Published in final edited form as:
Neuroimage. 2013 December ; 83: 493–504. doi:10.1016/j.neuroimage.2013.06.060.

Controlling automatic imitative tendencies: Interactions between mirror neuron and cognitive control systems

Katy A. Cross[a,b], **Salvatore Torrisi**[a,c], **Elizabeth A. R. Losin**[a,b,1], and **Marco Iacoboni**[a,b,c,d]

[a]Interdepartmental Neuroscience Program, University of California, Los Angeles

[b]Ahmanson-Lovelace Brain Mapping Center, University of California, Los Angeles

[c]Department of Psychiatry & Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles

[d]Brain Research Institute, University of California, Los Angeles

Abstract

Humans have an automatic tendency to imitate others. Although several regions commonly observed in social tasks have been shown to be involved in imitation control, there is little work exploring how these regions interact with one another. We used fMRI and dynamic causal modeling to identify imitation-specific control mechanisms and examine functional interactions between regions. Participants performed a pre-specified action (lifting their index or middle finger) in response to videos depicting the same two actions (biological cues) or dots moving with similar trajectories (non-biological cues). On congruent trials, the stimulus and response were similar (e.g. index finger response to index finger or left side dot stimulus), while on incongruent trials the stimulus and response were dissimilar (e.g. index finger response to middle finger or right side dot stimulus). Reaction times were slower on incongruent compared to congruent trials for both biological and non-biological stimuli, replicating previous findings that suggest the automatic imitative or spatially compatible (congruent) response must be controlled on incongruent trials. Neural correlates of the congruency effects were different depending on the cue type. The medial prefrontal cortex, anterior cingulate, inferior frontal gyrus pars opercularis (IFGpo) and the left anterior insula were involved specifically in controlling imitation. In addition, the IFGpo was also more active for biological compared to non-biological stimuli, suggesting the region represents the frontal node of the human mirror neuron system (MNS). Effective connectivity analysis exploring the interactions between these regions, suggests a role for the mPFC and ACC in imitative conflict detection and the anterior insula in conflict resolution processes, which may occur through interactions with the frontal node of the MNS. We suggest an

Corresponding Author: Katy A. Cross, Ahmanson-Lovelace Brain Mapping Center, 660 Charles E. Young Dr. South, Los Angeles, CA 90095, (310) 206-4456.
[1]Present address: University of Colorado, Boulder

extension of the previous models of imitation control involving interactions between imitation-specific and general cognitive control mechanisms.

Keywords

Automatic imitation; spatial compatibility; cognitive control; mirror neurons; fMRI; dynamic causal modeling

1. INTRODUCTION

During social interactions humans tend to mimic the postures and gestures of others. This mimicry is automatic in that it occurs without will or awareness (Chartrand and Bargh, 1999; Niedenthal et al. 2005). It also seems to be beneficial, increasing positive feelings and successful communication between social counterparts (Chartrand and Bargh, 1999; Lakin et al. 2003). The prevailing neural explanation for automatic imitative tendencies is that observing actions activates the corresponding motor program through a direct matching mechanism (reviewed in Heyes, 2011). This direct matching between observed and performed actions is thought to be mediated by the mirror neuron system (MNS) (Iacoboni et al. 1999; Ferrari et al. 2009; Heyes, 2011), which responds both to the observation of specific actions and the execution of similar actions. The strongest support for this model of automatic imitation comes from single-pulse transcranial magnetic stimulation (TMS), a technique that can be used to measure the cortico-spinal excitability of specific response representations. Many studies have now demonstrated that passive action observation causes increased cortico-spinal excitability specific to the muscles involved in producing the observed action (Fadiga et al. 1995; Baldissera et al. 2001; Gangitano et al. 2001; Gangitano et al. 2004; Clark et al. 2004; Montagna et al. 2005; Borroni et al. 2005; D'Ausilio et al. 2009). In other words, observing actions causes sub-threshold activation of the imitative response. This so-called "motor resonance" is reduced after the ventral premotor cortex (a putative MNS region) is disrupted with repetitive TMS, providing evidence that the frontal node of the MNS plays a causal role in the effect (Avenanti et al. 2007). In addition, TMS disruption of the same premotor region also reduces automatic imitation (Catmur et al. 2009), and social priming manipulations that modulate automatic imitation also modulate motor resonance (Obhi et al. 2011). Thus, there is increasing evidence for a link between motor resonance, the MNS and automatic imitation.

While the neural substrates leading to automatic imitation are relatively well-studied, it is less clear how these automatic tendencies are brought under intentional control. Action observation automatically activates the corresponding motor representation, yet under normal circumstances we do not overtly imitate all observed actions. This is likely due to an active control system that inhibits unwanted imitation; the observation of patients who imitate excessively after large lesions in the frontal lobe (Lhermitte et al. 1986; De Renzi et al. 1996) suggests a disruption of this active imitation control mechanism. If imitation is supported by a specialized action-observation matching system (Iacoboni et al. 1999), imitation control may rely on neural systems distinct from other commonly studied control mechanisms. Specifically, imitative control may be different from control employed in Stroop, flanker and spatial compatibility tasks, where automatic response tendencies are

evoked by non-social, symbolic stimuli. This hypothesis has received some support from neuroimaging (Brass et al. 2005) and neuropsychological (Brass et al. 2003) studies demonstrating dissociations between control processes in imitation and Stroop tasks and has led to the "shared representations" theory of imitative control (Brass et al. 2009a; Spengler et al. 2010).

The shared representations theory proposes that a central process in imitation control is distinguishing between motor activity generated by one's own intentions from motor activity generated by observing someone else perform an action. This is required because both perceived and internally planned actions are represented in the same neural system (the MNS; Rizzolatti and Craighero, 2004), yet the system itself does not distinguish between the source of the representations (i.e. whether activity is caused by one's own intentions or the observation of others' actions; Jeannerod, 1999). Therefore, when two different (conflicting) motor representations are simultaneously activated by intentions and action observation, an imperative first step to carrying out the intentional action (and avoiding imitation) is to attribute each motor representation to either self or other.

Early support for the shared representations hypothesis came from the observation that neural substrates of imitative control are similar to those observed in more complex social tasks that also require self-other distinctions and the representation of conflicting mental states (Brass et al. 2005; Brass et al. 2009a; Spengler et al. 2009). Specifically, the medial prefrontal cortex (mPFC) and temporo-parietal junction (TPJ) were shown to be involved in imitation control across a variety of studies (Brass et al. 2001; Brass et al. 2005; Brass et al. 2009a; Spengler et al. 2009; Wang et al. 2011b) and these regions are also involved in mentalizing, self-referential processing and determining agency (Ruby and Decety, 2001; Farrer and Frith, 2002; Farrer et al. 2003; Amodio and Frith, 2006; Nahab et al. 2011). Subsequent behavioral (Spengler et al. 2010b), neuropsychological (Spengler et al. 2010a; Spengler et al. 2010) and neuroimaging (Brass et al. 2009a; Spengler et al. 2009) research provided more direct links between higher social cognitive functions and imitative control. Based on this work, Brass and colleagues proposed that in the context of imitative control the TPJ distinguishes between self- and other-generated motor activity by signaling that the observed action is related to another agent (regardless of the presence of conflict), whereas the mPFC enforces the self-generated action when it conflicts with an externally-generated action representation (Brass et al. 2009b).

While the shared representations theory has gained traction, it does not describe mechanisms of imitation control beyond the involvement of mPFC and TPJ. For example, it is not clear how the mPFC resolves conflict between observed and intended actions after self-other distinctions are made. Furthermore, the mPFC and TPJ are not the only regions associated with imitative control tasks. The frontal operculum (Bien et al. 2009a; Wang et al. 2011b) and ventral premotor cortex (Brass et al. 2005; Spengler et al. 2009) have also been observed to be active during imitation control. The inferior frontal regions have been interpreted as the frontal node of the human mirror neuron system (MNS) (Spengler et al. 2009; Wang et al. 2011b), suggesting that imitation control involves modulation of the MNS. However, this hypothesis has only received indirect support.

To build on previous models of imitative control we used dynamic causal modeling (DCM) for fMRI to examine causal interactions between regions involved in imitative control and to test the hypothesis that resolving imitative conflict involves MNS modulation. In an imitation interference task, subjects performed a finger-lifting action while simultaneously watching a video clip depicting either the same action or a different action. Numerous studies have demonstrated that subjects are slower to respond on incongruent trials—when the observed and performed action differ—compared to congruent trials—when the observed and performed action are the same (Brass et al. 2000; Stürmer et al. 2000; Brass et al. 2001; Kilner et al. 2003; Bertenthal et al. 2006a; Bird et al. 2007; Longo et al. 2008; Gillmeister et al. 2008; Press et al. 2008; Catmur and Heyes, 2010; Wang et al. 2011a). This slowing is attributed to the recruitment of imitative control processes on incongruent trials; since the imitative response is incorrect, it needs to be inhibited to allow execution of the correct non-imitative response. Therefore, regions more active during incongruent compared to congruent trials are likely involved in imitation control.

In addition to the imitation interference task, we included a spatial interference paradigm that was identical except the stimuli depicted moving dots instead of moving fingers. The rationale for including the spatial task was twofold. First, it allowed us to identify regions that are involved specifically when conflict arises from action observation, in line with an imitation control mechanism that is distinct from mechanisms for overcoming automatic responses evoked by non-social, symbolic stimuli. In addition, comparing the imitation and spatial compatibility tasks provided a way to localize regions activated selectively for action observation so that we could identify putative mirror neuron regions within the same paradigm and subjects (Friston et al. 2006).

With a standard activation analysis based on the General Linear Model (GLM), we initially identified a specific imitation control network that was consistent with previous studies and included the frontal node of the MNS. Following this, we used dynamic causal modeling (DCM), a method of modeling effective connectivity, to test a set of alternative hypotheses about causal interactions between imitation control regions. We tested set of models aiming to determine (1) whether the mPFC detects imitative conflict, as proposed by the shared representations model; (2) whether coupling between prefrontal regions and the MNS is stronger when control is required, as would be expected if imitation control involves modulation of MNS activity; and (3) which prefrontal control region interacts with the MNS.

2. METHODS

2.1 Participants

25 adult subjects (15 female; age 19–39) were recruited through advertisement in the university newspaper and free online bulletins. All subjects were right-handed, had normal or corrected-to-normal vision, no history of neurologic or psychiatric disorders and were not taking psychoactive medications. Subjects were compensated for their participation and the study was approved by the UCLA Institutional Review Board. One subject was excluded from analyses for a structural abnormality and four additional subjects were excluded based on quality control criteria: two reported falling asleep during scanning and failed to respond

on more than 15% of trials in two or more runs and two had excessive head motion (more than 10% of volumes with motion artifacts detectable by visual inspection in 2 or more runs). The remaining 20 subjects were included in data analysis, with 17 subjects entering the DCM analysis (3 did not show reliable activation maxima in one or more of the 4 ROIs).

2.2 Behavioral Paradigm

Participants performed a simple reaction time task modified from Brass et al. (2001) to include both automatic imitation and spatial compatibility components (Figure 1). Subjects lifted their index or middle finger as soon as they detected movement in a video stimulus. The required response (index or middle finger) was indicated by a written instruction before each block of videos. For the automatic imitation blocks, videos depicted a hand lifting either the index or middle finger, such that the video was either imitatively congruent with respect to the predefined response finger (e.g. index finger video on a trial where the subject was instructed to lift their index finger) or incongruent (e.g. middle finger video on a trial where the subject was instructed to lift their index finger). Spatial compatibility blocks were identical except that videos depicted a moving black dot instead of a finger. The trajectory of the dot was similar to the trajectory of the fingertip in the imitative stimuli. Thus, the action was congruent or incongruent with respect to the left-right spatial location of the dot, but no action observation or imitation was involved. The resulting 2 ×2 design (cue type ×congruency) consists of four conditions: Imitative Congruent (ImC), Imitative Incongruent (ImI), Spatial Congruent (SpC), and Spatial Incongruent (SpI).

The first frame of all four trial types was the same, and the duration was jittered between 500 and 2000ms in 500ms steps so that participants could not anticipate movement onset (i.e. the go signal). Then, the movement of either a finger or dot was presented as three 34ms frames, followed by a final frame showing the finger or dot in the raised position for 900ms. A blank blue screen marked the end of the response window and trial. This blue inter-trial interval (ITI) was between 500 and 2000 ms (again in 500 ms steps) depending on the length of the first frame, so that the inter-stimulus interval was always 3.5 seconds. In addition to the 4 task conditions, “null” trials were included for measurement of a passive baseline and to improve detection power by jittering the interval between successive trial onsets. Null trials were the same length as task trials (3.5 s) and identical to the blue ITI. Therefore, they were perceived simply as longer ITIs and were not explicitly signaled to subjects. The trial order was optimized using a genetic algorithm (Wager and Nichols, 2003) for the efficiency of Incongruent > Congruent contrasts for each cue type (simple effects of congruency) with the following constraints: Within each cue type, each trial type followed every other type with equal probability and no more than 3 trials of the same condition occurred in a row.

Trials were presented in a mixed block/event-related design (Figure 1B). Each 16-second block began with a 2 second instruction screen (“Lift your INDEX FINGER when the FINGER[DOT] moves” or “Lift your MIDDLE FINGER when the FINGER[DOT] moves”) followed by four 3.5-second trials. Blocks consisted of all imitative or all spatial cues, but middle and index stimuli were presented randomly within a block so that the congruency (i.e. the need for control) was unpredictable. Imitation and spatial blocks alternated and the instructed finger movement changed every two blocks, so that subjects lifted the same finger

for an imitative and a spatial block and then switched fingers for the next two blocks. The response finger and cue type for the first block in each run were counterbalanced across runs and subjects.

Response times for each condition were measured with respect to the onset of movement in the video. Subjects held down two buttons on a response box with the index and middle fingers whenever they were not responding. A button was released when subjects performed the finger lifting response, and the stimulus presentation computer recorded button release times.

2.3 Procedure

Immediately prior to scanning, each subject was familiarized with the task in a brief practice session. The experiment comprised 80 trials of each of the four conditions as well as 80 null trials during a single scanning session. The session was divided into 5 runs lasting 5 minutes 20 seconds each, between which the subjects were allowed a short break. Each scan was preceded by a reminder of the instructions: “Remember, as soon as you see movement of either the fingers or dots in the video, lift the designated finger as quickly as you can.” In addition, two structural scans were acquired.

2.4 MRI data acquisition

Images were acquired on a Siemens 3-T Trio MRI scanner (Erlangen, Germany). The five functional runs consisted of 160 T2*-weighted echoplanar images (EPIs) [repetition time (TR) 2000 ms; echo time (TE) 28 ms; flip angle=90°; 34 slices; slice thickness 4 mm; interleaved slice acquisition; matrix 64 ×64; FOV 192 mm]. To allow for T1 equilibrium, the first 2 volumes of each run were automatically discarded by the scanner before task initiation. Two sets of structural images were also acquired for registration of functional data: a T2-weighted matched-bandwidth high-resolution scan with the same slice prescription as the EPI [repetition time (TR) 5000 ms; echo time (TE) 34 ms; flip angle=90°; 34 slices; slice thickness 4 mm; matrix 128 ×128; FOV 192 mm]; and a T1 weighted magnetization prepared rapid-acquisition gradient echo image (MPRAGE) [TR, 1900 ms; TE 2.26 ms; flip angle = 9°; 176 sagittal slices; slice thickness 1 mm; matrix 256 ×256; FOV 250 mm]. Visual stimuli were timed and presented with Presentation software (Neurobehavioral Systems, Albany, CA) through magnet-compatible LCD goggles.

2.5 fMRI Activation: General Linear Model (GLM)

In the first stage of analysis, a conventional GLM was performed to identify regions involved specifically in controlling automatic imitation. Image preprocessing and data analysis were performed with FSL version 4.1.4 (Centre for Functional Magnetic Resonance Imaging of the Brain software library, www.fmrib.ox.ac.uk/fsl) (Smith et al. 2004). Functional images were realigned to the middle volume to compensate for any head motion using MCFLIRT (Jenkinson et al. 2002). After motion correction, volumes were visually inspected for motion artifacts. Runs in which greater than 10% of volumes displayed striping artifacts were excluded from analysis. As previously mentioned, 2 subjects who had 2 runs meeting this criteria were excluded from analysis. In three subjects only one run met exclusion criteria; the remaining 4 runs for these subjects were included in analysis. After

motion correction, data were temporally filtered with a high-pass filter cutoff of 50 s and spatially smoothed with a 6 mm full width half maximum Gaussian kernel in three dimensions.

Statistical analyses were performed separately for each run using a general linear model (GLM) with fMRI Expert Analysis Tool (FEAT). Each trial type, convolved with a canonical double-gamma hemodynamic response function, was included as a regressor in the GLM. In addition, nuisance regressors were included for error trials, block instructions and the reaction time for each trial. The reaction time regressor was demeaned and orthogonalized with respect to EVs of interest. Trials for each condition were modeled as one-second events starting at video movement onset. Temporal derivatives were included for each regressor to account for variability in the hemodynamic response.

To identify regions involved in controlling automatic response tendencies for the two cue types we specified 3 contrasts. The simple effects of congruency (ImI-ImC and SpI-SpC) identified regions involved in overcoming the automatic response tendency evoked by each stimulus type. The cue type by congruency interaction, [i.e. the difference between congruency effects (ImI-ImC)-(SpI-SpC)], identified regions involved specifically in control of imitation, since this would subtract out the activation of any non-specific control regions involved in overcoming the spatially-compatible response tendency. Finally, we examined the main effect of cue type (Imitate - Spatial) for regions sensitive to action observation, regardless of congruency, to identify the MNS.

After contrast estimates were computed for each run in native subject space, they were registered to standard space (Montreal Neurological Institute, MNI) in three stages. The middle volume of each run of individual EPI data was registered first to the co-planar matched-bandwidth high-resolution T2-weighted image and subsequently, the co-planar volume was registered to the T1-weighted MPRAGE. Both of these steps were carried out using FLIRT (affine transformations: EPI to co-planar, 6 degrees of freedom; co-planar to MPRAGE, 6 degrees of freedom) (Jenkinson et al. 2002). Registration of the MPRAGE to MNI space (FSL's MNI Avg152, T1 2×2×2mm) was carried out with FLIRT (affine transformation, 12 degrees of freedom) and refined using FNIRT (non-linear transformation) (Jenkinson and Smith, 2001; Jenkinson et al. 2002). Contrast estimates for each subject were then computed by averaging over runs, treating runs as fixed effects.

The group level analysis was performed with a one sample t-test for each contrast using FSL's FLAME (FMRIB's local analysis of mixed effects) stage 1 and stage 2 with outlier de-weighting (Beckmann et al. 2003; Woolrich et al. 2004; Woolrich, 2008). Group images were thresholded at $Z > 2.3$ corrected for multiple comparisons using cluster-based Gaussian random field theory controlling for familywise error across the whole brain at $p = 0.05$. All analyses were performed across the whole-brain. However for the interaction analysis, discussion is limited to regions showing a significant simple effect of congruency, so that only regions showing a robust congruency effect for at least one cue type are considered control regions. This was accomplished by inclusively masking the interaction contrast by both simple effects of congruency after whole-brain statistical inference.

2.6 fMRI Effective Connectivity: Dynamic Causal Modeling (DCM)

With the cue type x congruency interaction contrast [(ImI-ImC)-(SpI-SpC) masked inclusively by the congruency effect for each cue type] (see Results) we identified 4 regions (mPFC, ACC, aINS and IFGpo) specifically involved in imitation control. We used DCM to examine effective connectivity between these regions and test a number of different models of imitative control. In the DCM approach used here, the brain is treated as a deterministic dynamic system. Models of causal interactions between task-relevant brain regions are compared within a Bayesian statistical framework to identify the most likely model out of those examined (Friston et al. 2003; Stephan et al. 2010). A bilinear state equation models neuronal population activity in each region of interest. Activity in a region is influenced by neuronal inputs from one or more connected regions and/or by exogenous, experimentally controlled inputs (i.e. task stimuli). Experimental inputs can influence the system in two ways: as “driving” inputs that elicit responses by directly affecting activity in a region (i.e. stimulus-evoked responses); or as “modulatory inputs” that change the strength of connections between regions (i.e. task-related changes in effective connectivity). Thus, with DCM one can compare a set of models differing in (1) which regions receive driving inputs (stimulus-evoked activity), (2) which regions are connected with one another and how they are connected (the endogenous connectivity structure) and (3) which of these connections receive modulating inputs (task-related changes in effective connectivity). Multiple models (hypotheses) are compared within a Bayesian statistical framework to identify the most likely model out of those examined given the observed data (Friston et al. 2003; Stephan et al. 2010).

Because DCM is not implemented in FSL, we used DCM10 within SPM8. To ensure that preprocessing of the data was consistent with the modeling procedures, we re-processed the data using a standard SPM processing stream and used this new preprocessed data for all DCM analysis steps. Although the SPM analysis showed very similar patterns to the FSL-derived GLM described above, it was not as sensitive, especially in the interaction contrast (Supplementary Figure 1 & Supplementary Table 1). Nonetheless, based on similarities with previous imitation control studies discussed in detail below, it is unlikely that this difference reflects false positives in the FSL analysis. While stronger group effects less sensitive to small differences in processing streams would be ideal, we did not have trouble locating individual subject peaks in our regions of interest using typical methods, so we proceeded with the DCM analysis even though SPM group effects were not as robust as FSL group effects. Several differences in FSL and SPM processing streams may have contributed to the difference in sensitivities. The methods for estimating autocorrelation differ between the packages, and differences in the estimation and success in modeling autocorrelation can affect variance and therefore t-value estimates. In addition, we employed a 2-stage model estimation analysis (Flame 1&2) in FSL, which increases sensitivity by refining variance estimates for all near-threshold voxels in the second stage (Beckmann et al. 2003; Woolrich, 2008).

For the DCM analysis data were preprocessed as follows: functional images were slice-time corrected (Kiebel et al. 2007), motion corrected with spatial realignment to the mean volume of the first run and coregistered to the MPRAGE structural scan. The MPRAGE was

processed using a procedure that combines grey and white matter segmentation, bias field correction and spatial normalization. The normalization parameters were then applied to the functional images. Finally the images were smoothed with a 6mm full-width half-maximum Gaussian kernel and resampled to 3×3×3mm voxels. In order to identify individual subject regions of interest in the reprocessed data, we again fit a GLM using SPM8 for each subject with separate regressors for each condition, errors, block instructions and reaction time. Temporal derivatives and motion parameters were also included in the model. An F-test across all conditions and temporal derivatives was specified to correct extracted timeseries, effectively removing variance associated with motion parameters.

2.6.1 Hypotheses and Model Specification—We constructed models defining exogenous inputs to and endogenous connections between four regions of interest (ROI) identified to be involved specifically in imitation control (Figure 2C). As described in detail in the Results section, these ROIs included a "prefrontal control network"—medial prefrontal cortex (mPFC), anterior cingulate cortex (ACC) and left anterior insula/frontal operculum (aINS)—and the frontal node of the MNS—left inferior frontal gyrus, pars opercularis (IFGpo). The construction of our model space was motivated by three central questions: (1) Does conflict detection occur in the mPFC (consistent with the shared representations hypothesis), in the ACC (consistent with the conflict monitoring hypothesis) or in the MNS? (2) Which prefrontal control region interacts with the MNS? (3) Is coupling between the control network and MNS node stronger when control is required than when it is not?

In all models (see Figure 3A), the MNS node (IFGpo) received action observation (i.e. imitative trials) as a driving input consistent with the response of this region and functional properties of the MNS and IFGpo (di Pellegrino et al. 1992; Iacoboni et al. 1999). In addition, the three regions comprising the control network were connected to one another with all combinations of either 2 or 3 bidirectional connections consistent with anatomical evidence for connections between these regions in primates (Augustine 1996; Petrides and Pandya 2007; Yeterian et al. 2012). This allowed identification of the most likely functional connectivity structure within the prefrontal control network before turning to questions about imitative conflict detection and resolution. Thus, there were 4 base models (Figure 3A and Supplementary Figure 2A), across which we varied which prefrontal region was connected to the IFGpo (Figure 3B), and which regions and connections were affected by imitative conflict (Figure 3C), to answer our three questions (see Supplementary Figure 2B for depiction of the expanded model space).

First, endogenous connectivity structures were defined to determine which of the prefrontal control regions interacts with the MNS. Three separate variations were created in which each one of the three control regions was connected directly to the IFGpo (Figure 3B). When crossed with the 4 base models detailed above (Figure 3A and Supplementary Figure 2A), this yielded a total of 12 possible endogenous connectivity structures in the full model space.

Next, we varied which node detects (i.e. which region is responsive to) imitative conflict (defined as the difference between incongruent and congruent trials) (Figure 3C). To test the

shared representations theory, conflict drove activity in mPFC, because this region is thought to be engaged when observed and executed actions activate conflicting motor representations (Brass et al. 2009b). In a variation of this model, conflict acted as a driver of the ACC. This was based on the influential conflict monitoring theory from the broader cognitive control literature in which the ACC is proposed to detect response conflict (Botvinick et al. 2004; Carter and van Veen, 2007) and provide a signal to lateral prefrontal regions to implement conflict resolution. In addition, we included models in which conflict drove both the mPFC and ACC to test the possibility that these regions act in concert in the detection of imitative conflict. This would be consistent with a scenario in which the mPFC detects imitative conflict specifically, whereas the ACC is a more general response conflict detector and therefore contributes across a variety of tasks. Finally, we tested a fourth alternative hypothesis in which conflict is detected in the MNS. The IFGpo receives inputs representing both the observed action and the conflicting planned action, so it is possible that conflict is detected where conflicting representations first arise. The presence of this conflict could then signal prefrontal cortex to reinforce the intended action or inhibit the externally-evoked action. These 4 variations in the location of conflict as a driving input (mPFC, ACC, mPFC+ACC, IFGpo) were crossed with the 12 endogenous connectivity structures creating 48 models.

Finally, we included another set of the identical 48 models but with the addition of conflict as a modulator of the connection from the prefrontal control network to the IFGpo (Figure 3C, dotted lines). This allowed us to determine whether the influence of prefrontal control regions on the frontal node of the MNS is greater when imitative control is implemented, as would be expected if the interaction effect relates to resolving the imitative conflict. Thus, the total model space was comprised of 96 models built as a factorial combination of 12 connectivity structures, 4 locations of conflict driving input, and 2 modulating inputs (i.e. the presence or absence of conflict as a modulator).

2.6.2 Time series extraction—The selection of subject-specific ROIs in the mPFC, ACC, aINS and IFGpo was based on local maxima of the relevant contrasts from the GLM analysis (Stephan et al. 2010). For the prefrontal control network we identified the local maxima in the imitative congruency contrast (ImI-ImC) nearest the interaction peaks (mPFC: −3 44 22; ACC: −3, 14 34; aINS: −39, 17 −5). Although guided by the interaction, we used the imitative congruency contrast for localization of individual subject ROIs so that control nodes were defined by their contribution to imitative control and not influenced by any effect of spatial congruency. For the IFGpo we used the main effect of cue type to define the node by its mirror properties, again locating the local maxima nearest the interaction peak (MNI −39, 14, 25). Nonetheless, parameter estimates from the resulting IFGpo individual subject ROIs still showed the imitative congruency effect as expected based on the GLM [$t(16)=2.5$, $p = 0.02$].

Individual subject ROIs were defined for each region as all supra-threshold voxels ($p<0.05$, uncorrected) within a 6mm sphere centered on the peak nearest to the group coordinate. Peaks were required to be within 16mm of the group coordinate and the four peaks for each subject were separated by at least twice the smoothing kernel (12mm). Finally, peaks were also within the following anatomical regions as defined by the Harvard-Oxford Probabalistic

Atlas: mPFC – cingulate or paracingulate gyrus; ACC – anterior cingulate gyrus (more posterior than mPFC peaks); IFGpo – inferior frontal gyrus, pars opercularis; aINS – anterior insula or frontal operculum. Using this procedure, one or more peaks could not be identified for 3 of the 20 subjects, so these subjects were excluded from the DCM analysis. This number is typical (e.g. Wang et al. 2011b) for a study including several ROIs. The resulting mean coordinates for each ROI were: mPFC (−2, 42, 23); ACC (−3, 15, 34); aINS (−35, 16, −4); and IFGpo (−39, 15, 25). Regional timeseries were extracted from each ROI as the first eigenvariate of responses and adjusted for effects of interest F-test (variance due to motion removed).

2.6.3 Model Selection—We used Bayesian model selection (BMS) amongst individual models (Stephan et al. 2009; Stephan et al. 2010) with inference over families of models (Penny et al. 2010) to identify the most likely model structure from the model space described above. This was done in two stages. First, for each subject the model evidence was computed for each model and each run using the negative free-energy approximation to the log-model evidence. The free-energy metric for model evidence balances model fit and complexity taking into account interdependencies amongst parameters and has been found to outperform other more conventional methods of model scoring for model comparison (Penny et al., 2012). The subject-specific sums of log evidences across runs (equivalent to a fixed effects analysis across runs) were entered into group random effects (RFX) BMS to identify the most likely model across subjects (Stephan et al. 2009). This procedure requires that all subjects have the same number of runs (c.f. SPM DCM manual), so only the first four runs were used for DCM for all subjects (as mentioned previously, three subjects had only four usable runs due to motion artifacts).

The RFX approach to group model selection was preferred over fixed effects because it does not assume that the optimal model is the same for all subjects. This is appropriate in studies of higher cognitive functions where there may be heterogeneity in strategy or neural implementations of task performance (Stephan et al. 2010). Results from random effects model comparisons are understood in terms of the exceedance probability (the probability that a particular model is more likely than any other model tested) and the expected posterior probability (the likelihood of obtaining the model for a random subject from the population) (Stephan et al., 2009). Both measures sum to 1, so the exceedance and expected posterior probabilities are reduced as the model space increases. As such, including multiple models makes it less likely that a single model will dominate the RFX analysis. Family level inference has been introduced as a technique to deal with this issue of dilution from a large number of models, which is particularly problematic when different models have many shared parameters and when different subjects use slightly different models (Penny et al. 2010). With this technique, models are divided into groups (families) according to the presence of shared features, which allows inference about these general features and can be used narrow the search for a best model.

Here, we divided models into families based on the intrinsic connectivity structure in a stepwise manner. First, we identified the family with the preferred prefrontal connectivity structure (see Supplementary Figure 2A), limiting further inference about MNS interactions and conflict detection to the set of most plausible models. Next, we entered models from the

winning family (fully connected prefrontal network) into a second set of BMS analyses to answer the questions outlined previously. The remaining models were divided into 3 families each of which included models sharing the same prefrontal→MNS connection (aINS→IFGpo, ACC→IFGpo, or mPFC→IFGpo depicted in Figure 3B; rows in Supplementary Figure 2B), but differing in the location of conflict driving and modulatory inputs. This allowed us to determine which prefrontal control region is most likely interacting with the MNS, removing uncertainty about the influence of conflict on the system. Models in the winning family were then compared to examine conflict processing in the system. To summarize individual parameters of the winning model, we performed one-sample t-tests on the maximum a posteriori parameter estimates across subjects to determine whether the parameters were significantly different from zero.

3. RESULTS

3.1 Behavioral Results

Mean reaction time (RT) and accuracy were calculated for correct responses in each condition for each subject, and then averaged across subjects. Trials with RT greater than 2 standard deviations above the mean were considered outliers and excluded from analysis (1.1–3.8% of trials per subject). RT analysis was carried out using a 2 (Cue type: imitative, spatial) ×2 (Congruency: congruent, incongruent) repeated measures ANOVA. This revealed a main effect of congruency [$F(1,19)=38.1$, $p<0.001$], demonstrating that responses for incongruent trials (mean=311ms, SD=40) were slower than congruent trials (mean=298ms, SD=32) (Figure 4). There was also a main effect of cue type [$F(1,19)=36.0$, $p<0.001$], with responses being faster for spatial (mean=298ms, SD=36) than imitative cues (mean=310ms, SD=36ms). Earlier detection of movement onset may have occurred for the dots due to greater contrast between the dot and background. Importantly, there was no interaction between cue type and congruency [$F(1,19)=0.27$, $p=0.6$)], confirming that congruency effects were of similar size regardless of the cue type (spatial: 12ms; imitative: 13ms). As such, differences in congruency effects in brain activation cannot be attributed to differences in the presence or magnitude of the interference effect. In a similar ANOVA on accuracy data, no significant effects were observed as accuracy was near ceiling in all four conditions (>97%).

3.2 GLM Results

Neuroimaging data revealed a dissociation between congruency effects for the two cue types. For imitative stimuli, the simple effect of congruency (ImI - ImC) showed activation in frontal and parietal regions, as well as the cerebellum and caudate (Figure 2A, Supplementary Table 2). Consistent with previous studies of imitation control (Brass et al. 2001; Brass et al. 2005; Brass et al. 2009a; Bien et al. 2009b; Spengler et al. 2009; Wang et al. 2011b), large clusters in the frontal lobes were observed in medial prefrontal cortex (mPFC) extending into the frontal pole, anterior cingulate cortex (ACC) and bilateral anterior insula (aINS) extending into the frontal operculum and orbito-frontal cortex. In addition there was bilateral activation in the IFG pars opercularis (IFGpo) extending posteriorly into precentral gyrus. In contrast to findings for imitative cues, no regions showed a significant congruency effect for spatial cues. This was true even when the

threshold was lowered to $z > 1.7$ to be more sensitive to small differences and when using a most liberal post-hoc ROI approach: One-sample t-tests on the parameter estimates for the contrast (SpI-SpC) were extracted from each of the regions showing an imitative congruency effect. No regions approached significance for spatial congruency effects even by this liberal method (all p-values greater than 0.2).

Consistent with the qualitative difference between imitative and spatial congruency effects, a direct comparison of the congruency effects confirmed a dissociation between control processes depending on the cue type. Significantly greater congruency effects for imitative compared to spatial cues [assessed with the Cue Type x Congruency interaction contrast (ImI-ImC) - (SpI-SpC)] were detected in multiple frontal regions: the ACC, mPFC extending into the frontal pole, left IFGpo and left aINS extending into the frontal operculum and OFC (Figure 2C), Supplementary Table 3).

Finally, to localize potential mirror neuron regions, we examined the cue type main effect (Imitate - Spatial). As expected, a fronto-parietal network commonly observed during action observation and imitation tasks was more active for imitative cues compared to spatial cues (Iacoboni et al. 1999). The network involved bilateral inferior frontal gyrus, pars opercularis (IFGpo) extending into ventral premotor cortex (PMv) and the superior parietal lobes (Figure 2B; Supplementary Table 4). To determine whether these mirror neuron regions were modulated during resolution of imitative conflict, we compared the cue type main effect with the imitative congruency effect. An overlay of the two contrasts demonstrates that the right parietal and bilateral IFGpo regions were sensitive to action observation and also modulated by conflict. The main effect of cue type strongly suggests that IFGpo represents the frontal node of the human MNS, especially in the context of previous work. The IFGpo is causally involved in both automatic imitation (Catmur et al. 2009) and motor resonance phenomena (Avenanti et al. 2007) and this region is also likely to be a human homologue of monkey area F5 where mirror neurons have been recorded in monkeys (Rizzolatti and Arbib, 1998). The imitative congruency effect observed in the same region suggests that this frontal MNS node is modulated during imitation control.

3.3 DCM Results

We sequentially partitioned the model space into families (groups of models which shared common features) to zero in on a winning model via bayesian model comparisons. We first used family level inference to find the preferred prefrontal connectivity structure by partitioning models into four families with each family sharing the same set of prefrontal connections. Results indicated that the fully connected prefrontal control network was more likely than the more sparsely connected prefrontal networks (exceedance probability= 0.88; expected posterior probability = 0.48; Table 1). An exceedance probability more than 10 times higher than the next highest family provides strong evidence that the fully-connected prefrontal network is better than other prefrontal connectivity structures.

Next, we entered models from the winning family—those with fully connected prefrontal nodes—into a second family-level comparison to determine which of the 3 prefrontal control regions (mPFC, ACC and aINS) interacted with the frontal MNS node (IFGpo). Models in each family shared the same prefrontal→MNS connection (aINS→IFGpo,

ACC→IFGpo or mPFC→IFGpo). Results demonstrated that the IFGpo is substantially more likely to be connected to the aINS (exceedance probability p=0.82; expected posterior probability = 0.50) than either the ACC (exceedance probability = 0.14; expected posterior probability = 0.30) or the mPFC (exceedance probability p=0.03; expected posterior probability = 0.20) (Figure 5, top left; Table 1).

Finally, we performed BMS on the 8 models in the winning family—models with the aINS to IFGpo connection—to determine more specifically how conflict processing occurs within the system. The models varied according to which region is driven by conflict (IFGpo, ACC, mPFC or ACC+mPFC) and whether top-down influence of the prefrontal control network on the IFGpo is modulated by conflict. Model 8 clearly out-performed the other 7 models, with an exceedance probability of 0.88 and expected posterior probability of 0.40 (Figure 5, bottom left; Table 1). In this model (Figure 5, right) both the ACC and mPFC are driven by conflict. Furthermore, the connection between the aINS and IFGpo is modulated by conflict, with greater connectivity when conflict resolution is required than when there is no conflict. This model is more likely than any of the alternatives, however it is interesting to note that the second highest model was identical except conflict drove only the ACC (model 7). The total exceedance probability of these two models together was greater than 0.99 with an expected posterior probability together of 0.73, providing strong evidence that conflict detection occurs in the medial frontal regions rather than first being detected in the MNS and then propagating to the frontal cortex. Similarly, these models both include conflict modulation of the aINS to IFGpo connection whereas the identical models without this modulation have exceedance probabilities much lower than 0.01.

For completeness, averages of posterior parameter estimates across subjects for the winning model are depicted in Figure 5. The endogenous connections from the mPFC→aINS and ACC→aINS were significantly greater than zero (both p = 0.001). In addition, all driving inputs were significant: conflict driving input to the ACC (p = 0.001); conflict → mPFC ($p<0.001$); action observation → IFGpo (p = 0.048). Conflict modulation of the aINS→IFGpo connection also approached significance (p=0.073). All but the latter two tests (conflict modulation and action observation → IFGpo) survive Bonferroni correction for the multiple parameters tested ($p<0.004$), however Bonferroni correction is quite a conservative approach in this case, since the parameter estimates are not independent (Stephan et al. 2010). Other individual parameters did not reach significance, including the aINS→IFGpo connection.

4. DISCUSSION

We examined neural mechanisms of imitation control using fMRI and dynamic causal modeling. Subjects performed a predefined finger movement in response to video stimuli depicting either an action (finger movement) or a dynamic spatial stimulus (a moving dot). As expected, for both cue types people were slower to respond when the stimulus and response were imitatively or spatially incongruent compared to when they were congruent, presumably due to the recruitment of additional resources to control the automatic response tendency on incongruent trials. In contrast to the very similar behavioral congruency effects, neural activity demonstrated a dissociation between imitative and spatial congruency effects,

revealing a set of regions involved specifically in imitation control. We used dynamic causal modeling to explore interactions between these regions and test several hypotheses about mechanisms of imitation control. Our results suggest that the mPFC and ACC detect conflict between observed and planned actions and the anterior insula interacts with the MNS, with some evidence for stronger coupling in the face of imitative conflict. Below, we begin by discussing results from the GLM analyses in the context of previous literature and then propose an expansion of the shared representations model of imitation control to incorporate the DCM findings.

Four regions—the ACC, mPFC, aINS and IFGpo—showed a significant interaction between congruency and cue type, demonstrating a congruency effect for imitative cues but not for symbolic spatial cues that moved with a similar trajectory. This cannot be attributed to an absence of response conflict altogether for the spatial cues. Congruency effects for the two cue types were intentionally equated to rule out the possibility that differences in neural correlates of the congruency effects are due to different degrees of conflict and control (Aicken, 2007). Instead, similar behavioral manifestations of conflict suggest that similar degrees of automatic response tendencies were evoked by the two stimulus types, and therefore, neural correlates of this conflict are likely related to the particular content of the stimuli rather than to the degree of conflict. Thus, the role of these regions in imitation control is distinct from any potential role in controlling prepotent response tendencies induced by non-social, symbolic stimuli.

This dissociation between imitation and spatial compatibility is in line with previous behavioral work demonstrating distinctions between imitative and spatial compatibility (Brass et al. 2001; Heyes et al. 2005; Bertenthal et al. 2006b; Catmur and Heyes, 2010; Jiménez et al. 2012). However, previous neuroimaging support of these findings has been mixed. Crescentini and colleagues (Crescentini et al. 2011) compared imitation and spatial congruency effects in similar tasks. However, they did not find behavioral congruency effects for half of responses and also did not observe fMRI congruency effects for either cue type. This may have been due to the task instructions, which required that participants withhold their response until the end of the video stimulus rather than responding immediately. In this situation, it is possible that inhibition of the prepotent response occurred on both congruent and incongruent trials, as subjects waited for the appropriate time to respond. In another study comparing imitative and spatial compatibility (Bien et al. 2009b) only the frontal operculum was demonstrated to show a greater imitative than spatial congruency effect. The relevant interaction contrast, however, was not performed across the whole brain so it is possible that a wider network similar to the present study showed similar effects.

The regions identified here as specifically involved in imitation control are consistent with previous studies that did not control for spatial compatibility. Although the mPFC has received the most attention (Brass et al. 2001; Brass et al. 2005; Brass et al. 2009a; Spengler et al. 2009; Wang et al. 2011b), the other regions have also been implicated in studies reporting whole brain imitation congruency effects (Brass et al. 2001; Brass et al. 2005; Bien et al. 2009a; Wang et al. 2011b). The anterior insula/frontal operculum region observed here is similar to that found in multiple previous studies (Brass et al. 2005; Bien et al. 2009a;

Wang et al. 2011b) despite receiving relatively little attention in theories of imitation control. The consistency of involvement of this region in imitation control may have been obscured by differences in nomenclature. For example, a cluster with MNI coordinates (45, 26, −7) falls within our aINS cluster, but was hypothesized to be part of the MNS and thus labeled the IFG (Wang et al. 2011b). Similarly, Brass and colleagues reported activation in Talairach (41 5 3), which is slightly posterior to the anterior insula cluster we observed. Bien and colleagues (2009a) also identified a region in the frontal operculum, however coordinates are not reported. Thus, activity around the junction of the anterior insula and frontal operculum seems relatively consistent across a variety of imitation control tasks.

The observation of IFGpo involvement in imitative control is especially intriguing in the context of previous literature on imitation and the MNS. The anatomical location of the congruency effect—the very posterior part of the inferior frontal gyrus and extending into the ventral premotor cortex—is one of the proposed human homologues of the frontal node of the monkey MNS (Rizzolatti and Arbib, 1998) and the region is commonly activated in studies of action observation and imitation in humans (Caspers et al. 2010). However, even more importantly, in our task the same region showed a main effect of cue type, indicating sensitivity to action observation as one would expect of a mirror neuron region. This finding is consistent with several previous imitation control studies that have argued for modulation of the MNS (Spengler et al. 2009; Wang et al. 2011b). However, these claims were based on anatomical parallels to previous studies of the MNS rather than identifying the MNS in the same study. The inclusion of a spatial compatibility task that was very similar to the imitation task except for the absence of action observation, allowed us to test the hypothesis that the MNS is involved in imitation control more directly. Our results support this hypothesis, and led us to explore functional interactions between the prefrontal control regions and the frontal node of the MNS using dynamic causal modeling.

We were interested specifically in how the set of 3 prefrontal control regions (mPFC, ACC, aINS) interacts with the MNS during imitation control and how conflict processing occurs in the network. In the winning model the aINS interacted with the IFGpo, this connection was modulated by imitative congruency, and activity in the mPFC and ACC was driven by imitative conflict. This model of imitative control is consistent with the shared representations theory in that the mPFC is involved in detecting conflict between self-generated and other-generated motor activity (Brass et al. 2009b). However the DCM suggests an extension of the shared representations model, which has not provided a detailed account of how conflict between the observed and intended action is subsequently resolved.

In the winning model the aINS input to the MNS is modulated by conflict. Although a univariate test of the parameter did not quite reach significance, the fact that the top models included the modulation suggests that it does contribute to model fit, and provides at least some support for the hypothesis that this interaction is involved in resolving conflict. A closer look at the aINS→IFGpo interaction provides some insight into potential prefrontal-MNS interactions in conflict resolution. The endogenous connectivity between aINS and IFGpo was not different from 0, but a modulation of this connection occurs in response to conflict. This provides at least tentative evidence that the aINS interacts with the MNS activity only when conflict occurs. Furthermore, the direction of modulating input was

negative, suggesting that aINS suppresses MNS activity in response to conflict. Further support for this hypothesized interaction is necessary given that we observed only a trend in the parameter, but this pattern would be consistent with models of conflict processing which often argue for inhibitory mechanisms, both in the context of automatic imitation (Brass et al. 2009b) and in more general response conflict tasks (Kornblum et al. 1990; de Jong, 1995; Miller and Cohen, 2001; Burle et al. 2004; Ridderinkhof et al. 2004).

Within the prefrontal control network, both the ACC and mPFC were driven by conflict in the winning model. In the next best model, the ACC alone was driven by conflict. Thus, both medial prefrontal regions seem to play some role in detecting imitative conflict. While mPFC seems to be involved only for the more specific case of imitation in which conflict is related to agency (Brass et al. 2001; Brass et al. 2005; Brass et al. 2009a; Spengler et al. 2009; Wang et al. 2011b), the ACC is activated by a wide range of conflict tasks (van Veen et al. 2001; Bunge et al. 2002; Egner and Hirsch, 2005; Wendelken et al. 2009; Botvinick et al. 2004; Carter and van Veen, 2007) and therefore may represent a more multi-modal and general conflict detector. In addition, the aINS region could also represent a more domain-general node of the network, as this region is also implicated in both response inhibition and conflict resolution, including stop-signal, go/no-go, Stroop and flanker tasks (Wager et al. 2005; Nee et al. 2007; Levy and Wagner, 2011).

Based on these similarities, control of imitation may involve interactions between general cognitive control mechanisms and a more specific imitation-relevant network. The ACC and aINS may be involved in detecting and resolving conflict regardless of the source of the conflict, but interact with different networks depending on the nature of conflict. In the context of imitation and action observation, the mPFC would be responsible for determining agency and thereby indicate to the aINS which representation reflects the intended action; the MNS—where conflict first arises—would be the target of top-down mechanisms of conflict resolution. This model is in line with a parsimonious and generalizable framework whereby a general conflict resolution system interacts with the system in which the conflicting representations occur. Indeed this is consistent with several previous studies aiming to dissociate conflict processes. Egner and others have demonstrated modulation of the visual system in tasks involving stimulus conflict (Egner and Hirsch, 2005; Egner et al. 2007), modulation of the amygdala in tasks with emotional conflict (Etkin et al. 2006; Egner et al. 2008), and motor modulation in tasks with response conflict (Egner et al. 2007; Stürmer et al. 2002).

Finally, we should note that our model of imitation control differs somewhat from a recent study that also used DCM to examine imitation control mechanisms, albeit in the context of direct and averted gaze (Wang et al. 2011b). That study was motivated by the observation that imitation interference effects were reduced when a video showed someone looking at the participant as compared to when someone was looking away from the participant. This behavioral effect was proposed to reflect reduced top-down control on automatic imitation in response to the social gaze stimulus (Wang et al. 2011a). Results from their DCM suggested that the interaction between imitation control and gaze was due to mPFC-mediated modulation of visual inputs to the frontal node of the MNS. The interpretation of MNS involvement in this study is tenuous, given that an inferior frontal region assumed to

be the frontal MNS was identified in an interaction between imitative congruency and eye gaze and was quite far anterior. However, a more interesting explanation for potentially different control mechanisms in the two studies is the difference in the timing of imitative control. In the gaze experiment, gaze was directed toward or away from the participant before the imitative task. Thus, the effect of gaze on imitative control is likely to occur in advance of the imitative stimulus, in a preparatory manner. In contrast, in the current study congruency effects must reflect control exerted in response to the imitative conflict rather than in preparation for conflict, since the need for control was unpredictable. Differences between preparatory and reactive control mechanisms have been observed in other domains (Braver et al. 2007; Boy et al. 2010; Braver, 2012) and are plausible in this context as well. For example, in a situation where imitation control can be implemented in advance (e.g. Cross and Iacoboni, 2011), it could occur by changing motor system sensitivity to action observation through modulation of input to the MNS (as described by Wang et al. 2011b; see also Heyes, 2011). However, when preparation to avoid imitation is not possible or is incomplete then some reactive control mechanism must deal with the unwanted motor activation that arises in response to action observation—in this case it may be too late to modulate the visual input, and instead the motor output of the MNS may be modulated as described in the current study.

4.3 Conclusions

In summary, our results support the view that imitative control relies on neural mechanisms that are at least partially distinct from those involved in overcoming automatic response tendencies evoked by non-social stimuli due to spatial compatibility. In addition, we propose an extension of the shared representations hypothesis of imitation control (Brass et al. 2009a): Once the mPFC and ACC detect conflict between the planned and the observed actions, enforcing the intended action involves interactions between the anterior insula and frontal node of the human MNS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

For generous support the authors wish to thank the Brain Mapping Medical Research Organization, Brain Mapping Support Foundation, Pierson-Lovelace Foundation, The Ahmanson Foundation, William M. and Linda R. Dietel Philanthropic Fund at the Northern Piedmont Community Foundation, Tamkin Foundation, Jennifer Jones-Simon Foundation, Capital Group Companies Charitable Foundation, Robson Family and Northstar Fund. The project described was supported by Grant Numbers RR12169, RR13642 and RR00865 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH); and by the NIH under Ruth L. Kirschstein National Research Service Award (1F30MH091808-01A1). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH.

References

Aicken MD, Wilson AD, Williams JH, Mon-Williams M. Methodological issues in measures of imitative reaction times. Brain Cogn. 2007; 63:304–308. [PubMed: 17070640]

Amodio DM, Frith CD. Meeting of minds: The medial frontal cortex and social cognition. Nat Rev Neurosci. 2006; 7:268–277. [PubMed: 16552413]

Augustine JR. Circuitry and functional aspects of the insular lobe in primates including humans. Brain Res Brain Res Rev. 1996; 22:229–244. [PubMed: 8957561]

Avenanti A, Bolognini N, Maravita A, Aglioti SM. Somatic and motor components of action simulation. Curr Biol. 2007; 17:2129–2135. [PubMed: 18083517]

Baldissera F, Cavallari P, Craighero L, Fadiga L. Modulation of spinal excitability during observation of hand actions in humans. Eur J Neurosci. 2001; 13:190–194. [PubMed: 11135017]

Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in FMRI. Neuroimage. 2003; 20:1052–1063. [PubMed: 14568475]

Bertenthal BI, Longo MR, Kosobud A. Imitative response tendencies following observation of intransitive actions. J Exp Psychol Hum Percept Perform. 2006a; 32:210–225. [PubMed: 16634666]

Bertenthal BI, Longo MR, Kosobud A. Imitative response tendencies following observation of intransitive actions. J Exp Psychol Hum Percept Perform. 2006b; 32:210–225. [PubMed: 16634666]

Bien N, Roebroeck A, Goebel R, Sack AT. The brain's intention to imitate: The neurobiology of intentional versus automatic imitation. Cereb Cortex. 2009a; 19:2338–2351. [PubMed: 19153108]

Bien N, Roebroeck A, Goebel R, Sack AT. The brain's intention to imitate: The neurobiology of intentional versus automatic imitation. Cereb Cortex. 2009b; 19:2338–2351. [PubMed: 19153108]

Bird G, Leighton J, Press C, Heyes C. Intact automatic imitation of human and robot actions in autism spectrum disorders. Proc Biol Sci. 2007; 274:3027–3031. [PubMed: 17911053]

Borroni P, Montagna M, Cerri G, Baldissera F. Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. Brain Res. 2005; 1065:115–124. [PubMed: 16297887]

Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: An update. Trends Cogn Sci. 2004; 8:539–546. [PubMed: 15556023]

Boy F, Husain M, Sumner P. Unconscious inhibition separates two forms of cognitive control. Proc Natl Acad Sci U S A. 2010; 107:11134–11139. [PubMed: 20534462]

Brass M, Bekkering H, Prinz W. Movement observation affects movement execution in a simple response task. Acta Psychol (Amst). 2001; 106:3–22. [PubMed: 11256338]

Brass M, Bekkering H, Wohlschläger A, Prinz W. Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. Brain Cogn. 2000; 44:124–143. [PubMed: 11041986]

Brass M, Derrfuss J, Matthes-von Cramon G, von Cramon DY. Imitative response tendencies in patients with frontal brain lesions. Neuropsychology. 2003; 17:265–271. [PubMed: 12803432]

Brass M, Derrfuss J, von Cramon DY. The inhibition of imitative and overlearned responses: A functional double dissociation. Neuropsychologia. 2005; 43:89–98. [PubMed: 15488909]

Brass M, Ruby P, Spengler S. Inhibition of imitative behaviour and social cognition. Philos Trans R Soc Lond B Biol Sci. 2009a; 364:2359–2367. [PubMed: 19620107]

Brass M, Ruby P, Spengler S. Inhibition of imitative behaviour and social cognition. Philos Trans R Soc Lond B Biol Sci. 2009b; 364:2359–2367. [PubMed: 19620107]

Brass M, Zysset S, von Cramon DY. The inhibition of imitative response tendencies. Neuroimage. 2001; 14:1416–1423. [PubMed: 11707097]

Braver TS. The variable nature of cognitive control: A dual mechanisms framework. Trends Cogn Sci. 2012; 16:106–113. [PubMed: 22245618]

Braver, TB.; Gray, JG.; Burgess, GB. Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In: Conway, ARA.; Jarrold, C.; Kane, MJ.; Miyake, A.; Towse, JN., editors. Variation in Working Memory. Oxford: Oxford University Press; 2007. p. 76-106.

Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, Gabrieli JDE. Dissociable contributions of prefrontal and parietal cortices to response selection. Neuroimage. 2002; 17:1562–1571. [PubMed: 12414294]

Burle B, Vidal F, Tandonnet C, Hasbroucq T. Physiological evidence for response inhibition in choice reaction time tasks. Brain Cogn. 2004; 56:153–164. [PubMed: 15518932]

Carter CS, van Veen V. Anterior cingulate cortex and conflict detection: An update of theory and data. Cogn Affect Behav Neurosci. 2007; 7:367–379. [PubMed: 18189010]

Caspers S, Zilles K, Laird AR, Eickhoff SB. ALE meta-analysis of action observation and imitation in the human brain. Neuroimage. 2010; 50:1148–1167. [PubMed: 20056149]

Catmur C, Heyes C. Time course analyses confirm independence of imitative and spatial compatibility. J Exp Psychol Hum Percept Perform. 2010

Catmur C, Walsh V, Heyes C. Associative sequence learning: The role of experience in the development of imitation and the mirror system. PHILOS T R SOC B. 2009; 364:2369–2380.

Chartrand TL, Bargh JA. The chameleon effect: The perception-behavior link and social interaction. J Pers Soc Psychol. 1999; 76:893–910. [PubMed: 10402679]

Clark S, Tremblay F, Ste-Marie D. Differential modulation of corticospinal excitability during observation, mental imagery and imitation of hand actions. Neuropsychologia. 2004; 42:105– 112. [PubMed: 14615080]

Cook JL, Bird G. Atypical social modulation of imitation in autism spectrum conditions. J Autism Dev Disord. 2012; 42:1045–1051. [PubMed: 21833823]

Crescentini C, Mengotti P, Grecucci A, Rumiati RI. The effect of observed biological and non biological movements on action imitation: An fmri study. Brain Res. 2011; 1420:80–92. [PubMed: 21959173]

Cross KA, Iacoboni M. Optimized neural coding? Control mechanisms in large cortical networks implemented by connectivity changes. Hum Brain Mapp. 2011

D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L. The motor somatotopy of speech perception. Curr Biol. 2009; 19:381–385. [PubMed: 19217297]

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. Understanding motor events: A neurophysiological study. Exp Brain Res. 1992; 91:176–180. [PubMed: 1301372]

Egner T. Multiple conflict-driven control mechanisms in the human brain. Trends Cogn Sci. 2008; 12:374–380. [PubMed: 18760657]

Egner T, Delano M, Hirsch J. Separate conflict-specific cognitive control mechanisms in the human brain. Neuroimage. 2007; 35:940–948. [PubMed: 17276088]

Egner T, Etkin A, Gale S, Hirsch J. Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. Cereb Cortex. 2008; 18:1475–1484. [PubMed: 17940084]

Egner T, Hirsch J. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci. 2005; 8:1784–1790. [PubMed: 16286928]

Etkin A, Egner T, Peraza DM, Kandel ER, Hirsch J. Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. Neuron. 2006; 51:871–882. [PubMed: 16982430]

Fadiga L, Fogassi L, Pavesi G, Rizzolatti G. Motor facilitation during action observation: A magnetic stimulation study. J Neurophysiol. 1995; 73:2608–2611. [PubMed: 7666169]

Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M. Modulating the experience of agency: A positron emission tomography study. Neuroimage. 2003; 18:324–333. [PubMed: 12595186]

Farrer C, Frith CD. Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. Neuroimage. 2002; 15:596–603. [PubMed: 11848702]

Ferrari PF, Bonini L, Fogassi L. From monkey mirror neurons to primate behaviours: Possible 'direct' and 'indirect' pathways. Philos Trans R Soc Lond B Biol Sci. 2009; 364:2311–2323. [PubMed: 19620103]

Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003; 19:1273–1302. [PubMed: 12948688]

Friston KJ, Rotshtein P, Geng JJ, Sterzer P, Henson RN. A critique of functional localisers. Neuroimage. 2006; 30:1077–1087. [PubMed: 16635579]

Gangitano M, Mottaghy FM, Pascual-Leone A. Phase-specific modulation of cortical motor output during movement observation. Neuroreport. 2001; 12:1489–1492. [PubMed: 11388435]

Gangitano M, Mottaghy FM, Pascual-Leone A. Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. Eur J Neurosci. 2004; 20:2193–2202. [PubMed: 15450099]

Gillmeister H, Catmur C, Liepelt R, Brass M, Heyes C. Experience-based priming of body parts: A study of action imitation. Brain Res. 2008; 1217:157–170. [PubMed: 18502404]

Hamilton AF. Emulation and mimicry for social interaction: A theoretical approach to imitation in autism. Q J Exp Psychol (Colchester). 2008; 61:101–115.

Heyes C. Automatic imitation. Psychol Bull. 2011; 137:463–483. [PubMed: 21280938]

Heyes C, Bird G, Johnson H, Haggard P. Experience modulates automatic imitation. Brain Res Cogn Brain Res. 2005; 22:233–240. [PubMed: 15653296]

Hommel B. The simon effect as tool and heuristic. Acta Psychol (Amst). 2011:189–202. [PubMed: 20507830]

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. Cortical mechanisms of human imitation. Science. 1999; 286:2526–2528. [PubMed: 10617472]

Jansson E, Wilson AD, Williams JH, Mon-Williams M. Methodological problems undermine tests of the ideo-motor conjecture. Exp Brain Res. 2007; 182:549–558. [PubMed: 17593359]

Jeannerod M. The 25th bartlett lecture. To act or not to act: Perspectives on the representation of actions. Q J Exp Psychol A. 1999; 52:1–29. [PubMed: 10101973]

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17:825–841. [PubMed: 12377157]

Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Med Image Anal. 2001; 5:143–156. [PubMed: 11516708]

Jiménez L, Recio S, Méndez A, Lorda MJ, Permuy B, Méndez C. Automatic imitation and spatial compatibility in a key-pressing task. Acta Psychol (Amst). 2012; 141:96–103. [PubMed: 22864312]

de Jong R. Strategical determinants of compatibility effects with task uncertainty. Acta Psychol (Amst). 1995; 88:187– 207.

Kiebel SJ, Klöppel S, Weiskopf N, Friston KJ. Dynamic causal modeling: A generative model of slice timing in fmri. Neuroimage. 2007; 34:1487–1496. [PubMed: 17161624]

Kilner JM, Paulignan Y, Blakemore SJ. An interference effect of observed biological movement on action. Curr Biol. 2003; 13:522–525. [PubMed: 12646137]

Kornblum S, Hasbroucq T, Osman A. Dimensional overlap: Cognitive basis for stimulus-response compatibility--a model and taxonomy. Psychol Rev. 1990; 97:253–270. [PubMed: 2186425]

Lakin JL, Jefferis VE, Cheng CM, Chartrand TL. The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. J NONVERBAL BEHAV. 2003; 27:145–162.

Levy BJ, Wagner AD. Cognitive control and right ventrolateral prefrontal cortex: Reflexive reorienting, motor inhibition, and action updating. Ann N Y Acad Sci. 2011; 1224:40–62. [PubMed: 21486295]

Lhermitte F, Pillon B, Serdaru M. Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: A neuropsychological study of 75 patients. Ann Neurol. 1986; 19:326–334. [PubMed: 3707084]

Longo MR, Kosobud A, Bertenthal BI. Automatic imitation of biomechanically possible and impossible actions: Effects of priming movements versus goals. J Exp Psychol Hum Percept Perform. 2008; 34:489–501. [PubMed: 18377184]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Ann Rev Neuroscience. 2001; 24:167.

Montagna M, Cerri G, Borroni P, Baldissera F. Excitability changes in human corticospinal projections to muscles moving hand and fingers while viewing a reaching and grasping action. Eur J Neurosci. 2005; 22:1513–1520. [PubMed: 16190904]

Nahab FB, Kundu P, Gallea C, Kakareka J, Pursley R, Pohida T, Miletta N, Friedman J, Hallett M. The neural processes underlying self-agency. Cereb Cortex. 2011; 21:48–55. [PubMed: 20378581]

Nee DE, Wager TD, Jonides J. Interference resolution: Insights from a meta-analysis of neuroimaging tasks. Cogn Affect Behav Neurosci. 2007; 7:1–17. [PubMed: 17598730]

Niedenthal PM, Barsalou LW, Winkielman P, Krauth-Gruber S, Ric F. Embodiment in attitudes, social perception, and emotion. Pers Soc Psychol Rev. 2005; 9:184–211. [PubMed: 16083360]

Obhi SS, Hogeveen J, Pascual-Leone A. Resonating with others: The effects of self-construal type on motor cortical output. J Neurosci. 2011; 31:14531–14535. [PubMed: 21994369]

Penny WD. Comparing dynamic causal models using AIC, BIC and free energy. Neuroimage. 2012; 59:319–330. [PubMed: 21864690]

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP. Comparing families of dynamic causal models. PLoS Comput Biol. 2010; 6:e1000709. [PubMed: 20300649]

Petrides M, Pandya DN. Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J Neurosci. 2007; 27:11573–11586. [PubMed: 17959800]

Press C, Bird G, Walsh E, Heyes C. Automatic imitation of intransitive actions. Brain Cogn. 2008; 67:44–50. [PubMed: 18077067]

De Renzi E, Cavalleri F, Facchini S. Imitation and utilisation behaviour. J Neurol Neurosurg Psychiatry. 1996; 61:396–400. [PubMed: 8890779]

Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn. 2004; 56:129–140. [PubMed: 15518930]

Rizzolatti G, Arbib MA. Language within our grasp. Trends Neurosci. 1998; 21:188–194. [PubMed: 9610880]

Rizzolatti G, Craighero L. The mirror-neuron system. Annu Rev Neurosci. 2004; 27:169–192. [PubMed: 15217330]

Ruby P, Decety J. Effect of subjective perspective taking during simulation of action: A PET investigation of agency. Nat Neurosci. 2001; 4:546–550. [PubMed: 11319565]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 2004; 23(Suppl 1):S208–S219. [PubMed: 15501092]

Spengler S, Bird G, Brass M. Hyperimitation of actions is related to reduced understanding of others' minds in autism spectrum conditions. Biol Psychiatry. 2010a; 68:1148–1155. [PubMed: 21130224]

Spengler S, Brass M, Kühn S, Schütz-Bosbach S. Minimizing motor mimicry by myself: Self-focus enhances online action-control mechanisms during motor contagion. Conscious Cogn. 2010b; 19:98–106. [PubMed: 20116291]

Spengler S, von Cramon DY, Brass M. Control of shared representations relies on key processes involved in mental state attribution. Hum Brain Mapp. 2009; 30:3704–3718. [PubMed: 19517530]

Spengler S, von Cramon DY, Brass M. Resisting motor mimicry: Control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. Soc Neurosci. 2010; 5:401–416. [PubMed: 20401807]

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. Neuroimage. 2009; 46:1004–1017. [PubMed: 19306932]

Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ. Ten simple rules for dynamic causal modeling. Neuroimage. 2010; 49:3099–3109. [PubMed: 19914382]

Stürmer B, Aschersleben G, Prinz W. Correspondence effects with manual gestures and postures: A study of imitation. J Exp Psychol Hum Percept Perform. 2000; 26:1746–1759. [PubMed: 11129371]

Stürmer B, Leuthold H, Soetens E, Schröter H, Sommer W. Control over location-based response activation in the simon task: Behavioral and electrophysiological evidence. J Exp Psychol Hum Percept Perform. 2002; 28:1345–1363. [PubMed: 12542132]

van Veen V, Cohen JD, Botvinick MM, Stenger VA, Carter CS. Anterior cingulate cortex, conflict monitoring, and levels of processing. Neuroimage. 2001; 14:1302–1308. [PubMed: 11707086]

Wager TD, Nichols TE. Optimization of experimental design in fmri: A general framework using a genetic algorithm. Neuroimage. 2003; 18:293–309. [PubMed: 12595184]

Wager TD, Sylvester CY, Lacey SC, Nee DE, Franklin M, Jonides J. Common and unique components of response inhibition revealed by fmri. Neuroimage. 2005; 27:323–340. [PubMed: 16019232]

Wang Y, Newport R, Hamilton AF. Eye contact enhances mimicry of intransitive hand movements. Biol Lett. 2011a; 7:7–10. [PubMed: 20427328]

Wang Y, Ramsey R, de C, Hamilton AF. The control of mimicry by eye contact is mediated by medial prefrontal cortex. J Neurosci. 2011b; 31:12001–12010. [PubMed: 21849560]

Wendelken C, Ditterich J, Bunge SA, Carter CS. Stimulus and response conflict processing during perceptual decision making. Cogn Affect Behav Neurosci. 2009; 9:434–447. [PubMed: 19897796]

Woolrich M. Robust group analysis using outlier inference. Neuroimage. 2008; 41:286–301. [PubMed: 18407525]

Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modelling for FMRI group analysis using bayesian inference. Neuroimage. 2004; 21:1732–1747. [PubMed: 15050594]

Yeterian EH, Pandya DN, Tomaiuolo F, Petrides M. The cortical connectivity of the prefrontal cortex in the monkey brain. Cortex. 2012; 48:58–81. [PubMed: 21481342]

HIGHLIGHTS

- Biological and non-biological stimuli invoke similar reaction time interference effects
- Neural correlates of these interference effects are dissociable, indicating a dedicated imitation control mechanism
- The frontal node of the mirror neuron system—the inferior frontal gyrus, pars opercularis—is modulated by imitative conflict
- Dynamic causal modeling suggests the medial prefrontal cortex detects imitative conflict.
- The anterior insula interacts with the mirror neuron system during imitative conflict resolution.

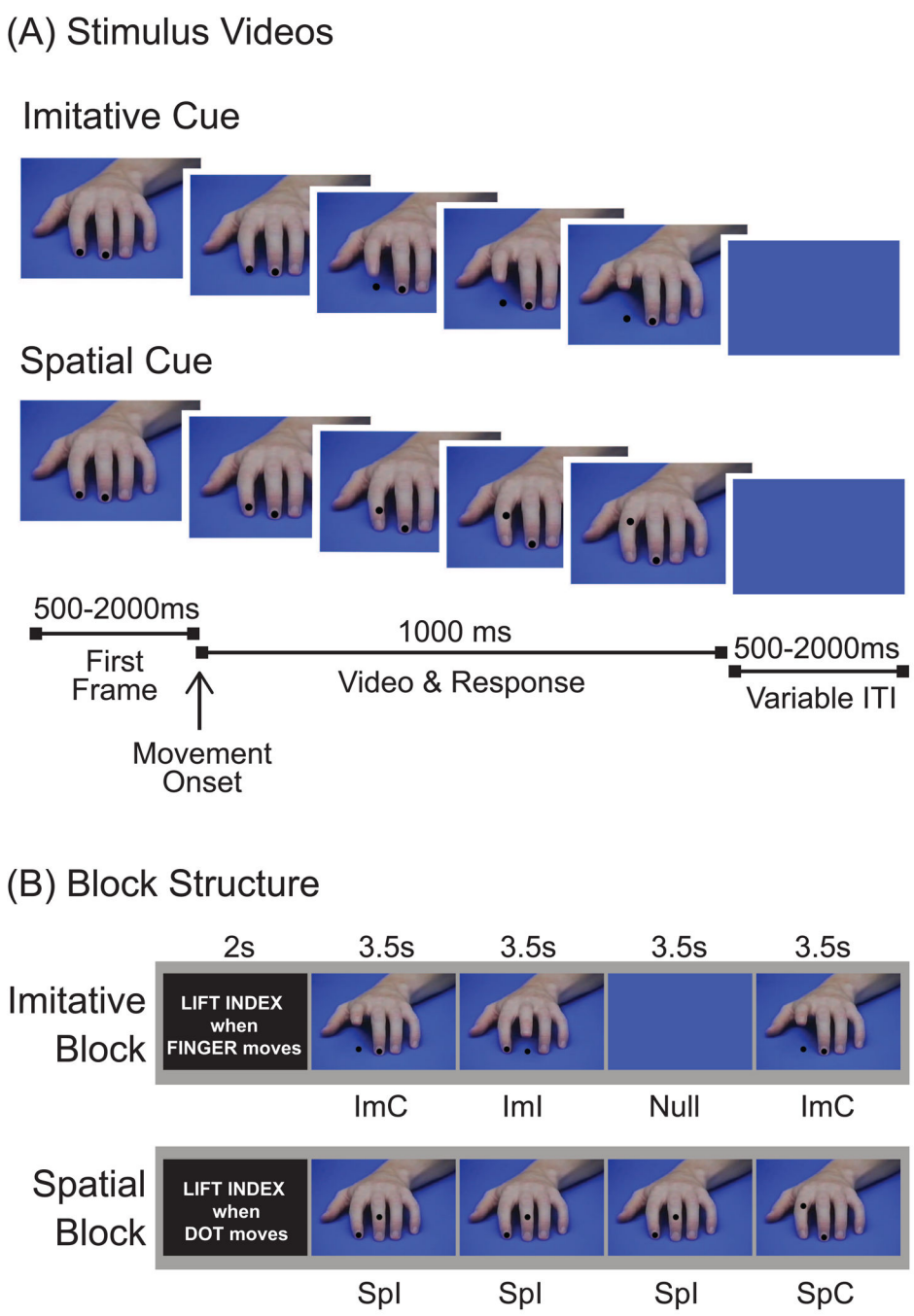

Figure 1.

Behavioral paradigm. (A) Example video stimuli and timing of one trial for imitative (top) and spatial (bottom) interference tasks. (B) Two example blocks are shown with time progressing from left to right and images depicting the last frame of the video for each trial. Conditions are listed under each frame (ImC = Imitate congruent; ImI = Imitate incongruent; SpC = Spatial congruent; SpI = Spatial incongruent). The congruency is defined with respect to the instructed action (lift index finger, in these examples).

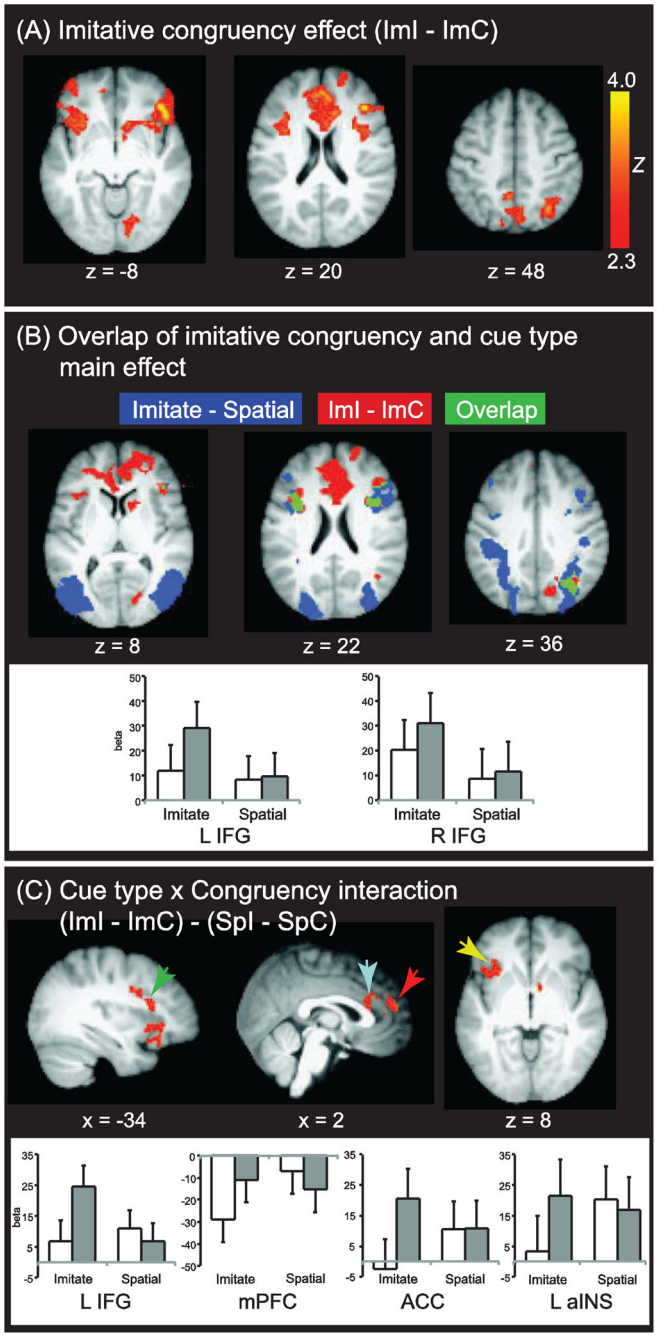

Figure 2.

Regional activation results. (A) Regions with greater activation for incongruent than congruent trials for imitative cues. No regions showed a significant congruency effect for spatial cues. (B) Overlap (green) of imitative congruency effect (red) and main effect of cue type (blue) demonstrate the IFGpo is both modulated by congruency and more active during action observation than observation of moving dots (C) Interaction effect showing regions where congruency effect is significantly greater for imitative than spatial cues. These regions represent the regions of interest in the DCM analysis (Green = IFGpo; Blue = ACC; Red = mPFC; Yellow = aINS). Bar graphs depict parameter estimates extracted from significant clusters, with error

bars representing standard error of the mean across subjects. All contrasts are thresholded at z>2.3 corrected across the whole brain for multiple comparisons ($p < 0.05$ FWE).

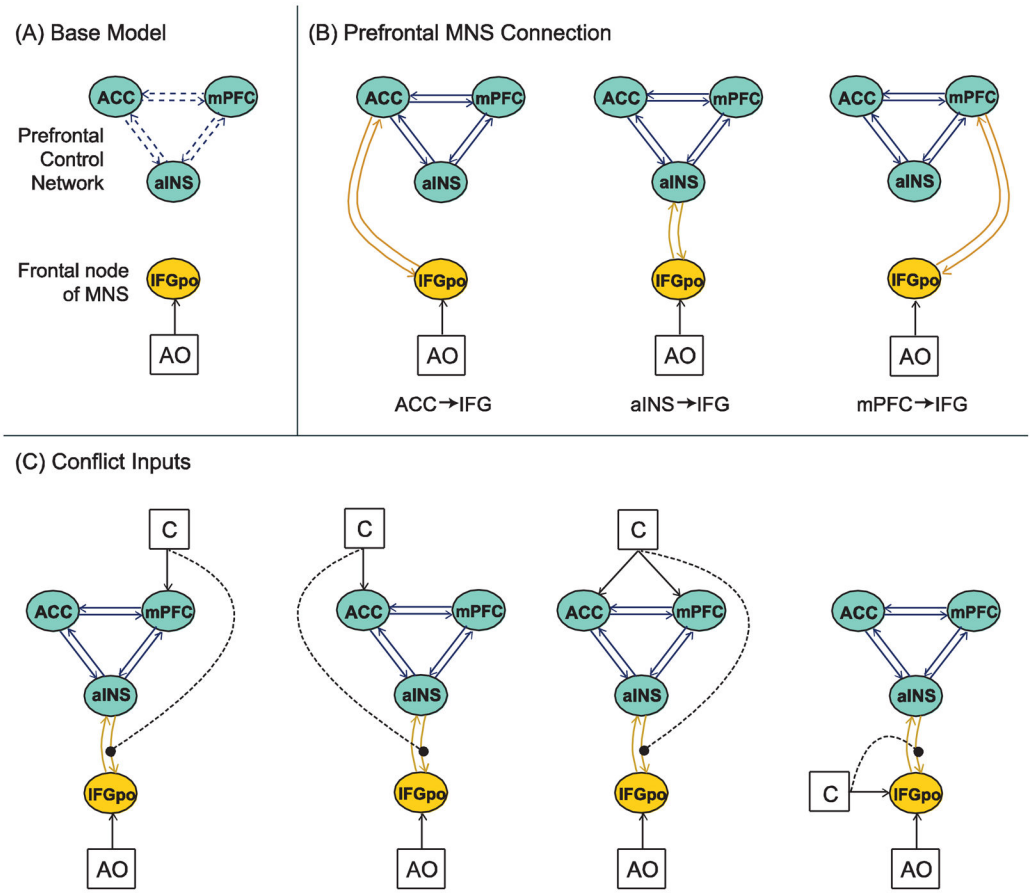

Figure 3.

Model space. (A). Schemata of parameters that made up the base models. The mirror neuron system is driven by action observation; and the three prefrontal nodes are connected by all combinations of 2 or 3 of the bidirectional connections, which are depicted by dotted lines (see Supplementary Figure 2A for expansion of 4 possible prefrontal connection models). (B). Three variations of prefrontal-MNS interactions were included. The prefrontal network was connected to the frontal node of the mirror neuron system (IFGpo) via one of the 3 prefrontal control nodes by varying the connectivity structure as shown. This allowed us to identify which prefrontal region interacts with the MNS. (C). Variations of conflict input to the system are depicted on one single connectivity structure (fully connected prefrontal network and aINS→IFGpo connection). Solid arrows show variations in the nodes receiving conflict as a driving input (from left to right: mPFC, ACC, ACC & mPFC, IFGpo). These variations test conflict detection hypotheses. Dotted lines depict conflict as a modulator of prefrontal input to the MNS. The same models excluding the modulating input were also included creating a total of 8 variations of conflict inputs. An expanded depiction of the model space showing the factorial combinations of the models depicted here can be found in Supplementary Figure 2. ACC = anterior cingulate cortex; mPFC = medial prefrontal cortex; aINS = anterior insula; IFGpo = inferior frontal gyrus, pars opercularis; AO = action observation; C = imitative conflict.

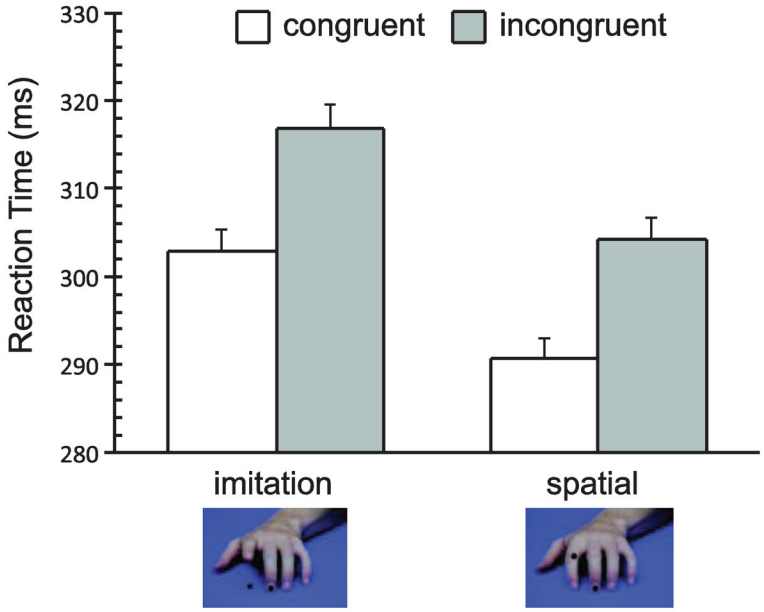

Figure 4.

Behavioral Results: Mean reaction time for each condition. Error bars represent within subject standard error of the mean, calculated with Cousineau's adaptation of Loftus & Masson's method {Cousineau 2005, Loftus 1994}. Main effects of congruency and cue type were significant ($p < 0.01$), but the interaction between cue type and congruency was not.

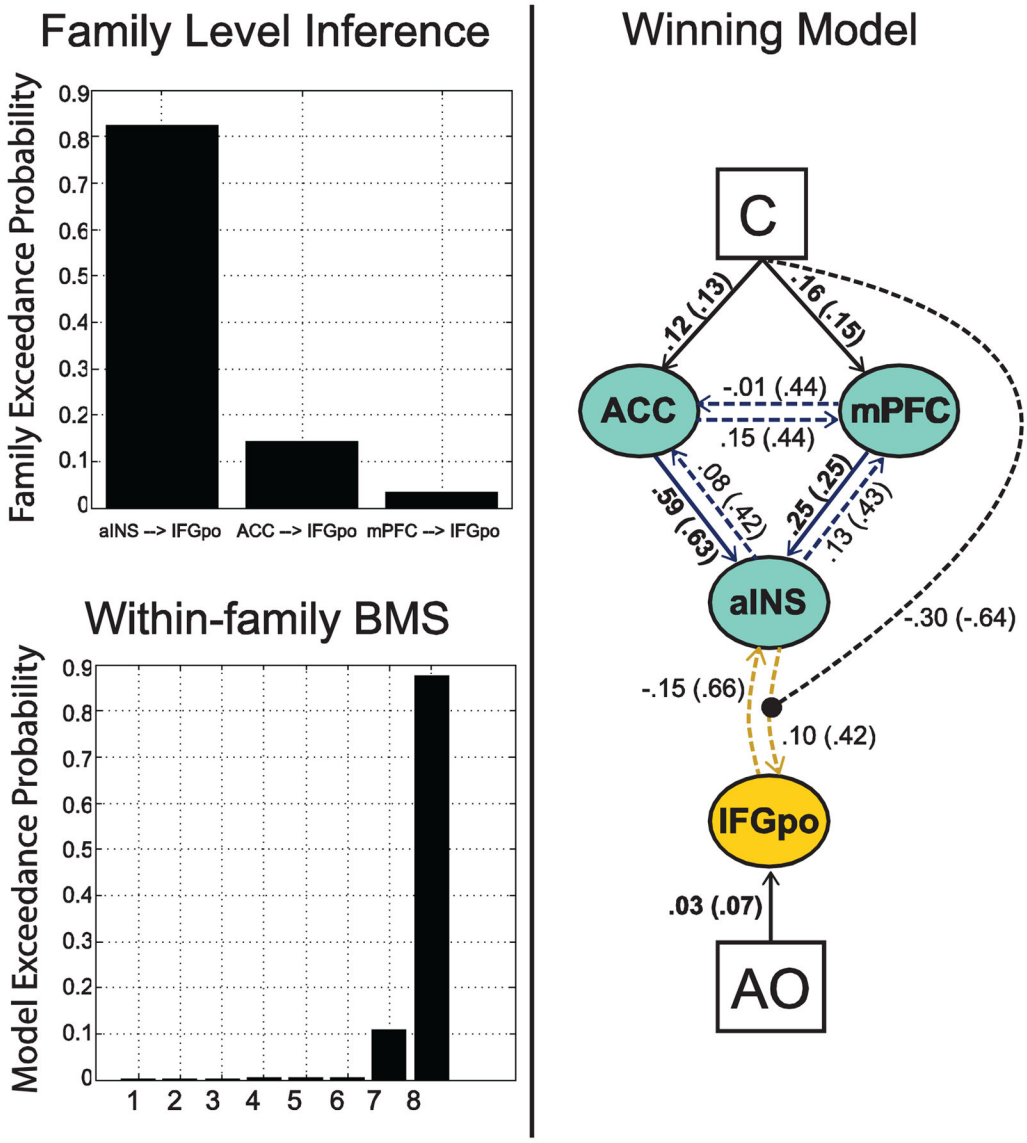

Figure 5.

DCM results. Family level inference performed on models within the fully-connected prefrontal family demonstrates exceedance probability of 0.82 for models including the aINS→IFGpo connection (top left). Model selection comparing models within this family shows only 2 models receiving any evidence (bottom left). The winning model (model 8) is shown at right. Values next to each connection or input show the mean and standard deviation (in parentheses) of the parameters across subjects. Parameters significantly different from 0 ($p<0.05$) are depicted with solid lines and bold parameter values. The modulation of aINS→IFGpo connection by conflict also approached significance ($p=0.07$).

Table 1

BMS results.

	Exp. Posterior Probability	Exceedance Probability
Prefrontal Family Inference		
Fully connected prefrontal nodes	0.48	0.88
mPFC-aINS & ACC-aINS connections	0.14	0.02
mPFC-ACC & ACC-aINS connection	0.24	0.08
mPFCACC & mPFC-aINS connections	0.14	0.02
Prefrontal → IFGpo Family Inference		
aINS → IFGpo	0.50	0.82
ACC → IFGpo	0.30	0.14
mPFC → IFGpo	0.20	0.03
Fully connected/aINS→IFGpo BMS (top 3)		
mPFC + ACC detection, with modulation	0.40	0.88
ACC detection, with modulation	0.22	0.11
mPFC detection, with modulation	0.09	0.005
IFGpo detection, with modulation	0.08	0.005
mPFC + ACC, no modulation (4 models)	0.08	0.004
ACC detection, no modulation (4 models)	0.04	< 0.001
mPFC detection, no modulation (4 models)	0.04	< 0.001
IFGpo detection, no modulation (4 models)	0.04	< 0.001