the middle intra-parietal sulcus (mIPS). Significantly greater activation for the familiar category was detected from 650 ms to 850 ms in the vPM, and in the mIPS and the aIPS from 1100 ms to 1300 ms. The dPM showed significantly enhanced activation for the unfamiliar category from 350 ms to 550 ms. These results provide evidence that familiarity with a given action results in an enhanced mirror neuron (MN) response during its observation in brain regions located lower in the action understanding hierarchy. The temporal modulation of activity observed in our study provides evidence that action observation involves both the spatial and temporal recruitment of areas within the AON.

Introduction

Defining Action Understanding

Human action understanding is a difficult concept to define, and frankly there is no correct way to do it. Literature reviews and studies concentrating on different aspects of action understanding view it in light of their distinct interests. For instance, a study on object use states that action understanding involves the "application of learned knowledge about proper object use" (Back, Pellen & Tipper, 2010, p.2798), while another study emphasizes "the recognition of a goal, and, from another perspective, the understanding of the agent's intentions" (Iacobini et al., 2005, p.0529). In fact, all definitions found in literature may be used to define action understanding, as over the course of our lives we perform and observe a wide range of actions that vary in complexity, context, significance, ubiquity, and so on. Furthermore, two individuals observing the same exact action might interpret it in two distinct ways depending on their unique background and experience. Five hundred years ago an individual living in Europe might have used his fork as often as an individual living in Asia might have used his chopsticks. Consequently, would a European observing someone eat with the familiar fork process this action differently than when observing someone eat with the unfamiliar chopsticks? The question we set out to evaluate in the present study is whether one's familiarity with an object used to perform an object-directed action plays a role in the way they comprehend this action when it is performed by another individual. The following sections of the introduction highlight key aspects of research on action observation that make our predictions and conclusions possible, including research on the mirror neuron system, action observation hierarchies and the roles of expertise.

<u>Theories of action understanding</u>

Most action understanding theories assume the existence of two interdependent hierarchies. The first is the action hierarchy that breaks down the action into levels composed of goals, actions and motor elements. The second is the control hierarchy that accounts for the neural processes behind these action levels. Uithol, Rooij, Bekkering & Haselager (2012) describe the two most common action hierarchies proposed in literature. The first is the part-whole relations hierarchy that decomposes actions into sub-actions and sub -sub-actions, from goal-level to action kinematics. According to this theory, the goal of 'getting milk' would consist of sub-goals such as 'walking to the fridge', 'opening the door' and 'grasping milk'. Each of these sub-goals would then be comprised of actions such as 'swinging the right leg' and 'standing on left leg' which are made up by motor acts like 'flexing knee' and so on. The second action hierarchy is the causal relations hierarchy. Like the part-whole relations hierarchy, the causal relations hierarchy breaks down every action into levels. However, the levels higher up in the hierarchy *cause* all lower levels to occur. Thus, the goal of 'getting milk' *causes* both 'walking to the fridge' and 'swinging the right leg'. Uithol et al. argue that the two aforementioned hierarchies are combined to explain neural mechanisms in literature but they cannot coexist because they assume incompatible interdependence between action levels.

In another hierarchical theory, known as the 'temporal extension hypothesis' action levels cooperate during motor activity and link together following training (Uithol et al., 2012). The temporal extension hypothesis was first modeled through the development of a humanoid robot motor system composed of neurons with fast reactivity and slow reactivity (Yamashita & Tani, 2008). After motor training, the fast neurons continuously repeated various patterns while the slow neuronal activity remained constant. In this simulation, the slow neurons were able to

represent goal level representations because of their slow dynamics, while the fast neurons represented lower levels responsible for actions or motor elements. Yamashita & Tani observed that eventually after training continuous sensorimotor activity divided into motor primitives which could be reused and integrated into novel actions.

A theory developed by Bernstein (1996) also proposes a complex relationship between action levels. Grafton, Antonia & Hamilton (2007) outline and name Bernstein's five ideas essential to action representation; chaining, adaptive variability, recursion, goal and chunking. Bernstein proposes that there is a set of sensory driven reflexes that can be combined in a specific order to attain a goal; chaining. This previously developed chain of motor elements can be recombined into a new action in order to fit a novel context; adaptive variability. However, the movements comprising the new action are formed from sets of previously learned priors, or primitives that are retrieved by the nervous system; recursion. All actions are formed in order to achieve a certain goal or solve a problem; goal. Eventually, independent motor elements are integrated into a single unit which involves an increase in co-articulation and a decrease in cognitive demands; chunking. According to Grafton et al. Bernstein's 'chaining' and 'chunking' allow for the combinations of motor elements according to an outcome representation that acts as a reference during action planning.

<u>Mirror neurons in monkeys</u>

The mirror neuron system has since its discovery been at the core of most theories on action understanding, although its exact function is still undergoing scrutiny and debate. Mirror neurons were originally observed in the F5 area in monkeys (*macaca menestrina*) by Di Pellegrino et al. (1992) and later found in other areas in the cortex. Di Pellegrino's single neuron

recordings revealed that the same neurons that fire when a monkey performs an object-directed action also fire when the monkey observes another monkey or human performing a similar object-directed action.

Mirror neurons are considered to be a class of visuomotor neurons, which also includes the canonical neurons that fire upon presentation of an object on its own. (Rizzolatti & Craighero, 2004). The mirror neurons in monkeys are activated when the observed action involves the interaction between a biological effecter (hand, mouth etc) and an object. About two thirds of the mirror neurons are more flexible, differentiating between action goals, these are called the 'broadly congruent' mirror neurons. The remaining one third fire in response to specific combinations of action goals and means for reaching these goals, these are called the 'strictly congruent' neurons. Whether the action is rewarded, and the identity of the object involved, do not appear important to mirror neurons activation.

Mirror neurons in humans

For ethical reasons, the same type of single neuron experiments used in studying monkey mirror neurons cannot be performed in humans. Therefore, there is no true evidence that mirror neurons exist in humans. However, neurophysiological and brain-imaging experiments have shown that there is most likely a mirror-neuron system in humans. The human mirror system is thought to include the inferior frontal gyrus (IFG) and the inferior parietal lobule (IPL) (Rizzolatti and Craighero, 2004) as well as the anterior intra-parietal sulcus (aIPS) and the premotor cortex (PMC) (Overwalle & Baetens, 2009). This system differs from the one observed in monkeys in several ways. First, the human mirror-neuron system appears to have a different physiology, although the complex homology between monkey and human brains complicates the

comparison of the two systems. Second, whereas the mirror-neurons in monkeys are only activated in response to the observation of goal-directed actions, human mirror neurons also respond to meaningless movements. Third, whereas monkey mirror-neurons code only for a specific action, human mirror- neurons appear to also code for the reduced movements that form each specific action. (Rizzolatti & Craighero,2004)

The degree to which mirror neurons contribute to action understanding in humans has been a source of much debate. Numerous studies have shown that mirror neurons are essential to action understanding. They have been implicated in speech perception, music perception, empathy, altruism, emotion, theory of mind, imitation, autism spectrum disorder and others (Hickok, 2008). On the opposite end of the spectrum are theories suggesting that mirror neurons are formed as byproducts of our everyday experiences and do not possess evolutionary significance. Recent years have witnessed multiple attempts to merge both hypotheses into one would allow mirror neurons to be essential to, but not solely responsible for action understanding.

<u>Mirror neurons play a role in action understanding</u>

The results of one study in particular are commonly used in literature to support a higher function for mirror neurons in action understanding. In the study, an experimenter performed object-directed actions in front of a monkey, while monkey single mirror neuron responses were recorded. Monkey mirror neurons responded when the entire action was visible to the monkey as well as when the monkey could see the experimenter pick up the object but not the hand-object interaction. When the monkey could not see the object or the experimenter's hands, mirror neurons showed no activation. According to these findings monkey mirror neurons are activated

not only when the monkey sees the action but also when it can infer the end of the action without

actually seeing it (Umilta et al., 2001). Another study using fMRI response in humans found that

the observation of grasping actions embedded in context produced greater activity in inferior

frontal mirror neurons than did the observation of grasping actions without context, or context

alone. Both studies provide evidence that mirror neurons are involved in the understanding of

action intentions, and are not simply an action recognition mechanism (Iacobini et al., 2005).

Arguments against the mirror neuron hypothesis

In a critical review, Hickok (2008) proposes eight points that undermine the essential role

of mirror neurons in action understanding. Hickok points out that there is no true evidence in

monkeys that mirror neurons support higher function action understanding, which may be

achieved via non-mirror neuronal mechanisms. The relation between mirror neurons in monkeys

and humans is undefined, a problem since mirror neuron research in humans is based on monkey

single neuron recordings. Furthermore mirror neurons have been discovered in the human motor

cortex responsible for lower-function actions and action understanding dissociates from the

human 'mirror system' and from brain areas involved in action production. Damage to certain

parts of the human mirror neuron network is not correlated with action understanding deficits

and the many studies that link the mirror system to speech recognition are not empirically

supported.

In the same review Hickok proposes two possible roles of the mirror neuron system in

action understanding. The first is that the mirror system plays no role in action understanding and

is just a Pavlovian-type association between sensory and motor information. He supports this

idea with a study conducted by Catmur, Walsh and Heyes (2009) in which participants were

asked to move their index finger while observing little finger movements. This resulted in the eventual activation of index finger representation in the primary motor cortex when participants observed little finger movements. According to Hickok, these results provide evidence for the associative sequence learning (ASL) model that the mirror neuron system develops through association and is therefore a meaningless sensorimotor action byproduct. Hickok's second proposal is that although abstract concept level representations are sufficient for action understanding, the sensory-motor information from mirror neurons enriches this information and provides it with a 'relational context'. Hickok provides an example in which someone who has never played a saxophone can still understand the idea of 'saxophone playing' but someone experienced with the instrument might recognize that the player is holding the saxophone incorrectly, with the help of his 'saxophone playing' experienced mirror neurons. Hickok's second proposal is a reconciliation of research ascribing mirror neurons higher cognitive function, and his first proposal that neuronal mirroring is a conditioned response.

Repetition suppression (RS) method provides evidence for a neuronal action representation hierarchy in the brain

Despite Hickok's critiques, there is growing evidence that although mirror neurons are not alone responsible for action understanding, they are an essential component of the action understanding network (Oztop, Kawato & Arbib, 2012) . Every human mirror neuron homologue part of the monkey 'mirror system' contains mirror and non-mirror neurons, it functions alongside with other brain regions and in the context of the environment. Disorders involving compulsive repetition of observed actions such as echopraxia and echolalia are examples of incorrectly integrated mirror neurons, evidence that healthy mirror neuron activity is essential for normal behavior, but also dependent on its interaction with higher brain centers.

Based on the action hierarchy theory proposed by Bernstein (1996), Grafton et al. (2007) conducted a series of experiments testing whether parts of the MNS (mirror neuron system) and AON (action observation network) could be activated differentially depending on the hierarchical level of an action being processed. They used the method of fMRI repetition suppression (RS) of the blood-oxygen-level-dependent (BOLD) signal. This method is based on the idea that repeated stimuli decrease the physiological response in the neural populations that represent them. This type of analysis would allow for the localization of brain areas responsible for processing different action levels. In their first two experiments participants watched an actor reach for, and grasp an object with their hand. Action goal was defined by the identity of the object grasped and kinematics was defined by the actor's hand trajectory. In experiment one, RS was observed in the left anterior-parietal sulcus (aIPS) when object grasped was manipulated. Manipulation of trajectory, resulted in RS in the left lateral occipital sulcus and the right superior pre-central sulcus.

In experiment two, Grafton et al. manipulated how an actor's hand grasped either a wine bottle or a dumbbell placed on-end. Alternating the wine bottle and the dumbbell would replicate results from experiment 1 for object grasped. Moreover, because the wine bottle and dumbbell have a thin and thick part, alternating between the hand grasping the thin and thick part of the dumbbell or bottle would test whether the aIPS is responsible for local kinematics, a higher goal process, or both. In line with experiment 1, RS for goal-object was localized to the left aIPS extending into the IPL, also the right aIPS, and also the left IFG with a lower statistical threshold. RS for how the object was grasped identified three clusters in the inferior and middle occipital regions, one cluster in the inferior frontal gyrus (IFG) and the middle inferior parietal sulcus (IPS), effects were also seen in the supplementary motor area (SMA) and the middle

frontal gyrus (MFG). The identification of RS in different areas supports the idea of a neuronal action hierarchy, localizing the representation of kinematic, goal-object and higher goal-object action features to different brain areas.

In the previous two experiments, defining the goal level by the identity of the object grasped did not alter the ultimate outcome of the action, which in all cases was grasping an object. However, the actor's one object-goal of grasping a cup can have different consequences on whether it is picked up for a drink or to be thrown at someone, for example (Grafton et al., 2007). Reasoning that action outcomes are higher in the hierarchy than object goals and may therefore be represented in different brain areas, in the next study Grafton et al. defined outcome as 'desired consequence of an action in the world'. In order to test whether action kinematics are found in a different part of the MNS from action outcomes fMRI RS was measured while participants watched movies with novel and repeated outcomes and novel and repeated kinematics. Stronger response was observed for novel outcomes in the right IPL and right IFG. Therefore, outcomes but not kinematics are represented in the right IPL and IFG. The lateralization for action outcome to the right hemisphere shows that it likely the predominant region for outcome representation. These effects generalize across a variety of actions, and provide evidence that the MNS in the inferior frontal and inferior parietal brain areas represent higher order intentions in addition to action kinematics. The presence of mirror neurons in the IFG was supported in a study comparing RS by executed and observed actions (Kilner Neal, Weiskopf, Friston, Frith, 2009). In the IFG, responses were suppressed when executed action was followed by the same observed action and when observed action was followed by the same executed action.

<u>Mirror neuron function within the action representation hierarchy explained using the predictive coding (PC) model</u>

Grafton et al. were able to use the RS method to localize regions responsible for several levels of the action understanding hierarchy and many studies have yielded similar findings. Using brain event related potentials (ERPs) Schie & Becking (2007) demonstrated that immediate goals are coded by parietal regions and final goals by frontal regions. Liepelt, Cramon & Brass (2008) differentiated between regions involved in goal inference of non-stereotypic actions and intentional actions. Iacobini et al. (2005) found that observing grasping actions in context resulted in greater mirror neuron activation in the inferior frontal cortex than observing grasping actions without context, suggesting that the MNS may code intentions. These, along with many other studies support the idea that mirror neurons play a role in action understanding.

The predictive coding (PC) model proposes an explanation for how mirror neurons function within the action observation hierarchy. As described by Kilner Friston and Frith (2007) The PC model is based on "minimizing prediction error at all levels of the cortical hierarchy that are engaged during action observation". The predictive coding model explains the sensorimotor mapping that occurs in the brain between the cause and the sensory expression of the cause. According to this model each level of the hierarchy predicts the representations in the level below. Backward connections convey prediction to the lower level where it is compared to produce a prediction error. The prediction error is then sent back to the higher level via forward connections, where it changes the prediction by adjusting the neuronal representation according to the new sensory information. This cycle continues until the prediction error is minimized. The mirror neuron systems in these motor regions are engaged in movement observation. Because, as demonstrated by Graton et al. 2007),these brain regions are arranged hierarchically, prediction

error signals in the higher levels will manifest as responses in cortically higher MNS neurons. During action observation we can use a prior expectation of a goal to predict a person's motor commands. Using the motor commands encoded in our own action system we can then predict their kinematics. By comparing predicted kinematics with the observed kinematics we generate a prediction error that updates our representation of the observed person's motor commands. (Kilner, Friston & Frith, 2007).

<u>Associative sequence learning, predictive coding and action familiarity</u>

When Hickok argues that mirror neuron system cannot play a role in action understanding because it develops through association, Press, Heyes & Kilner (2011) point out that the associative sequence learning (ASL) model only supports models of mirror neuron function within the action hierarchy. The ASL model proposes that mirror neurons form visuomotor connections between sensory-motor activity used to perform an action and the neural representation of this action. According to this model, individuals more experienced in performing a specific action would have stronger visuomotor connections in the brain region that encodes it. These connections form the network described by the predictive coding (PC) model. Due to continuous feedback through the PC neurons, the strength of the sensorimotor connections would become stronger and more fine-tuned. Over time the person would become more and more sensitive to the subtleties and nuances of the action, both when it is executed and observed. As proposed by Hickok's second hypothesis this would allow them to understand the action on a more professional level during its observation.

Although little to no studies have looked at this effect in everyday activities, it has been demonstrated that the mirror neuron systems of professional musicians and dancers show an

enhanced response compared to controls when they observe performances in their area of expertise. For example, Margulis et al. (2009) found that when musicians listened to the instrument of their expertise they exhibited an increased BOLD (blood-oxygen level dependent) in the left premotor and inferior parietal cortices, than when they listened to other instruments. In another study, dancers were trained to perform a dance piece. The reported ability of the dancers to execute the piece correlated with fMRI activation of their premotor and parietal areas when they observed the piece (Cross, Hamilton & Grafton, 2006). In a study by Haslinger et al (2005)professional pianists had an enhanced BOLD response in the inferior frontal gyrus (IFG) and in the bilateral inferior parietal cortices when they observed piano playing finger movements (Haslinger et al, 2005).

The present study

The majority of studies investigating the action observation network (AON) have focused on the identification of brain regions responsible for pre-specified levels involved in action observation, such as outcome, goal, kinematics and grip. The predominant methods used to localize these areas have been functional resonance imaging (fMRI) and repetition suppression (RS). However, the activity in the human motor system has been shown to modulate temporally (Press et al., 2011), and the study of this aspect of action observation is limited with both fMRI and RS. In the present study we used magnetoancephalography (MEG) techniques to study the temporal features of neural activations within 7 different regions of the AON when participants viewed goal-directed real-world video clips featuring familiar and unfamiliar target objects. Videos were divided into two conditions. Videos in which a goal was completed in a normal fashion (usual) were presented along with videos in which a goal was achieved using a novel combination of actions and objects (novel). Participants were asked to press one button if the

action in the video made sense to them and another button if it didn't. This paradigm encouraged participants to evaluate all aspects of the goal-directed action. We expected that due to participants' increased sensorimotor experience with familiar objects, mirror neuron areas of the left lateralized parietal AON (mIPS, SSC) involved in kinematics would display greater activation while participants watched movies in the familiar condition. We expected this increase in activation to proceed throughout the entire two seconds of object-directed action, and to level off at the completion of the goal. Toward the end of the video clip, we expected increased activation for the familiar category in the anterior intra-parietal sulcus (aIPS), a part of the mirror neuron system shown to differentiate between object-goal. We hypothesized that participants would show increased activation for videos in the unfamiliar category in frontal brain areas such as the right inferior frontal gyrus (rIFG) and the ventral premotor area (vPM), responsible for levels higher in the action observation hierarchy. According to the predictive coding model, participants would not have strong enough sensorimotor connections to understand the familiar movies through their mirror system and would instead rely directly on brain areas encoding outcome representation and action planning. We expected this difference in activation to occur directly after the onset of the final scene.

Methods

Participants

For the MEG recording 21 healthy volunteers (15 male, 6 female, M = 20.62) were recruited via advertisement. Exclusion criteria for participants included severe head trauma, neurological damage or illness, substance dependence at any time or substance abuse based on *DSM-IV* criteria within the previous 3 months, and treatment with psychotropic medication

within the previous 6 months. 31 healthy volunteers were recruited via Tufts sona system for the familiarity norming study.

Stimuli and Task

MEG stimuli

Participants were instructed to watch 180 pairs of video clips depicting goal-directed activities within the context of the real world. The video clips included common everyday activities such as cleaning and cooking as well as less common activities such as changing a light bulb and oaring. Each video clip consisted of a context and a final scene presenting a goal-directed action using a target object. All video clips were categorized as either congruous or incongruous. In congruous video clips the target object used in the final scene fit the context of the entire video. The ability of the context to prepare the viewer for the single target object was confirmed by the Cloze probability test (>70% on average). Each Incongruous video clip was contextually identical to a congruous video but featured a target object from a separate congruous video clip that was not expected based on context. A cinematographic cut separated the context from the final scene of each video (see Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008), which was designated by a yellow frame. Target objects were engaged in the central action and clearly presented at the onset of the final scene, but did not appear in the context.

Video clips were organized into two sets, each consisting of 180 congruous and 180 incongruous clips. No context or final scene was included twice in one set. Between the two sets, all contexts and target objects were included in both the congruous and incongruous conditions. Half of the participants viewed video clips from the first set, and the other half from the second set. The present study evaluated MEG recordings only from congruous video clips. The presence

of incongruous videos encouraged participants to consciously evaluate the actions presented in the congruous videos in order to come up with action goals.

The 180 movies in the congruous condition were divided into two categories based on the target object familiarity rating using a median split. Movies that included target objects rated most familiar were included in the familiar category and movies that included target objects rated least familiar were included in the unfamiliar category. No significant difference was found in the visual complexity rating between objects in the familiar and unfamiliar categories; $t(177) = 1.47, p > .05$.

Video clips were centered on a black background and subtended 4° of visual angle. Clips were without sound and were presented at a rate of 30 frames per second. All contexts lasted 6 seconds and all final scenes 2 seconds. 100 ms after the offset of each final scene, a green question mark appeared in the center of the screen for 900 milliseconds. Participants were given a button box and asked to respond accordingly upon seeing the question mark. Participants were instructed to press one button if a reasonable goal could be inferred from the video clip, and another button if not. Six additional clips were used during a practice session prior to the task. Participants were asked not to blink during the second half of each movie. After every few video clips, a white plus sign appeared on the screen for 1 second. Participants were instructed to rest during the plus signs, but asked to remain still for the duration of the imaging task.

Familiarity norming

Participants were shown the last 2 seconds of the 180 incongruous video clips described previously on a computer screen. Video clips included all target objects present in the MEG

recording. Participants were instructed to identify the target object and to rate the complexity and familiarity of the object by pressing keys 1-7 on the computer keyboard.

Data Acquisition

MEG data were recorded using a 306-channel Vectorview MEG system (Elekta Neuromag Oy, Helsinki, Finland) contained in a six-layer magnetically shielded room. Head position was monitored during the task using head position indicator (HPI) coils. Points on each participant's scalp surface, as well as the positions of the HPI coils, were digitized using a Polhemus FastTrak digitizer; this permitted the spatial alignment of the MEG sensor array with the participant's structural MRI data. Sampling rate was 600 Hz and bandwidth of acquisition was 0.1-200 Hz. Electro-occulogram recordings were used to correct for physiological noise, such as eye blinks, and eye saccading.

Structural MRI data were recorded with a 3T scanner (Siemens Medical Solutions, Erlangen, Germany). One multi-echo T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) sequence (TR, 2530 msec, TE, 1.64/3.5/5.36/7.22 msec for four consecutive images, flip angle, 7°, matrix size, 256 x 256) and two fast low-angle shot (FLASH) scans (651 hz/pixel, 5° and 30° flip angle, respectively, TR, 20 ms, TE = (1.8 + 1.82 x n) ms, n, 0-7) were performed. FLASH sequences were used to construct individual boundary-element model (BEM) surfaces.

Data Analysis

The minimum norm estimate (MNE) distributed model was used to estimate cortical sources of MEG activity. The forward model was calculated using a BEM. To correct for movement between runs, the average of the forward solutions for each run was used in the

analysis. MEG signal at each point on the cortical surface was estimated every 5 ms. To determine where MNE signals was greater than the noise level, dynamic Statistical Parametric Map (dSPM) was calculated.

Regions of interest (ROIs) were based on a quantitative meta-analysis of consistently activated cortical regions associated with action observation across 104 experiments (1061 subjects, 1390 activation foci) conducted by Caspers et al. (2010). Each ROI measured 400mm *2. MEG signal for each condition was averaged across vertices within each ROI. Time windows showing significant variability in activation due to object familiarity between groups were determined using a 2-tailed independent-samples t-test with a 0.05 alpha level in each vertex. To confirm that significant time windows were not determined by chance, 1000 Monte Carlo simulations were performed. For each subject and simulation, numbers with a Gaussian distribution were generated and analyzed according to the same parameters as for the experimental data. Time windows that were obtained in less than 5% of simulations were considered significant.

Results

In order to examine activation within the action observation network we constrained our analysis to 7 regions of interest constituting the action observation network (AON) in the right and left hemispheres; middle intra-parietal sulcus (mIPS), somatosensory cortex (SI, BA 1/3), dorsal premotor area (dPM, BA6), anterior intra-parietal cortex (aIPS), superior temporal sulcus (STS,), inferior frontal gyrus/ventral premotor area (IFG/vPM, BA 44/45) and middle temporal sulcus (mTS).

For each ROI, MEG signal was averaged across vertices at each time point. Plots of time-course for each ROI show the difference in activity between the familiar and unfamiliar conditions between 300 and 1500 ms. Plots are time-locked to the onset of the final scene. Red highlighted areas designate regions of significant difference in activation between movies that include familiar and unfamiliar objects. Significant difference in activation was found in four out of the fourteen ROIs tested; the left dorsal premotor area (dPM), the left anterior intra-parietal sulcus (aIPS), the left ventral premotor area (vPM) and the middle intra-parietal sulcus (mIPS). Significantly greater activation for the familiar category was detected from 650 ms to 850 ms (mid-final scene) in the vPM, and in the mIPS and the aIPS from 1100 ms to 1300 ms (toward end of final scene). The dPM (noise normalized)showed significantly enhanced activation for the unfamiliar category from 350 ms to 550 ms (beginning of final scene).

Discussion

The study of action understanding is made possible by two hierarchies. The first is the action hierarchy that decomposes actions into action levels; outcomes, goals, kinematics, motor events. The second is the neuronal hierarchy of brain areas, with each area in the AON playing a role in the comprehension of action understanding. This hierarchy is supported by research findings that brain areas are activated differentially depending on the level of action being processed. According to the predictive coding (PC) model these areas are connected to one another through reciprocal connections. Areas responsible for processing action levels higher in the action hierarchy predict the responses at the lower levels, and upon receiving feedback from the these levels, update their representations according to prediction error. When actions are continuously processed in this manner, the sensorimotor mirror neuron connections become increasingly fine tuned to nuances and exhibit an enhanced response to these actions during

action observation. Such increased mirror neuronal firing has been demonstrated in professional musicians and dancers. In the present study we sought to evaluate the temporal modulation of activity in response to goal-directed actions encountered in everyday real life scenarios.

We found significant differences in activation between the familiar and unfamiliar conditions in 4 regions within the AON. Each region exhibited pronounced temporal modulation throughout the two seconds of the final scene, during which the target object was presented and used to complete an object-directed action. Temporal modulation of the difference in activation between familiar and unfamiliar conditions may be used to determine the temporal aspect of the object directed action that is encoded by the region displaying this difference. Significantly greater activation for familiar movies was detected in the middle of the final scene (during the action) in the ventral premotor area (vPM) and in the medial intra-parietal sulcus (mIPS), as well as in the anterior intra-parietal sulcus (aIPS) toward the end of final scene (at the completion of object-directed action). Enhanced activation in the dorsal premotor area (dPM) for the unfamiliar condition was found at the beginning of the final scene (initiation of action). Consistent with our hypothesis, we found increased activation in the rIFG to unfamiliar movies, and in the somatosensory cortex to familiar movies, but neither difference was significant.

Three regions of the AON showed an enhanced response to the familiar condition; the aIPS, vPM and mIPS. All three areas are located within the mirror neuron (MN) regions of the action observation network (AON). This finding is consistent with our hypothesis that participants' mirror neuron (MN) AON regions would exhibit enhanced response to the familiar condition by activating the participants' own sensorimotor connections while they watch familiar actions that they may have performed in the past. Increased activation of MN to familiar actions suggests that sensorimotor connections are not unique to specialized activities such as playing an

instrument or dancing. Sensorimotor connections are developed as we perform everyday tasks, for example brushing our teeth or cleaning.

Although the vPM, aIPS and mIPS all showed increased response to the familiar condition, the temporal aspect of this activation differed. In the aIPS and mIPS the difference in activation between the familiar and unfamiliar categories significantly increased 300 ms before the end of the final scene. Because this time frame corresponds to the completion of the object-directed action, we can infer that the aIPS and mIPS are involved in the completion of the action goal. This finding is consistent with the observed response suppression (RS) of the blood oxygen level dependent methodd (BOLD) in the left aIPS and mIPS recorded by Grafton et al. (2006) when participants were presented with the same object-goal. In an fMRI study by Gritkes et al. (2004) monkeys used a joystick to execute precise movements toward a visual target. Consistent with our findings, the study concluded that the mIPS controls goal-directed hand movements. It is noteworthy that although the aIPS and the mIPS were activated for both the familiar and unfamiliar actions during the final scene, it is only at the end of the scene, during the object directed action, that the difference in activation for the familiar and unfamiliar movies emerged. At this point the activation for familiar movies increased in both the aIPS and the mIPS, but leveled off for unfamiliar movies.

Activity in the ventral premotor area (vPM) also increased in response to familiar movies, but unlike the activity of the mIPS and aIPS, it peaked in the middle of the final scene. The vPM is a part of the parietal MN AON homologue of the monkey F5 area and is implicated in multiple action components. Ferrari, Rozzi and Fogassi (2005) recorded neuronal activity in the monkey vPM area while they made wrist movements in varying forearm postures. They found that 94% of the neurons in the vPM were selective for direction of movement in space. In an fMRI study

in humans, Chao & Martin (2000) demonstrated increased activation in the vPM when

participants viewed pictures of tools. Another study demonstrated that lesion of the vPM

resulted in deficits of grasping objects (Fogassi et al., 2001). These results demonstrate that the

vPM matches motor acts to sensory inputs and is most active during action execution (Hoshi &

Tanji, 2007). In our study the vPM was active during the middle part of the final scene, while the

final action was still being executed. Since the vPM is involved in matching inputs to motor acts,

it is according to the PC model on an intermediate level action hierarchy, and is dependent on

constant sensorimotor input. This is consistent with our hypothesis that the enhanced vPM

response to the familiar condition is due to strengthened mirror association with the familiar to

participants actions.

The dorsal premotor area (dPM) is the only area that displayed a significant increased

response to the unfamiliar condition. It was also the only area with the difference in activation

manifesting at the onset of the final scene. The early activation for the unfamiliar condition is

consistent with our hypothesis that the dPM is involved in movement preparation and is therefore

on a higher level according to the PC model. According to Hoshi and Tanji (2007) "dPM activity

reflects the motor significance of the instructional cue rather than its sensory or attentional

significance (Pellegrino & Wise, 1993). They propose that the difference in activation between

the dPM and vPM can be explained using the direct and indirect matching hypotheses. The vPM

involved in direct matching receives information about a motor target and sends outputs to

achieve the desired action. In contrast to the vPM, the dPM "has a major role in indirect

sensorimotor processing, retrieving multiple sets of motor information from sensory signals and

integrating components of a required action to formulate a motor program for the intended

action"(Pellegrino & Wise, 1993, p.239). Thus, due to its integrative role, the dPM is higher on

the action hierarchy and is not as dependent on direct sensorimotor connections. In our study the dPM exhibited increased activity for unfamiliar actions, instead of mirroring these actions through past experience, the dPM was generating a 'motor program for the predicted action' before its onset. Our detection of enhanced dPM activity in the beginning of the final scene further supports this idea.

Together, our findings support the hypothesis that familiarity with real-life goal-directed actions enhances the sensorimotor connections in the mirror neuron (MN) action observation network (AON) brain regions, as is proposed by the predictive model (PC) and the associative sequence learning (ASL) models. The extensive sensorimotor connections result in enhanced mirror neuron response to increasingly detailed action components, within brain regions lower in the action understanding hierarchy. Unfamiliar actions that have not formed strong sensorimotor connections are processed by brain regions higher in the AON, such as those involved in action planning. Moreover, the temporal modulation in activity reported in our study provides evidence that in studying the AON, it is imperative to acknowledge that the spatial and temporal aspects of action understanding are interdependent, and cannot exist one without the other.

References

Bach, P., Peelen, M.V., Tipper, S.P. (2010) On the role of object information in action observation: an fMRI study. *Cerebral Cortex*, 20, 2798-2809.

Bernstein, N.A., On dexterity and its development. In: Latash ML, Turvey MT, editors. Dexterity and its development. New Jersey: Lawrence Erlbaum Associates;1996.

Caspers, S., Zilles, K., Laird, A.R., Eickhoff, S.B. (2010) ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, 50, 1148-1167.

Catmur, C., Walsh, V., Heyes, C. (2009) Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society*,

Chao, L.L., Martin, A. (2000) Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, 12, 478-484.

Cross, E.S, Hamilton, A.F.C., Grafton, S.T. (2006) Building a motor simulation de novo: Observation of dance by dancers. *Neuroimage*, 31, 1257-1267

di Pellegrino, G., Fadiga, L., Fogassi L., Gallese, V., Rizzolatti, G. (1992) Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91, 176-180.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., Rizzolatti, G. (2001) Cortical mechanisms for the visual guidance of hand grasping movements in the monkey: a reversible inactivation study. *Brain*, 124, 571-586.

Gritkes, C., Ritzl, A., Zilles, K., Fink, G.R. (2004) Human medial intraparietal cortex subserves visuomotor coordinate transformation. *Nueroimage*, 23, 1494-1506.

Grafton, S.T., Hamilton, A.F.C. (2007) Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26, 590-616.

Hamilton, A.F.C, Grafton, S.T. (2007) Action outcomes are represented in human inferior fronto-parietal cortex. *Cerebral Cortex*, 18, 1160-1168.

Hamilton, A.F.C., Grafton, S.T. (2006) Goal representation in human anterior intraparietal sulcus. *Journal of Cognitive Neuroscience*, 26, 1133-1137.

Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H., Ceballos-Baumann, A.O. (2005) Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, 17, 282-293.

Hickok, G. (2008) Eight problems for the mirror neuron theory of action understanding in monkeys and in humans. *Journal of Cognitive Neuroscience*, 21, 1229-1243.

Hoshi, E, Tanji, J. (2007) Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties. *Current Opinion in Neurobiology*, 17, 234-242.

Iacaoboni, M., Molnar-Szakacs I., Gallese, V., Buccino, G., Mazziotta, J.C, Rizzolatti, G. (2005) Grasping intentions of others with one's own mirror neuron system. *PLoS Biology*, 3, 0529-0534.

Kakei, S., Hoffman, D.S., Strick, P.L. (2001) Direction of action is represented in the ventral premotor cortex. *Journal of Cognitive Neuroscience*, 17, 212-226.

Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D. (2009) Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neroscience*, 29(32), 10153-10159.

Kilner, J.M., Friston, K.J., Frith, C.D., (2007) Predictive coding: an account of the mirror neuron system. *Cognitive Processes*, 8, 159-166.

Liepelt, R., Cramon, D.Y.V., Brass, M. (2008)How do we inver others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *NeuroImage*, 43, 784-792.

Margulis, E.H., Misna, L.M., Uppunda, A.K., Parrish, T.B., Wong, P.C.M. (2009) Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Human Brain Mapping*, 30, 267-275.

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

Oztop, E., Kawato, M., Arbib, M.A. (2012) Mirror neurons: functions, mechanisms and models. *Neuroscience Letters*, 13, 1-13.

Pellegrino, G., Wise, S.P. (1993) Visuospatial versus visuomotor activity int he premotor and prefrontal cortex of a primate. *The Journal of Neuroscience*, 13, 1227-1243.

Press, C., Heyes, C., Kilner, J.M. (2010) Learning to understand others' actions. *Biology Letters,*7, 457-460.

Press, C., Cook, J., Blakemore, S.J., Kilner, J. (2011) Dynamic modulations of human motor activity when observing actions. *The Journal of Neuroscience*, 31, 2792-2800.

Rizzolatti, G., & Craighero, L. (2004) The mirror neuron system. *Annual Review Neuroscience*, 27, 169-192.

Sheng, L., Ostwald, D., Giese, M., Kourtzi, Z. (2007) Flexible coding for categorical decisions in the human brain. *The Journal of Neuroscience*, 27, 12321-12330.

Shmuelof, L., Zohary, E., (2007) Watching others' actions: Mirror representations in the parietal cortex. *The Neuroscientist*, 13, 667-672.

Sitnikova, T., Holcomb, P.J., Kiyonaga, K.A, Kuperberg, G.R., Two neurogocnitive mechanics of semantic integration during the comprehension of visual real-world events. Journal of Neuroscience, 20, 2037-2057.

Uithol, S., Rooji, I.V., Beckkering, H., Haselager, P. (2012) Heirarchies in action and motor control. *Journal of Cognitive Neuroscience*, 24, 1077-1086.

Umilta, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers C., Rizzolatti, G. (2001) I know what you are doing: a neurophysiological study. *Neuron*, 31, 155-165.

Yamashita, Y., Tani, J., (2008) Emergence of functional hierarchy in a multiple timescale new network model: A humanoid robot experiment. *PLoS Computational Biology*, 4(11).
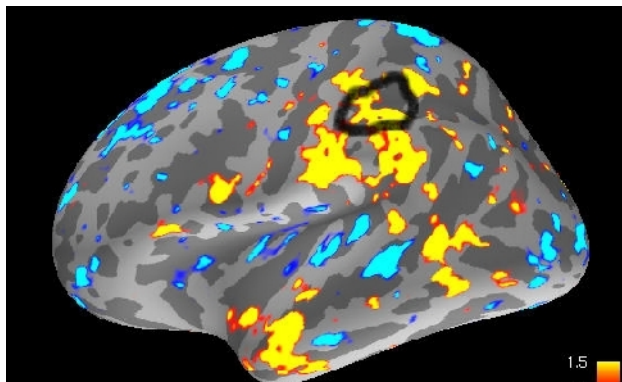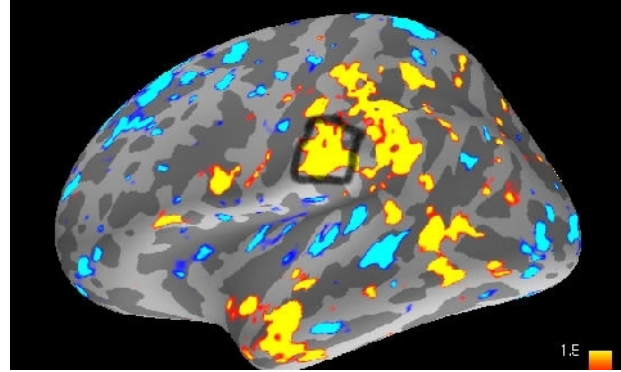
A) middle intraparietal sulcus (mIPS) 1200 ms
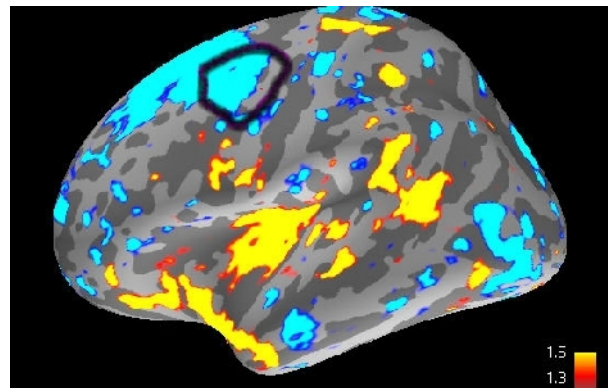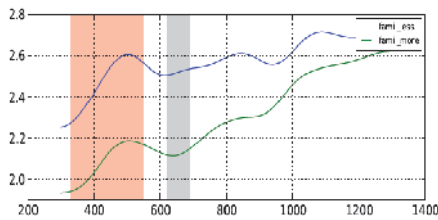
B) dorsal premotor area (dPM) 400 ms
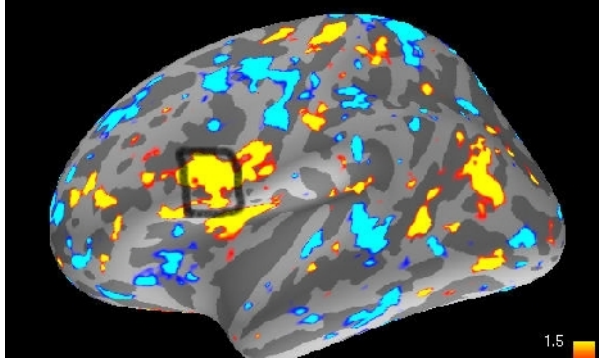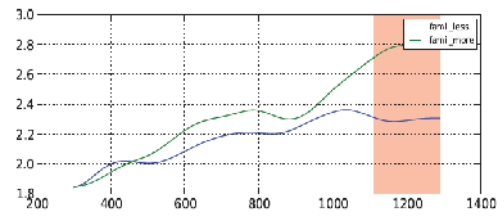
C) ventral premotor area (vPM) 800 ms

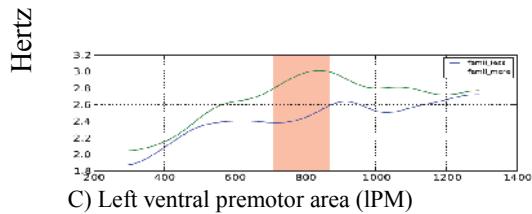D) anterior intraparietal sulcus (aIPS) 1200 ms

Figure 1 - Yellow indicates areas with greater activation for familiar category, blue indicates greater activation for unfamiliar category. Labeled regions of interest (ROI) are circled in black.
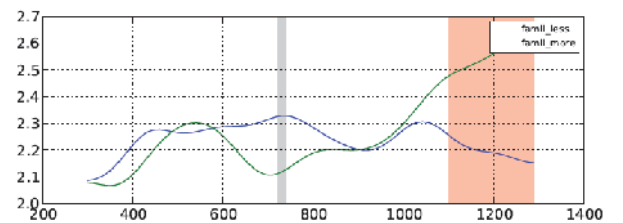

A) Left dorsal premotor cortex (dPM)

B) Left anterior intraparietal sulcus (aIPS)

Hertz

C) Left ventral premotor area (lPM)

D) Left medial intraparietal sulcus (mIPS)

Time (ms)

Figure 2 - Time course of activation in the four regions of interest that yielded significant results. Blue lines represent the time course of activation for unfamiliar category, green lines represent time course for familiar category. Pink areas highlight the time course during which activation for familiar and unfamiliar movies was significantly different.