Whenever we observe a movement of a conspecific, our mirror neuron system becomes activated, urging us to imitate the observed movement. However, because such automatic imitation is not always appropriate, an inhibitive component keeping us from imitating everything we see seems crucial for an effective social behavior. This becomes evident from neuropsychological conditions like echopraxia, in which this suppression is absent. Here, we unraveled the neurodynamics underlying this proposed inhibition of automatic imitation by measuring and manipulating brain activity during the execution of a stimulus–response compatibility paradigm. Within the identified connectivity network, right middle/inferior frontal cortex sends neural input concerning general response inhibition to right premotor cortex, which is involved in automatic imitation. Subsequently, the fully prepared imitative response is sent to left opercular cortex that functions as a final gating mechanism for intentional imitation. We propose an informed neurocognitive model of inhibition of automatic imitation, suggesting a functional dissociation between automatic and intentional imitation.

Keywords: fMRI, imitation, mirror neurons, response inhibition, TMS

Introduction

Whenever we observe a movement of a conspecific, our mirror neuron system becomes activated, urging us to imitate the observed movement. Action observation leads to an activation in the premotor cortex normally involved in the execution of the same action, a process facilitated by mirror neurons; motor neurons that also respond to sensory information concerning biological motion (Fadiga et al. 1995; Rizzolatti and Arbib 1998; Iacoboni et al. 1999, 2001). This motor activation induces a tendency to imitate the observed action (Chaminade et al. 2002, 2005; Decety et al. 2002; Buccino et al. 2004), a neural mechanism considered to be involved in an array of processes ranging from simple action recognition and action understanding to high cognitive functions such as language, theory of mind, social behavior, or even disorders like autism (Rizzolatti and Arbib 1998; Blakemore and Decety 2001; Umiltà et al. 2001; Wicker et al. 2003; Keysers et al. 2004).

Numerous previous studies have aimed at revealing the exact neural mechanisms underlying our natural tendency for imitation (Fadiga et al. 1995; Gallese et al. 1996; Rizzolatti et al. 1996; Buccino et al. 2001; Grafton et al. 1997; Iacoboni et al. 1999, 2001; Heiser et al. 2003). However, the potential existence of an equally important and closely related neural mechanism enabling us to inhibit such automatic imitative response tendencies has been largely neglected. In many daily-life situations, however, such suppression is crucial for an effective social behavior, for example, when action imitation is not appropriate or even maladaptive.

Neuropsychological support for the existence of such an inhibitory mechanism of automatic imitation comes from a condition in which imitation is positively disturbed. Luria described a condition called echopraxia, in which patients suffering from frontal lobe damage repeated every observed movement, preferably in a mirror-image fashion (Luria 1966). Lhermitte reported that many frontal syndrome patients showed "obstinate imitation behavior", a severely increased urge to imitate every observed action, even if they were explicitly instructed not to do so (Lhermitte et al. 1986). Even though these patterns of behavior seem similar in nature because they both involve an increased urge to imitate observed movements, Lhermitte explicitly classifies them as different conditions. Whereas echopraxic imitation is immediate, involuntary, with the speed and abruptness of a reflex, and not relying on intellectual processes, obstinate imitation behavior is a voluntary act of higher cognitive nature. Lhermitte (Lhermitte et al. 1986) proposes that damage to the anterior inferior part of the frontal lobe results in a release of inhibition of activity from the parietal lobe, causing inappropriate imitation behavior to occur. Notwithstanding the possible differences, both conditions involve severely increased imitative behavior and both are being linked to frontal lobe damage. The implications following from these phenomena of increased imitation are 2-fold: first, in accordance with mirror neuron theory, there is an automatic tendency to imitate observed actions and second, under normal physiological circumstances, this tendency can be suppressed if needed.

Traditionally, processes of response inhibition have been investigated using overlearned responses, in paradigms such as the well-known Go–Nogo or Stroop tasks. In accordance with lesion studies, brain imaging studies reveal a network of several frontal and cingulate cortical regions involved in response inhibition. Using the Stroop task, areas in the left inferior frontal gyrus (Langenecker et al. 2004), right lateral prefrontal cortex (Vendrell et al. 1995), prefrontal and anterior cingulate cortex (ACC) (Harrison et al. 2005), superior and inferior parietal lobules, and inferior frontal gyrus and extrastriate cortex (Millham et al. 2002) were found to be of importance for the inhibition of the overlearned reading response. Using the Go–Nogo task, Dillon and Pizzagalli (2007) concluded that response inhibition depends upon fronto-basal ganglia networks and that right ventrolateral prefrontal cortex may support a general inhibitory process. Horn et al. (2003) found support for the hypothesis that the anterior lateral orbitofrontal cortex is activated during response inhibition. They also concluded that a network of higher order association and

The Brain's Intention to Imitate: The Neurobiology of Intentional versus Automatic Imitation

The Author 2009. Published by Oxford University Press. All rights reserved.
For permissions, please e-mail journals.permissions@oxfordjournals.org

Nina Bien, Alard Roebroeck, Rainer Goebel and Alexander T. Sack

Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, 6200 MD, Maastricht, the Netherlands
paralimbic areas is involved in response inhibition. In a functional magnetic resonance imaging (fMRI) and connectivity study, Koechlin et al. (2003) report that caudal lateral prefrontal regions are involved in selecting premotor representations associating stimulus and motor responses.

Most studies investigating response inhibition focus on a single paradigm. However, the downside of this approach is that it is difficult to disentangle activation caused specifically by response inhibition from complex, task-specific activations. In an attempt to specify brain regions involved in response inhibition in general, Wager et al. (2005) compared fMRI activation and connectivity patterns resulting from 3 different inhibition tasks: a Go–Nogo task, a flanker task, and a stimulus–response compatibility task. They found a set of frontal and parietal regions commonly engaged in response inhibition across the 3 tasks, including bilateral anterior insula/frontal operculum and anterior prefrontal, right dorsolateral and premotor, and parietal cortices. Even though correlations among tasks both for brain activity and performance were low, probably because of individual biases unique to each task, they suggested that certain common interference detection and/or resolution mechanisms are engaged across tasks.

Only few neuroimaging studies have attempted to specifically identify the neural correlates underlying the inhibition of imitative response tendencies in healthy human volunteers. Brass, Zysset, and von Cramon (2001) explored the cortical mechanisms underlying the inhibition of imitation in an event-related fMRI study, using a simplified stimulus–response paradigm, a task frequently used in studies mapping the human mirror neuron system (Brass et al. 2000). A lack of activation in regions usually detected in inhibition studies, including the inferior frontal gyrus and the ACC, was interpreted by the authors as indirect evidence that the inhibition of imitative responses employed different cortical structures than inhibitive mechanisms involved in Go–Nogo and Stroop tasks. A second fMRI study aimed at finding a functional dissociation between the inhibition of learned and imitative responses (Brass et al. 2005). The authors indeed showed this dissociation, supporting the hypothesis that the inhibitive mechanisms in both processes consist of different components. In addition, this study revealed a region in the right inferior frontal gyrus that seemed to be involved in both inhibitive processes. The authors reasoned that this region is involved in generating a stop signal right before the response is executed. Thus, even though the cortical mechanisms preceding this decision are different, the eventual inhibition of the movement could be executed by a mutual region.

The results of these studies are insufficient to explain the process of the inhibition of imitation. Partly, this can be ascribed to the experimental design of the aforementioned studies. Brass, Zysset, and von Cramon (2001) compared cortical activity in response to execution of a finger action while participants were observing either the same (congruent) or a different (incongruent) action on a screen. The congruent condition was supposed to reflect imitation; the incongruent condition should reveal inhibition of an imitative response. However, in this design, actions were not actively imitated, or inhibited, for that matter. Rather, stimulus and response merely either coincided or not, which is not the same as intentionally imitating an observed action. In the second study (Brass et al. 2005), besides the previously described task, a Stroop task was used to detect functional dissociations between the inhibition of imitative and overlearned responses. However, because these tasks were utterly different, the resulting cortical activations could not be compared directly. Rather, congruent and incongruent conditions were contrasted for both tasks, and the produced cortical areas were merely compared qualitatively.

Besides design shortcomings, there are methodological issues. Using exclusively fMRI to investigate processes of response inhibition certainly has its limitations. The blood oxygen level–dependent (BOLD) signal allowing for visualization of local brain activity is likely to represent intracortical processing rather than neuronal spiking rate (Logothetis et al. 2001; Logothetis and Wandell 2004). An area influencing inhibition might thus appear in the functional image in an identical manner as an excited area, namely, as an area of increased synaptic activity. Indeed, it seems that complex inhibition can be reflected in a larger, smaller, or no BOLD signal at all (Ritter and Villringer 2002). Evidence supporting the notion of negative BOLD deflections reflecting decreased neuronal activity (Shmuel et al. 2006) and/or cortical inhibition (Hummel et al. 2002, 2004; Kastrup et al. 2008) has also been presented recently. However, the exact significance of negative BOLD responses is not yet known, and as of yet, there is no proof that this mechanism plays a role in the inhibition of automatic imitation.

Although investigating response inhibition using functional imaging appears to be a suspect to ambiguous signal interpretations, sensible hypothesizing and, consequently, construction of a clear cognitive model prior to experimental recording can lead to simple, testable predictions. In the current study, as in most studies concerning response inhibition, we considered inhibition an active process, resulting in an increase in BOLD signal. In addition, methods investigating connectivity and causality in the brain can be of great use when investigating processes of inhibition. Data-driven effective connectivity analysis is capable of revealing targets and sources within fMRI data, without having to rely on a priori specification of a model that contains preselected regions and connections between them. It provides temporal information about the direction of information flow through an activated network.

Using effective connectivity analysis, conclusions can be drawn with regard to the probable directed and task-related neural influence one cortical region exerts over another one (Friston et al. 1995), making it particularly suited for questions regarding cortical inhibition. However, it does not provide evidence with regard to causal relationships between brain activation and behavioral output. Brain areas showing up in effective connectivity maps are—at least—correlated to task execution, but whether they are functionally relevant for the actual performance cannot be assessed on the basis of these data alone.

A method particularly useful for investigating causal relationships between activity modulations in specific brain areas and actual behavior, and therefore potentially also for clarifying issues on cortical inhibition, is transcranial magnetic stimulation (TMS) (see, e.g., Hallett 2000; Pascual-Leone et al. 2000). By intervening with the functionality of a certain brain region and at the same time observing possible changes in behavioral responses, assumptions about the causal relevance of a certain brain area for the execution of a certain task can be experimentally tested.
In the current study, the existence, network connectivity, and functional relevance of the specific brain system enabling us to inhibit responses, particularly automatic imitative response tendencies, were investigated. Our explicit goal was to identify, functionally dissociate, and manipulate the neural correlates of this proposed inhibition of automatic imitation by using an elaborate combination of complementary noninvasive brain imaging and brain interference techniques.

A stimulus–response compatibility design was employed for assessing imitation and inhibition behavior. During brain imaging, participants had to imitate an observed finger movement while ignoring a simultaneously presented spatial cue (imitative trials) or move their finger as it was indicated by the spatial cue while ignoring finger movements (spatial trials). The imitative and the spatial cue could appear either on the same finger (congruent condition) or on different fingers (incongruent condition). We hypothesized that incongruent trials would mainly reflect response inhibition and that the condition in which a movement of the “wrong” finger would have to be suppressed (spatial incongruent) would reflect the inhibition of imitation, specifically.

Random effects (RFX) statistical activation maps were produced to allow mapping of the brain regions involved in response inhibition and specifically in the inhibition of imitative responses. In order to reveal the direction of information flow within the revealed network, RFX effective connectivity analyses were conducted using the concept of Granger causality (Granger 1980), producing cortical maps depicting effective connectivity, consisting of source and target activations. Source activations represent regions whose activation consistently predicts future activation (i.e., next time point) of the seed region and thus are hypothesized to have exerted task-related causal influence on the cortical activation of the seed region. Target activations on the other hand represent regions whose activation is consistently predicted by the activation of the seed region (i.e., previous time point) and thus are hypothesized to be causally influenced by the seed region. Finally, fMRI-guided TMS was applied to identify the underlying functional dissociation by revealing the specific functional contribution of each involved brain area for successful imitation and inhibition of imitation.

The converging evidence enabled us to propose a new neurobiological model of the inhibition of imitative response tendencies, distinguishing between automatic and intentional imitation.

Materials and Methods

Participants

A total of 15 healthy volunteers with normal or corrected to normal vision (mean age: 23 ± 1.5), 5 of which were male, participated in the fMRI part of this study. Eight of these participants (mean age: 22.3 ± 2.6), 3 of which were male, subsequently underwent TMS. All participants were unaware of the goal of the study until after having completed their participation. Ethical approval was given by the local medical ethical committee.

Experimental Design

Each stimulus consisted of a picture of a hand, followed by an imitative cue (a downward movement of either index or middle finger) and a spatial cue (a small white cross appearing underneath the index or middle finger), emerging at the same time. Subjects rested their index and middle fingers on their respective buttons, waiting for the appearance of the cues to indicate a response with either the index or the middle finger. Depending on the instruction, subjects were required to imitate the observed finger movement while ignoring the spatial cue (imitative trials) or move their finger as it was indicated by a spatial cue while ignoring finger movements (spatial trials). Please note that both responses involved active downward motion of the index or middle finger to press a button, whereas during baseline, subjects' fingers were resting on the response buttons without active extension of the fingers. The imitative and the spatial cue could appear either on the same finger (congruent condition) or on different fingers (incongruent condition) (see Fig. 6 or Stimulus Material for more details). A 2 by 2 design was constructed, containing the independent variables “congruency” and “target.”

fMRI Design

A fast event-related design was chosen for the fMRI study. A blocked presentation was necessary with regard to the target condition because the imitative and spatial trials required different task instructions for the subject. A psychophysical pilot study, however, revealed that a blocked design would be too predictable to disclose congruency effects. Consequently, a mixed blocked and event-related fMRI design was applied.

A total of 256 trials, 64 per condition, were assigned to 2 functional acquisition runs containing 8 blocks each. Each block contained 16 trials. The 8 blocks, presented in semirandomized order, consisted of 4 imitative and 4 spatial blocks, all containing both congruent and incongruent trials in randomized order.

Each block started with a 22.5-s period of fixation, visual instruction, and again fixation. Stimulus onset asynchrony, which was varied in a semirandom manner, varied between 4500 and 9000 ms (3, 4, 5, or time repetition [TR]'s). This added up to a block duration of almost 2 min and 50 s. The duration of a run was 22 min and 40 s, resulting in an experimental duration of 45 min.

TMS Design

Event-related TMS was applied with a figure-of-eight TMS coil, using 10-Hz triplets, starting at 200 ms after the onset of the imitative and spatial cue. Subjects were stimulated at 120% of their resting motor threshold. The experiment consisted of 3 sessions, during each of which a single target site was stimulated. Each session contained 8 blocks, during 4 of which participants received magnetic stimulation. The resulting 4 nonstimulation baseline blocks were alternated with stimulation blocks. An imitative TMS block was followed by an imitative non-TMS block or vice versa. Each block contained 12 trials with a duration between 6000 and 7000 ms.

Three target sites were selected based on fMRI group data. Two of these regions derived from functional brain imaging data. The third was detected as a result of effective connectivity analysis. For each participant, these target sites were detected within their individual brain imaging data set, using 5 selection criteria. First, the candidate region had to be directly accessible with TMS (Sack and Linden 2003). Second, using the same contrast that yielded this particular target region in the group data, the candidate target region had to show considerable activation in each single-subject analysis. Third, it had to display a BOLD response signal consistent with the conducted contrast analysis (deconvolution time courses), fourth, the BOLD signal characteristics had to be comparable to those resulting from group data analysis, and finally, stimulating the target region should not directly interfere with right hand motor execution.

Participants were coregistered to their structural brain imaging data using the BrainVoyager TMS Neuronavigator (Brain Innovation BV, Maastricht, the Netherlands), a system that allows for continuous online coil navigation directly above a cortical target point localized using fMRI activation (Fig. 7). TMS target points were defined on the cortical surface using individual functional brain imaging data. The figure-of-eight TMS coil was mounted in an adjustable arm and positioned directly over the intended target stimulation site, tangentially to the participant’s skull in both the horizontal and the vertical dimension. During the course of the stimulation sessions, the proximity of the coil’s “hotspot” to the target point was continuously monitored and, if necessary, the position was adjusted online. Subjects were positioned...
in a headrest and chin rest. Coil orientation was kept as constant as possible between subjects. For the right premotor and middle frontal gyrus target points, the coil was oriented in a 45 degree angle respective to the longitudinal fissure, with the coil handle pointing in a posterior-superior direction. For the left opercular region, preservation of intersubject homogeneity was more difficult because not only the headrest interfered with the positioning of the coil but also some subjects experienced muscle reactions to certain orientations of the coil above this region. In almost all the participants, however, the coil was positioned in a 90 degree angle respective to the longitudinal fissure, with the handle pointing in the medial-superior direction.

**Stimulus Material**

During the complete study, Presentation software (Neurobehavioral Systems, Inc., Albany, NY) was used for both stimulus presentation and recording of the behavioral responses. Each experimental stimulus started with a stationary picture of a hand, recorded from a viewing point such that it looked as the left hand of another individual sitting across from the subject (Fig. 6). Like in previous studies using a similar design, a stimulus mirror imaged to the participants' own right hand was used because this is the setting in which imitation occurs mostly during daily social engagements and in which imitation is easiest to perform (Luria 1966; Brass, Bekkering, and Prinz 2001). The index and middle finger of the hand were lifted in order to correspond to the position of the subjects’ fingers as they were held above their respective buttons. This was important to ensure that the observed movement would indeed correspond to the executed movement, thus ensuring stimulus-response compatibility.

After a randomly varied period of 1250, 1500, 1750, 2000, or 2250 ms, the imitative and spatial cues appeared (Fig. 6). A picture of the same hand replaced the initial picture, except now showing either the index or the middle finger in a downward position while the rest of the hand and background remained the same. At exactly the same time, the spatial cue, consisting of a small white cross, appeared. In congruent trials, this cross appeared on the same finger that had made the downward movement. In incongruent conditions, the cross appeared on the finger that remained stationary.

As stated earlier, a picture of the final posture of the “moving” finger followed the picture of the static hand, thus creating a powerful illusion of motion. Movie recordings of the same movement could have resulted in slight differences in the starting position, as well as the onsets of the trajectory of the different finger movements. Also, it would have created an imbalance between the imitative and the spatial cue: whereas the appearance of the movement would have started gradually and become more noticeable every few milliseconds, the cross would have appeared from one moment to the other. Timing those events such that they would be equivalent would have been impossible. Hence, the use of static pictures replacing each other was preferred over the use of a movie, to ensure maximum controllability with regard to cue onset and duration. Participants consistently reported that they did not notice seeing 2 consecutive pictures, instead of a movie of a natural movement.

Both the imitative and the spatial cue remained visible for 1000 ms. Trial duration was fixed at 3500 ms. In between trials, subjects were shown a white fixation cross on a monochrome black background.

**Apparatus and Data Acquisition**

**fMRI Acquisition**

Brain imaging data were gathered using a 3-T Siemens Allegra magnetic resonance (MR) scanner and a volume head coil at the Faculty of Psychology of Maastricht University, the Netherlands.

Functional images were acquired using single-shot gradient-recalled echoplanar imaging (710 volumes, TR = 1500 ms, time echo [TE] = 30 ms, flip angle = 90°). Twenty-four oblique slices were acquired without interslice gap. Voxel size was 3.5 × 3.5 × 4 mm³, and each slice contained a matrix of 64 by 64 voxels.

High-resolution (voxel size 1 × 1 × 1 mm³) 3D anatomical data were collected using a T₁-weighted modified driven equilibrium Fourier transform (MDEFT) sequence. This sequence consisted of 176 sagital slices of 1-mm thick and a matrix of 256 by 256 voxels (TR = 7.92 ms, TE = 2.4 ms, flip angle = 15°).

Stimulus material was displayed onto a frosted screen, positioned at the rear end of the scanner bore, using an LCD projector (PLC-XT11-16, Sanyo North America Corporation, San Diego, CA). An adjustable mirror mounted on the head coil allowed each subject a complete view of the display. Responses were recorded using a standard MR compatible button box (LUMItouch keypads, Photon Control, Burnaby, Canada).

**TMS Apparatus**

Biphasic magnetic stimulation was generated using a Medtronic MagPro X100 stimulator (Medtronic Functional Diagnostics A/S, Skovlunde, Denmark). Magnetic pulses were delivered with a figure-of-eight coil (Magnetic Coil Transducer MC-B70, Medtronic), mounted in an adjustable arm.

**Data Analysis**

**Brain Imaging Data Analysis**

Functional and anatomical brain imaging data were preprocessed and analyzed offline using BrainVoyager QX (Brain Innovation BV).

The first 5 volumes of each run were excluded to allow for T₁ saturation, permitting the T₂*-weighted MR signal to stabilize. The first functional volume served as a high-contrast prototype to which following functional volumes were aligned during preprocessing.

Preprocessing of the functional data existed of slice scan time correction using sinc interpolation, 3D motion correction involving trilinear interpolation, linear trend removal, and application of 3 cycles per time course (2.83 × 10⁻³ Hz) high-pass filter. Individual functional runs were inspected visually to ascertain that subjects had not moved excessively during data recording, which was not the case. Anatomical data were standardized using Talairach transformation (Talairach andTournoix 1988). Subsequently, the 2D functional runs were transformed into Talairached 3D volume time courses. Volume time courses were spatially smoothed using a 6-mm full-width half-maximum kernel.

The statistical analysis of variance (ANOVA) of the BOLD signal was based on the application of multiple regression analysis to time series of task-related functional activation (Friston et al. 1995). Predictor time courses were adjusted for the hemodynamic response delay by convolution with a hemodynamic response function (Boynton et al. 1996) (delta 3 and tau 2). Contrasts between conditions were carried out using BrainVoyager’s RFX general linear method (GLM) analysis tool, enabling generalization of the statistical inferences to the population level. Fixation periods at the start and end of each block were used as a baseline.

To obtain mirror neuron-related activation maps, the 2 experimental imitative conditions were contrasted with the 2 experimental spatial conditions because mirror neuron hypothesis would predict that the mirror neuron network is especially activated during instances of imitation.

Processes of response inhibition mutual to both imitative and nonimitative responses were considered congruency main effects and were visualized by subtracting congruent trials (imitative and spatial) from incongruent trials (imitative and spatial). Particularly important was a contrast that would identify regions specifically involved in the inhibition of imitative responses. The rationale was that observing a finger movement while the task demanded that an adjacent finger was moved in a similar way, as was the case in the spatial incongruent condition, would lead to an urge to imitate, which would have to be actively inhibited to complete the task. Subtraction of the imitative incongruent BOLD responses from the spatial incongruent ones would reveal areas involved in suppressing the urge to imitate the moving finger. To exclude areas related to response inhibition in general, in addition to a higher response to incongruent spatial trials, the response to imitative incongruent trials should be as low as, or lower than, the response to imitative congruent trials, resulting in the following contrast: (spatial incongruent - spatial congruent) – (imitative incongruent + imitative congruent). This contrast would reveal activation in response to suppressing the urge to imitate, corrected for general incongruency effects.
For visualization purposes, statistical activation maps were projected onto inflated and flattened template cortices and smoothed once with a repeat value of 5. Cluster size threshold was set at 50 voxels. Results from the RFX analysis were reported using \( P \) values. In addition to statistical contrasts, event-related deconvolutions were used to visualize relative activations between conditions within a certain cortical region of interest, as produced by statistical analysis.

**Functional and Effective Brain Connectivity Analysis**

Functional connectivity analysis was carried out offline on all 15 subjects, using RFX Granger causality mapping (GCM; for details, see Roebroeck et al. 2005), producing cortical maps depicting source and target activations. RFX analysis enables generalization of the statistical inferences to the population level.

Three seed regions were selected, based on their task-dependent BOLD characteristics as observed in the group GLM analysis and one based on one of the produced connectivity maps. Several criteria for the selection of seed regions were employed. First, they appeared as significant activation clusters in RFX group analysis (GLM or an already executed GCM). In addition to showing a reliable group effect, all these regions showed similar characteristics in each individual subject. The latter prerequisite has to be met in order to justify a RFX GCM, which demands individually localized seed regions. Each selected seed region was chosen such as to consist of 200-300 voxels; large enough to consider it a representative selection of the activated voxels but small enough to prevent heterogeneity from concealing the results.

Due to the mixed, randomized presentation of congruent and incongruent trials—a decision based on pilot studies that indicated that a blocked, separated presentation of congruent and incongruent trials would be too predictable—connectivity differences between these conditions could not be investigated using GCM. Thus, collapsing congruent and incongruent trials, only connectivity patterns from imitative trials could be compared.

**Behavioral and TMS Data Analysis**

Reaction times were recorded during the brain imaging session and during the TMS experiment. Data preprocessing consisted of removing false responses and inspecting the data for skewness and outliers.

Statistical analysis of the reaction times recorded during the MRI session consisted of a repeated measures ANOVA containing the factors target (imitative or spatial) and congruency (congruent and incongruent). Alpha values of pairwise comparisons were Bonferroni corrected to anticipate sphericity issues. In that subjects responded considerably slower in the incongruent condition (mean RT = 466.2 ms, SD = 19.4 ms) compared with the congruent condition (mean RT = 409.5 ms, SD = 15.1 ms). No significant interactions were found.

**Brain Imaging Results**

Regions supposedly reflecting response inhibition were revealed by showing the incongruency main effect, obtained by contrasting incongruent with congruent trials (Fig. 1a). This contrast revealed activations in the premotor cortex and middle frontal gyrus (dorsolateral prefrontal cortex), especially in the right hemisphere, intraparietal sulcus and inferior parietal gyrus bilaterally, right ACC, right pre-supplementary motor area (pre-SMA), right superior temporal sulcus (STS), and the anterior insula/frontal operculum, again slightly lateralized to the right. Although there is evidence that cortical inhibition correlates with negative BOLD changes, we did not find such patterns in the current data. A few negative clusters emerged from the analysis; however, closer inspection of their event-related deconvolutions showed that they were most likely the results of artifacts, not functional deactivations.

Brain regions showing an increased BOLD response to imitative relative to spatial trials consisted of the posterior part of the right inferior frontal gyrus and the occipitotemporal cortex (Fig. 1b). The activation was bilateral, although much more pronounced in the right hemisphere. No brain areas were found that showed a stronger response to spatial than to imitative cues.

The contrast hypothesized to reveal regions specifically involved in the inhibition of automatic imitation, namely, spatial incongruent versus all other 3 main experimental conditions, showed an extensive network, comprising premotor cortex, posterior parietal and frontal/parietal opercular cortex, all bilateral, and right STS (Fig. 1c). However, only few of these regions actually showed a BOLD response that was specifically increased during spatial incongruent trials. Only the left opercular region showed a clear BOLD response depicting a stronger activation in the spatial incongruent condition compared with the other 3 experimental conditions.

**Functional and Effective Connectivity Results**

In order to reveal the direction of information flow within the revealed network, effective connectivity analyses were conducted using GCM (Roebroeck et al. 2005), producing cortical maps depicting functional and effective connectivity, consisting of source and target activations. Source activations represent regions whose activation consistently preceded that of the seed region and thus are hypothesized to have exerted task-related causal influence on the cortical activation of the seed region. Target activations on the other hand represent regions that are consistently activated after the seed region and thus are hypothesized to be causally influenced by the seed region.

As starting points for effective connectivity, 3 seed regions were used, whose selection was based on the previously described brain imaging results. The first region of interest was located in right premotor cortex. It displayed a consistent incongruency effect in the overall GLM as well as in all 15 individual subjects and also showed corresponding event-related deconvolutions in each subject. The second region of interest used as a seed for effective connectivity analysis was...
produced by the connectivity analysis of right premotor cortex as an area sending neuronal input and was located in the right middle frontal gyrus. The third seed region was located in the left opercular cortex, displayed a preference for spatial incongruent trials, and was thus hypothesized to be crucial to the inhibition of imitative responses.

Effective connectivity analyses revealed that right premotor cortex receives neural input from predominantly right middle frontal gyrus and frontal opercular/insular regions and in turn projects neural input to bilateral posterior parietal, frontoparietal opercular, and occipitotemporal cortical target regions (Fig. 2b). These effective connectivity networks were similar in...
Figure 2. Effective connectivity maps. Effective connectivity maps of seed regions in (a) right middle frontal gyrus, (b) right premotor cortex, and (c) left opercular cortex, show the flow of neuronal information during task execution. Seed regions are shown in red, green colors display regions serving as an activation source of the seed region, and blue colors represent target regions, receiving neural input from the seed region. Maps obtained from imitative blocks are shown in the top row of each panel, and maps obtained from spatial conditions are in the bottom row of each panel. Right middle frontal gyrus (response inhibition) projects to right premotor cortex (imitation), which in turn projects to posterior parietal cortex (which possibly sends feedback) and finally to left opercular cortex (inhibition of imitation).
imitative and spatial blocks and thus in accordance with the hypothesis, based on this regions’ BOLD responses, that this network is involved in a general incongruency effect—probably reflecting response inhibition—in both imitative and spatial blocks. As a whole, and considering the heightened BOLD responses of the right premotor region during incongruent trials, this suggests that the network is involved in response inhibition and displays a flow of information from middle frontal and insular regions, via premotor cortex, to bilateral posterior parietal, frontoparietal opercular, and occipitotemporal cortices.

The same analysis was carried out over the right middle frontal gyrus, which was consistently present as a source of information for the activation in right premotor cortex. This confirmed the flow of neuronal information from right middle frontal to right premotor cortex (Fig. 2a). Furthermore, posterior parietal, opercular, and occipitotemporal cortices were shown to be targets of right middle frontal gyrus. Because GCM cannot distinguish between influence exerted directly or via an intermediate cortical area, these target activations probably reflect the aforementioned target activations of right premotor cortex, which thus functions as a mediator. Again, few differences were observed between imitative and spatial conditions. The presence of many cortical areas receiving information from middle frontal cortex, whereas almost none are sending information, indicates a relatively early time point of relevance within the activated network.

Effective connectivity analysis over the left opercular seed region, according to its BOLD pattern specifically involved in suppressing the urge to imitate, once more showed no large difference between imitative and spatial blocks, although the source activation in insular/opercular regions was slightly more pronounced in the spatial condition (Fig. 2c). In this case, compared with the right premotor and moreover the right middle frontal connectivity maps, much more source activations were observed, implying that this area was activated in a later stage than the right premotor and middle frontal areas. Among the source regions were areas along the precentral gyrus and sulcus, anterior insular/fronteral opercular regions, and occipitotemporal cortex, bilaterally. It is important to note that the aforementioned right premotor seed region, earlier seen to project information to left opercular cortex, was also included in these areas, confirming that its activation served as a source of information for the left opercular region. Taken together, these results suggest that, assuming that the left opercular region is indeed involved in the inhibition of imitation as its BOLD pattern suggests, it mainly receives its information from insular/opercular, occipitotemporal, and frontal regions, the latter including the middle frontal area, which is sending input to the right premotor area.

**TMS Results**

The 3 selected target regions consisted of the right premotor area that displayed an incongruency effect and was also used as a seed region for effective connectivity analysis, a right middle frontal region that was revealed by effective connectivity analysis as a source region of the aforementioned TMS target region, and the left opercular region that displayed a strong activation as a result of spatial incongruent trials (Fig. 3). Although the activation of premotor cortex was largely bilateral, right premotor cortex was chosen to prevent direct influence of TMS on motor performance during task execution. Furthermore, the selection criteria ruled out several interesting candidate regions in the parietal lobe, due to a very high degree of intersubject variability.

An overall view of the results showed that when neural activity in any of these regions was disrupted by TMS, participants on average responded significantly faster as compared with non-TMS blocks ($F = 135.9$, $df = 1$, $P < 0.001$) (Fig. 4). This effect was seen following stimulation of any of the 3 target regions and was thus nontask or area specific. However, an overall GLM consisting of the factors region (3 levels) and condition (4 levels) yielded a significant interaction ($F = 2.4$, $df = 6$, $P = 0.027$), showing that the effect of condition was significantly different for the 3 different TMS target regions.

Disturbing the right middle frontal gyrus with TMS resulted in a significant main effect of condition ($F = 4.75$, $df = 3$, $P = 0.003$) and, when broken down, in significant main effects of both target ($F = 4.2$, $df = 1$, $P = 0.041$) and congruency ($F = 7.4$, $df = 1$, $P = 0.007$). Post hoc comparisons showed that participants responded significantly faster to spatial incongruent trials, when compared with imitative ($P = 0.001$, mean difference $= 23.4$ ms) and spatial congruent ($P = 0.003$, mean difference $= 20.4$ ms) trials, but not when compared with imitative incongruent trials ($P = 0.147$; Fig. 4a). Thus, after stimulation of the right middle frontal cortex, participants responded faster to trials that required inhibition of automatic imitation, compared with when their right middle frontal cortex was not disturbed by TMS.

TMS of the right premotor cortex target region resulted in a main effect for the factor condition ($F = 3.2$, $df = 3$, $P = 0.022$). Broken down into separate factors, the factors target ($F = 3.5$, $df = 1$, $P = 0.062$) and congruency ($F = 3.7$, $df = 1$, $P = 0.056$) showed trends toward main effects. Post hoc pairwise comparisons between the 4 different conditions showed that subjects responded significantly faster in the spatial incongruent condition than to imitative congruent ($P = 0.005$, mean difference $= 21.4$ ms), imitative incongruent ($P = 0.014$, mean difference $= 18.6$ ms), and spatial congruent ($P = 0.017$, mean difference $= 18.0$ ms) trials (Fig. 4b). Hence, much like TMS of right middle frontal gyrus, TMS of right premotor cortex made participants respond faster when they had to inhibit automatic imitation, relative to when their right premotor cortex was not disrupted.

Administering TMS to the left opercular target region again resulted in a main effect of condition ($F = 5.0$, $df = 3$, $P = 0.002$). Broken down, only the factor target revealed a significant main effect ($F = 11.5$, $df = 1$, $P = 0.001$). Post hoc pairwise comparisons showed that compared with spatial congruent trials, participants responded significantly faster in the imitative congruent ($P = 0.014$, mean difference $= 25.0$ ms) and imitative incongruent ($P = 0.002$, mean difference $= 29.7$ ms) conditions (Fig. 4c). Thus, when left opercular cortex was disturbed, participants responded relatively faster in conditions requiring intentional imitation.

**Discussion**

Our goal was to reveal the existence, layout, connectivity, and functional relevance of the brain network enabling us to inhibit automatic imitation. In order to do this and based on previous brain imaging studies successfully employing the same design (Iacoboni et al. 1999), we used the stimulus–response paradigm.
to map the human mirror neuron system and compare different forms of cortical inhibition, including the one sought after inhibition of imitation. Using different complementary methods, we were able to not only observe but also experimentally manipulate the brain activity associated with response inhibition. This resulted in a new neurobiological model explaining the processes and mechanisms underlying the inhibition of imitative responses.

Figure 3. TMS target regions. The 3 target regions for the TMS experiment, visualized on an inflated Montreal Neurological Institute hemisphere (top panel) and as clusters in a 3D reconstruction of the head, color coded per subject (bottom panel). (a) Right middle frontal gyrus was produced consistently as a source of the cortical activation in right premotor cortex (shown in red) in effective connectivity maps. (b) Right premotor cortex showed an increased activation during incongruent relative to congruent stimuli, as shown by its event-related deconvolution. (c) A region in the left operculum shows a preference for spatial incongruent trials, the condition presumed to reflect inhibition of imitation, as shown in orange in the event-related deconvolution.
Conditions reflecting a general stimulus–response incongruency effect showed a slightly right-lateralized cortical network consisting mainly of premotor cortex, operculum/insula, posterior and intraparietal cortex, and middle/inferior frontal gyrus. This latter cortical area was predominantly involved in the processing of imitative stimuli. Effective connectivity analysis showed a flow of information through this network from middle frontal, to premotor, to parietal cortex. Possibly, feedback was sent from parietal cortex back to the premotor cortex. The left operculum, according to its BOLD pattern specifically involved in the inhibition of imitative responses, received information from right premotor cortex.

To unravel the contribution of the cortical areas involved in this network, we targeted 3 regions with TMS, enabling us to observe the causal relevance of these regions for the behaviorally controlled inhibition of imitative responses. TMS induced a general and consistent speeding of response times. As a consequence, it is difficult to say whether a shorter reaction time in a certain condition is indeed a sign of improved performance or whether it reflects deteriorated performance in the other conditions. However, the issue with regard to whether a change in performance was in absolute sense positive or negative can never be ruled out and does not interfere with drawing conclusions about the functional relevance of, in this case, brain areas targeted with TMS. Interestingly, and aside from this neither task- nor area-specific behavioral improvement effect, there were also task- and area-specific effects, as could already be concluded from the significant interaction between TMS target region and task condition.

Disturbing right premotor cortex resulted in relatively faster responses in the only condition during which an automatic imitative response would not be beneficial, even counterproductive: the spatial incongruent condition. We conclude from this that premotor cortex is involved in the process of automatic imitation and that disrupting premotor activity with TMS inhibited this tendency of automatic imitation, thereby resolving the incongruency effect resulting from competition between the spatial and the imitative stimulus. In other words, when the urge to imitate the wrong finger movement is reduced by TMS-induced premotor activity disruption, moving the “correct” finger as indicated by the spatial target stimulus becomes easier, hence faster.

Disturbing right middle frontal gyrus during task execution resulted in a pattern similar to that caused by disturbing the right premotor cortex, a significant speeding up of especially the responses to spatial incongruent trials. This behaviorally supports the results of the effective connectivity analysis, which suggests that right premotor and middle frontal cortices are, with regard to response inhibition, nodes of the same functional network. However, we did observe a difference between TMS of the right premotor versus middle frontal areas: in case of the latter, the difference between spatial and imitative incongruent conditions is not significant. This implies that, whereas premotor cortex serves to process imitative stimuli as discussed earlier, the target area in middle/inferior frontal cortex is possibly concerned with response inhibition in general. This conclusion seems in accordance with the findings of Brass et al. (2005), who reported a similar cluster showing mutual activation in 2 inhibition tasks, which they interpreted as an area generating a final stop signal concluding different instances of response inhibition. However, our combined evidence suggests a functional role for the inferior/middle frontal cortex fairly early in the process, rather than at the very end. The postulation is on the other hand challenged by the GLM results, which show involvement of premotor cortex in response inhibition and of middle frontal gyrus in the processing of imitative stimuli. Although interesting, this debate currently remains speculative because it is not backed up by statistical significance.

Stimulation of the left opercular region, which showed an increased BOLD in response to the condition during which the inhibition of imitation was of utmost importance, resulted in speeded response times in the imitative conditions, especially compared with the spatial congruent condition. It is important to note that during the spatial and imitative conditions, the visual display is entirely the same and that during the congruent conditions even the correct response is identical because both cues are always presented on the same finger. In this light, the observed difference between the imitative conditions and the spatial congruent condition is remarkable. The only difference between these conditions is that in the
former, the observed actions are intentionally imitated, whereas in the latter case, they are unintentionally imitated. In addition, effective connectivity analyses showed that this area within the left opercular region receives information from middle frontal cortex (which according to our findings is involved in response inhibition) through right premotor cortex (involved in imitation), which implies a seemingly late time point of relevance in the cortical network. We therefore suggest that the left frontoparietal operculum is involved in intentional imitation and serves as a final gating mechanism deciding on whether to execute the imitative action prepared by premotor cortex, or not. Thus, whereas automatic imitation is the unconsciously initiated imitative response of mirror neurons to the sensory perception of biological motion, intentional imitation occurs when a person consciously decides to imitate an observed movement. The mirror neuron system is most likely involved in both, and there is probably considerable continuity between automatic and intentional imitation. They share the first stages of the process, and in the case of intentional imitation, inhibition is released at the final stage of the procedure.

The concept of an inhibitive component serving as a final gating mechanism on an already prepared motor response may come across as counterintuitive and inefficient. However, one must keep in mind that this prepared imitative motor action is a result of the activation of the premotor areas in response to action observation, hence in essence an epiphenomenon of action understanding and deriving higher order cognitive processes associated with the mirror neuron system. Furthermore, a similar concept of a gating mechanism inhibiting an already prepared motor response was also brought up by Brass et al. (2005).

Subtracting spatial from imitative trials was expected to reveal areas usually associated with the mirror neuron system, especially Broca’s area (left inferior frontal gyrus and frontal operculum), left premotor cortex, and parietal cortex. In this light, the lack of activation in Broca’s area was unusual, contradicting among others the evidence described by Heiser et al. (2003), who concluded that this brain region is crucial for imitation. It has been shown that even static images conveying motion can induce cortical activation usually related to action observation (Nishitani and Hari 2002). In the current experiment, this would be the case, even increasingly so as time during the experiments elapsed and participants got more and more acquainted with the stimulus and its relation to motion. Preliminary activation would prevent mirror neuron-related areas from appearing in activation maps.

Thus, with regard to the intended mapping of the mirror neuron system, we conclude that the employed stimulus-response compatibility paradigm did not yield the typical set of cortical areas usually associated with mirror neuron activity. Neither did we find the reaction time advantage in response to imitative stimuli, again contrarily to the expectancy based on previous studies using the same paradigm. Both findings are, however, in accordance with several other studies not finding mirror neuron-related cortical activity as a result of action observation (Makuuchi 2005; Jackson et al. 2006), as well as a recent study by Jonas et al. (2007), putting forward the notion that, depending on the experimental context, the stimulus-response compatibility task does not always reliably activate the mirror neuron system. Following this reasoning, we propose that in the experimental context of the brain imaging part of our study, the stimulus-response compatibility task might have encouraged participants to, to a certain extent, follow a spatial matching strategy besides employing their mirror neuron system. In conclusion, our setup was not ideal for mapping the mirror neuron system. This was however not our primary goal and does not necessarily interfere with the gist of this study.

In this study, several methods were employed, each with their own strengths and weaknesses. Combining these methods in a way that they become complementary to each other resulted in a deepened understanding of observed brain activations, which could not have been obtained using any of these methods in isolation. In conclusion, our evidence converges to a new neurocognitive model for the mechanisms and cortical components associated with the inhibition of automatic imitative responses (Fig. 5). Information concerning general incongruency and response inhibition flows from right middle and inferior frontal cortices to a right premotor region involved in imitative responses. Subsequently, information is sent to the posterior parietal and intraparietal cortices and possibly sent back in the form of feedback to the premotor cortex. Finally, information is sent from right premotor cortex to left frontoparietal operculum, which subsequently acts upon this information as a final gating mechanism deciding on
intentional imitation, reflecting the inhibition of an already prepared imitative response.

Considering the previously described neuropsychological reports by Luria and Lhermitte concerning patients exhibiting increased imitation behavior after having sustained frontal lobe damage, it seems unexpected to conclude that the inhibition of imitation would be settled in this somewhat differently located cortical area. On the other hand, the participation of the middle and inferior frontal lobe in response inhibition in general has since long been established (Vendrell et al. 1995; Milham et al. 2002; Horn et al. 2003; Langenecker et al. 2004; Harrison et al. 2005; Langenecker et al. 2004; Wager et al. 2005; Dillon and Pizzagalli 2007) and is supported once more by the current results. With regard to the specific inhibition of imitative responses, the quantity of neuropsychological literature is rather limited and inconclusive with regard to the exact neurobiological mechanisms underlying this process.

We propose that, besides the premotor cortex which is involved in automatic imitation and the middle frontal cortex which subserves general response inhibition, the described frontoparietal area is part of the cortical network enabling healthy human beings to imitate in an intentional manner and to inhibit their already planned imitative responses when needed. This functional neural dissociation between automatic and intentional behavior, as revealed by our study, is also supported by the neuropsychological distinction, as proposed by Lhermitte (Lhermitte et al. 1986), between the involuntary, reflex-like form of imitation associated with echopraxia, and the voluntary, complex imitation behavior associated with obstinate imitation behavior. According to our neurobiological model, patients suffering from echopraxia or imitation behavior may have sustained defects at one or several levels of this
network, resulting in increased imitation, decreased inhibition of imitation, or a combination of both.

Funding
Dutch Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek [NWO], 452-06-003) to ATS.

Notes
We thank Peter de Weerd for his helpful comments on the manuscript, Teresa Schuhmann, our medical supervisor Cees van Leeuwen, and our independent physician Martin van Boxtel. Conflict of Interest: None declared.

Address correspondence to Dr Nina Bien, Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, PO Box 616, 6200 MD, Maastricht, the Netherlands. Email: nina.bien@psychology.unimaas.nl.

References


