Neural mechanisms, temporal dynamics, and individual differences in interference control

Forstmann, B.U.; van den Wildenberg, W.P.M.; Ridderinkhof, K.R.

Published in:
Journal of Cognitive Neuroscience

DOI:
10.1162/jocn.2008.20122

Citation for published version (APA):
Neural Mechanisms, Temporal Dynamics, and Individual Differences in Interference Control

Birte U. Forstmann, Wery P. M. van den Wildenberg, and K. Richard Ridderinkhof

Abstract

Functional magnetic resonance imaging (fMRI) methods may help in understanding processes of response capture and response inhibition in conflict tasks, such as the Simon task. However, data-driven approaches thus far have not yielded consistent insights into these processes. Here, a theory-driven approach is introduced that capitalizes on individual differences in the processes of central interest. Based on the so-called activation–suppression model, specific behavioral parameters for each individual derived from reaction time (RT) distribution analysis were computed and entered into model-based fMRI analyses. These parameters correspond closely to the processes of inappropriate location-driven response activation (capture) and the subsequent inhibition of this activation as detailed by the model. Data from 24 participants revealed activation in the pre-supplementary motor area, which covaried with the RT distribution measure of response capture. Activation in the right inferior frontal cortex was found to covary with the RT distribution measure of response inhibition. These results, which are consistent against the backdrop of the larger literature on cognitive control, could have been derived neither from the standard data-driven fMRI approach, nor from inspecting overall mean RT alone.

INTRODUCTION

The aim of the present investigation is to bring to light the neural bases and dynamics of inhibitory control processes using novel combinations of behavioral and neuroimaging data analysis techniques. In particular, this study combines reaction time (RT) distribution-analytical techniques with model-based functional magnetic resonance imaging (fMRI) to provide an integrated account of the neural mechanisms underlying individual differences in selective response inhibition. The term response inhibition is used here descriptively to refer to the mechanism or set of processes that result in the containment of prepotent behavioral responses when such responses are reflex-like, premature, inappropriate, or incorrect. Thus, response inhibition is a key process of cognitive control that is necessary for adequate adaptive behavior to meet the new goals in an ever-changing environment.

Inhibitory control has been investigated most widely in the context of the go/no-go paradigm. In this paradigm, participants respond to “go” stimuli but are required to withhold their response to “no-go” stimuli. fMRI studies have shown that inhibition on no-go trials can be associated with a predominantly right hemispheric network of brain areas, including bilateral superior, inferior, and dorsolateral prefrontal cortices, the supplementary motor area (SMA), the anterior cingulate, inferior parietal and temporal cortices, the caudate nucleus, and the cerebellum (e.g., Simmonds, Pekar, & Mostofsky, 2008; Durston, Thomas, Worden, Yang, & Casey, 2002; de Zubicaray, McMahon, Wilson, & Muthiah, 2001; Liddle, Kiehl, & Smith, 2001; Menon, Adleman, White, Glover, & Reiss, 2001; Kiehl, Liddle, & Hopfinger, 2000; Garavan, Ross, & Stein, 1999). Although go/no-go fMRI studies have provided valuable insights into the network of brain areas activated when a response has to be withheld, no-go signals also differ from go signals in terms of processes related to stimulus recognition, attention, and response selection. Hence, some of the observed differences in brain activity related to no-go and go stimuli may be related to processes other than response inhibition.

As in the no-go task, participants in the stop-signal paradigm are instructed to withhold a go response upon the presentation of an exogenous signal (i.e., the stop signal; Logan & Cowan, 1984). There is growing evidence that the neural network involved in stopping a response primarily relies on frontal brain structures (e.g., Aron, Behrens, Smith, Frank, & Poldrack, 2007; Aron & Poldrack, 2006; Aron, Robbins, & Poldrack, 2004) in tandem with the basal ganglia (Aron & Poldrack, 2006; van den Wildenberg et al., 2006; Casey, Tottenham, & Fossella, 2002).

Other paradigms that draw upon inhibitory control to resolve conflict situations in which multiple responses are simultaneously active are the Stroop task, the Eriksen flanker task, and stimulus–response (SR) compatibility tasks.
(e.g., Schumacher, Cole, & D’Esposito, 2007; Hazeltine, Bunge, Scanlon, & Gabrieli, 2003; Pardo, Pardo, Janer, & Raichle, 1990). In such tasks, responses are typically slowed on incongruent (IG) trials, that is, when the target aspect of the stimulus requires a response opposite to the response triggered by task-irrelevant stimulus aspects, rather than when the activated responses overlap, as on congruent (CG) trials. The magnitude of the behavioral interference effect of irrelevant stimulus position (i.e., $RT_{\text{IG}} - RT_{\text{CG}}$) is often considered as a measure for the ability to resolve response interference, which may involve selective response inhibition to reduce interference between competing actions (Simon, 1967). In general, recent meta-analyses studies indicate the activation of a fronto-parietal network, involving the lateral prefrontal cortex, the medial prefrontal cortex, and parietal regions in tasks that require the resolution of interference (Nee, Wager, & Jonides, 2007; Derrfuss, Brass, Neumann, & von Cramon, 2005; Neumann, Lohmann, Derrfuss, & von Cramon, 2005). More specifically, processes such as monitoring of response conflict and response selection are believed to play a crucial role in task performance and have been associated with functioning of the medial frontal cortex (MFC; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Botvinick, Braver, Barch, Carter, & Cohen, 2001).

However, there are also differences in subprocesses required in these conflict tasks that might account for differences in brain activation. In the Stroop task, besides the monitoring of response conflict and selection between competing responses, in addition to selective response inhibition, the relevant contrast in this task may capture other processes as well (e.g., conflict at the semantic level). A recent quantitative meta-analysis by Nee et al. (2007) that included fMRI data from several Stroop tasks revealed a strong left-lateralized network including the left middle frontal gyrus (MFG), the insula, as well as the left posterior parietal cortex. Contrary to Stroop stimuli, the target and flanker stimuli in the Eriksen flanker paradigm are presented within the same stimulus dimension without requiring semantic processes. With respect to Flanker task performance, meta-analyses show mainly right-lateralized activation clusters in the MFG and the insula (Nee et al., 2007). Interestingly, these patterns resemble activations found with SR compatibility tasks as well as go/no-go manipulations, whereas in the Stroop and Eriksen tasks, activation is also elicited in the SMA and the premotor cortex.

The heterogeneity of these results might be explained by different subprocesses or by the inseparability of subprocesses that subserve performance on these conflict tasks. Moreover, in a study by Hazeltine et al. (2005), evidence is provided for specific areas within the fronto-parietal network to code for material-dependent and material-independent selection processes (see also Schumacher, Elston, & D’Esposito, 2003). The results indicate that prefrontal areas such as the right inferior frontal cortex (IFC) reveal competition-related activation for color stimuli using an Eriksen flanker task, whereas regions within the MFG of both hemispheres exhibited more competition-related activation for letter stimuli. Contrary to this finding, conflict-monitoring areas such as the MFC and parietal regions exhibit stimulus-independent activations.

Taken together, results from neuroimaging studies investigating response inhibition and selection provide compelling evidence for the involvement of specific prefrontal and parietal regions. However, the exact nature and contribution of these areas in the process of selective response inhibition remains controversial. Stop and go/no-go studies have highlighted the neural circuits involved in global, but not selective, inhibition. Selective inhibition may be studied in the context of conflict tasks. Studies with the Eriksen flanker task appear less susceptible to the process-impurity objections that apply (in this regard) to the Stroop task. However, because task-relevant and task-irrelevant stimuli are presented within one and the same stimulus dimension in this paradigm, it is difficult to distinguish whether a particular response was activated by target or flankers and, consequently, which one should be selectively suppressed. Selective inhibition of inappropriate responses is likely to be more potent and direct, and less confounded with semantic or other processes, in a paradigm that, within the family of conflict tasks, stands somewhere in between the Stroop and Eriksen flanker tasks: the Simon task (Simon, 1967). Thus, in order to focus more specifically on the subprocesses of selective inhibitory control, we use the Simon task, in which the designated response is indicated by one aspect of the imperative signal (e.g., its color), but competing response tendencies may be elicited by the task-irrelevant spatial location of the stimulus, even though the latter is to be ignored.

To date, a data-driven fMRI approach in combination with the Simon task did not yield consistent patterns of brain activation with respect to selective inhibition. To better capture the temporal dynamics of response processing in conflict situations, we adopt recent RT distribution-analytical techniques to examine the interplay between response activation and response inhibition. In addition, we adopt a novel approach in neuroimaging that we will refer to as model-driven fMRI analysis to identify brain mechanisms associated with individual differences in interference control.

**Model-based fMRI**

Model-based fMRI capitalizes on individual differences in parameters that quantify the efficiency of specific cognitive processes. Such parameters can be estimated for each individual on the basis of formal (mathematical) models, or on the basis of theoretical (descriptive) models. As an example of the former, Haruno and Kawato...
(2006) employed the Q-learning model of reinforcement learning (Sutton & Barto, 1998) to quantify the effect of outcome processing and reward anticipation processes. More specifically, Haruno and Kawato (2006) used the Q-learning algorithm to estimate outcome and reward anticipation processes on a trial-by-trial basis for each participant. These estimates were convolved with a hemodynamic response function to model the blood oxygenation level-dependent (BOLD) signal representing outcome and reward anticipation processes, respectively (Haruno & Kawato, 2006, p. 951). The results revealed that individual differences in these estimates correlated with BOLD activation in the putamen and caudate nucleus, respectively. As an example of theoretical model-based fMRI, Aron et al. (2007) and Aron and Poldrack (2006) employed the empirical and analytical paradigm of the stop task to quantify the latency of global response inhibition processes (stop-signal reaction time, SSRT). This was done by correlating the individual brain activation derived from regions of interest with the individual SSRT. The results revealed a densely interconnected neural network consisting of the right IFC, the right pre-supplementary motor area (pre-SMA), and the subthalamic nucleus to covary with individual differences in inhibiting a response as measured with the SSRT.

**RT Distribution Analysis of Interference Effects**

Here, we will employ the dual-process model of interference effects in the Simon task and use the associated RT-distributional analyses to quantify the effects of response capture and selective response inhibition to examine which brain areas covary in their activation with individual differences in the respective parameter values. In examining the processes underlying interference effects in conflict tasks such as the Simon task, many authors have reported evidence in support of dual-route architecture (e.g., Ridderinkhof, 2002; Eimer, 1995; Kornblum, Hasbroucq, & Osman, 1990). A schematic representation of this type of model is depicted in Figure 1. Most significant, the controlled process of SR translation is paralleled by a direct response capture route. The two routes converge at the level of response activation processes.

In accordance with this general dual-route architecture, a recurrent and critical observation is that fast responses to IG stimuli tend to be associated with below-chance accuracy (e.g., Gratton, Coles, & Donchin, 1992). Conditional accuracy functions (CAFs), which plot accuracy as a function of response speed, reveal that very fast responses tend to be driven by the activation of the response system by task-irrelevant stimulus position. This direct response capture yields near-asymptotic accuracy when the response designated by the color is CG to stimulus position, but low accuracy with IG responses, as expressed in the fast portion of CAFs (e.g., Ridderinkhof, 2002). Individuals differ in the sensitivity to the capture of their initial action processing by task-irrelevant stimulus features, hence, in the extent to which stimulus position elicits activation of the associated response in the Simon task. The slope of the fastest portion of the CAF for IG responses may thus serve as a parametric index sensitive to individual differences in direct response capture (cf. Ridderinkhof, Scheres, Oosterlaan, & Sergeant, 2005).

Functional imaging studies provide evidence for the posterior MFC (pMFC) to be tightly linked to action selection and response conflict processes. In particular, the goal-directed selection of actions from among competing alternatives has been associated with activation of pMFC areas such as the pre-SMA (e.g., Aron et al., 2007; Nachev, Rees, Parton, Kennard, & Husain, 2005) and the dorsal anterior cingulate cortex (for a review see, e.g., Rushworth, Behrens, Rudebeck, & Walton, 2007). In the face of direct location-based response capture, in IG conditions, these areas have to work harder to select the appropriate response based on stimulus color. Processes related to the simultaneous activation of multiple competing responses have been associated with activation of the rostral cingulate zone (RCZ; for a review, see, e.g., Ridderinkhof et al., 2004), with more conflicting activation eliciting stronger RCZ activation. Hence, we hypothesized that individual differences in our index of response capture, the slope of the fastest portion of the CAF for IG responses, should covary with elevated activations in these pMFC areas.

**Figure 1.** Dual-process model. Architecture of the dual-process model. The increasing size of the arrows from the suppression module schematically represents the operation dynamics of this module (i.e., suppression is not operational immediately after signal onset, but takes some time to build up). S = stimulus; R = response.
To overcome direct response capture, many authors have assumed implicitly or explicitly that the response as activated by irrelevant stimulus features is subsequently inhibited (e.g., Eimer, 1999) and aborted (e.g., Logan & Cowan, 1984). Such inhibition is not operationally instantaneous but requires some time to develop (e.g., Eimer, 1999). Because of these dynamics (the gradual build-up of response inhibition as time progresses across a trial), slower responses will be more affected by selective response inhibition than faster responses (see also Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; Eimer, 1999). Although fast responses are dominated by direct location-driven response capture, an additional factor comes into play for slower responses: With slower responses, the selective inhibition process has had time to develop, and thus, the activation of the incorrect response along the direct capture route will be reduced. Consequently, correct slow responses to CG stimuli will be less facilitated by the position-driven response capture, whereas correct slow responses to IG stimuli will be less delayed. Thus, given these dynamics, interference effects are affected by selective response inhibition more in slow than in fast responses (for a review, see Ridderinkhof et al., 2005). These dynamics are well-captured by so-called delta plots, constructed by plotting the interference effect as a function of response speed (Wiegand & Wascher, 2007; Ridderinkhof, 2002; De Jong, Liang, & Lauber, 1994). Although delta plots prototypically have a positive slope (i.e., the effects of any experimental factor increase as a function of response speed), the notion that selective inhibition results in a reduction of the interference effect in slow responses (outlined above) implies a different delta-plot pattern: Here, the interference effect does not increase linearly as a function of response speed, but instead levels off and reduces for slow responses. If more effective selective inhibition results in a more pronounced reduction of interference effects in slow responses, as argued above, then the leveling off of the delta plot should be more pronounced in individuals that are more proficient in response inhibition than in less proficient individuals. Thus, the slope of the slowest portion of the delta plot may serve as a parametric index sensitive to individual differences in selective response inhibition. Using this index, such individual differences have been examined in children across different age groups (Bub, Masson, & Lalonde, 2006), children with attention deficit/hyperactivity disorder (Ridderinkhof et al., 2005), and individuals with mild cognitive impairment (Wylie, Ridderinkhof, Eckerle, & Manning, 2007).

Functional imaging as well as lesion and patient studies provide evidence for the right IFC, and more specifically, the right pars opercularis (BA 44), to play a crucial role in implementing top–down controlled inhibition (Aron et al., 2007; Aron & Poldrack, 2006; Casey et al., 2002; Garavan, Ross, Murphy, Roche, & Stein, 2002). Hence, we hypothesized that individual differences in our index of selective response inhibition, the slope of the slowest portion of the delta plot, should covary with elevated activations in this prefrontal brain area.

**The Present Study**

The aim of the present study is to combine RT distribution-analysis techniques with model-based fMRI to provide an integrated account of the neural mechanisms underlying individual differences of response capture and selective response inhibition. Therefore, we adopted a novel theoretical model-driven fMRI approach to covary indices of response capture and selective response inhibition with the BOLD signal derived from differences between IG and neutral trials (see Figure 2). Note that neutral trials were used to account for nonspecific speed effects, such as higher distraction or difficulty, on the BOLD signal (i.e., the BOLD signal might saturate more for slow than fast responses for reasons not pertinent for present purposes). This was important because we were interested in the covariation of the temporal dynamics derived from RT distributions and the BOLD signal. The IG versus neutral contrast values were entered in the regression analysis separately with the behavioral indices for response capture and selective response inhibition as covariates, respectively. Two hypotheses were formulated: First, we expected to find stronger response capture indices as derived from the fastest segment of the CAFs to yield elevated activation in the pMFC. Second, we expected stronger selective response inhibition indices as derived from the slowest segment of the delta plots to reveal elevated activation in the right IFC. Finally, we validate the specificity of these hypotheses by entering other, less specific control variables as a covariate in the regression analysis. For instance, the individual overall interference effect (RTIG – RTEG) lacks the specificity to expose the response capture and inhibition processes as expressed in the distribution indices. Hence, entering the individual overall interference effect as a covariate likely lacks the power and the specificity to reveal the specific brain areas associated with these subprocesses.

**METHODS**

**Participants**

Thirty healthy volunteers were recruited. We obtained written consent from all participants prior to the scanning session and the experiment was approved by the ethics committee of the University of Amsterdam. All subjects had normal or corrected-to-normal vision. No subject had a history of neurological, major medical, or psychiatric disorder. The data of five subjects were excluded from the analyses due to movement artifacts and one subject was excluded due to a cyst in the cerebellum. The remaining 24 subjects were 15 women...
and 9 men (age: $M = 24.2$ years, $SD = 2.76$) who were all right-handed as assessed by the Edinburgh Inventory (Oldfield, 1971).

**Behavioral Task**

A modified Simon task (Simon, 1967) with blue and green circles as stimuli was used. The blue or green circle was either associated with a left or right response button press with the index finger of the left or right hand, respectively. Participants practiced the relevant mapping in a short training session of approximately 10 min before entering the magnetic resonance imaging (MRI) scanner. Three different blocks were scanned, including one neutral block at the beginning and one at the end of the experiment, and an experimental block in between (see Figure 2). In the neutral block, participants responded to the color of a circle presented in the middle of a screen which was either associated with a left or right response button press, counterbalanced across subjects. In the experimental block, CG and IG trials were presented with the green or blue circle presented to the left or right of the central fixation cross. On CG trials, participants responded with a spatial CG mapping (e.g., a left green circle designates a left-hand response). On IG trials, participants responded with a spatial IG mapping (e.g., a left blue circle designates a right-hand response). On neutral trials, the circle was presented centrally on the screen and participants responded according to the corresponding mapping rule (e.g., green circle leads to a left-hand response).

During the variable oversampling interval and the following 500 msec, a fixation cross was presented. Participants were asked to maintain fixation. Then a warning signal (a red exclamation mark) appeared for 500 msec in the middle of the screen. Finally, the stimulus was displayed until the response interval exceeded 1500 msec. One trial lasted 4 sec.

The experiment consisted of 240 neutral trials and 240 experimental trials. The length of the experimental block was 18.6 min, whereas each of the neutral blocks lasted for 9.3 min, respectively. We also included 20 null events in each neutral block and 40 null events in the experimental block. These were pseudorandomly interspersed. The null events were included to compensate for the overlap of the BOLD response between adjacent trials. The experiment lasted about 60 min. Every block started out with two dummy trials that were excluded from further analysis.

**MRI Scanning Procedure**

The experiment was carried out on a 3-T scanner (Philips, the Netherlands). Thirty axial slices were acquired (222 $\times$ 2 mm FOV, 96 $\times$ 96 in-plane resolution, 3 mm slice thickness, 0.3 mm slice spacing) parallel to the AC–PC plane and covering the whole brain. We used a single-shot, gradient-recalled echo-planar imaging (EPI) sequence (repetition time [TR] = 2000 msec, echo time [TE] = 28 msec, 90° flip-angle [FA], transversal orientation). Prior to the functional runs, a 3-D T1 scan was acquired (T1 TFE, 25 $\times$ 2 cm FOV, 256 $\times$ 2 in-plane resolution, 182 slices, slice thickness 1.2, TR = 9.69, TE = 4.6, FA = 8°, sagittal orientation). Stimuli were presented to the participants on a back-projection screen, which
could be viewed via a mirror system attached to the MRI head coil.

fMRI Analysis

Analysis was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.63, a part of FSL (FMRIB’s Software Library; www.fmrib.ox.ac.uk/fsl). The first two volumes were discarded to allow for T1 equilibrium effects. The remaining images were then realigned to compensate for small head movements (Jenkinson, Bannister, Brady, & Smith, 2002). Data were spatially smoothed using a 5-mm full-width half-maximum Gaussian kernel. The data were filtered in the temporal domain using a high-pass filter with a cutoff frequency of 1/50 Hz to correct for baseline drifts in the signal. Finally, the functional data were prewhitened using FSL (Woolrich, Ripley, Brady, & Smith, 2001).

All functional datasets were individually registered into three-dimensional (3-D) space using the participants’ individual high-resolution anatomical images. This 3-D reference dataset was acquired for each participant at the beginning of each scanning session and used to normalize the EPI data toward Montreal Neurological Institute (MNI) space by linear scaling (affine transformations; Jenkinson & Smith, 2001). The statistical evaluation was performed using the general linear model. The design matrix was generated with a synthetic hemodynamic response function and its first derivative. Contrast maps were generated for each participant. Because the individual functional datasets were all aligned to the same stereotactic reference space, a group analysis was then performed. Higher-level analysis was carried out using FLAME (FMRIB’s Local Analysis of Mixed Effects) (Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004; Beckmann, Jenkinson, & Smith, 2003). For the whole-brain analysis of CG versus neutral trials and IG versus neutral trials, we only report cortical regions with a height threshold of \( z > 3.09 \) and a cluster probability of \( p < .05 \), corrected for whole-brain multiple comparisons (using Gaussian random field theory [GRFT]). The contrast of IG versus CG trials did not reveal any significant activation. For the whole-brain covariance analyses, we only report cortical regions with a height threshold of \( z > 1.69 \) and a cluster probability of \( p < .05 \), corrected for whole-brain multiple comparisons (using GRFT).

The aim of the covariance analyses of the present study was twofold. In a first step of the covariance analyses, individual contrast maps were computed for the contrast of IG compared to neutral trials in order to account for general RT effects on the BOLD signal, such as higher distraction, and difficulty. In a second step, the individual contrast maps on IG trials compared to neutral trials were included in three separate covariate analyses to investigate the dynamics of response capture and selective response suppression. Therefore, for each participant, the slope of the fastest IG CAF portion was incorporated as a covariate in the first regression model of the fMRI analysis procedure to index direct response capture. Next, to account for individual differences in selective response inhibition, the three delta-slope values (see also Behavioral Analysis) for each participant were included as individual covariates in a second covariance analysis. Finally, to test the functional significance of these two previous covariance analyses, the individual effect size of the Simon interference effect was also entered as a covariate in the fMRI regression model.

To compute the percent signal change of the hemodynamic response of the pre-SMA and the right IFC, all voxels of the relevant contrast exceeding the critical threshold in the mean \( z \)-map were determined. We then extracted the time course of the signal underlying these activated voxels for each participant from the modeled data. The percent signal change was calculated in relation to the mean signal intensity across all time steps for these voxels. The signal change was averaged for the IG condition beginning with the presentation of the stimuli for each participant. Finally, this value was entered in a Pearson correlation to compute correlations between the three different segments of the CAFs and the percent signal change in the pre-SMA as well as the three delta slopes and the percent signal change in the right IFC, respectively (see Figures 3 and 4).

RESULTS

Behavior

Direct Response Capture

The first set of analyses of variance focused on the main effects of congruence on error scores. Overall, error rates were low and did not distinguish between CG trials (1.7%) and IG trials (3.6%) \([F(1, 23) = 2.46, p = .13]\). Next, we used distributional analyses to test the specific prediction of the dual-route model that fast IG responses are more error-prone. As shown in Figure 3 (top), there was a significant interaction between congruence and segment \([F(3, 69) = 5.87, p = .003]\). Whereas error rates on CG trials did not differ between segments, \( F < 1 \), error rates of IG responses was affected by segment \([S1 = 8.9\%, \ S2 = 3.1\%, \ S3 = 1.3\%, \ S4 = 2.0\% \), \( F(3, 69) = 8.72, p < .001]\]. In line with the activation–suppression hypothesis, slope analysis revealed that fast IG responses were more error prone than slower IG responses \([Slope 1 vs. Slopes 2 and 3, F(1, 23) = 5.10, p = .034]\), reflecting location-driven response capture.

Selective Inhibition

A second set of analyses of variance confirmed that congruence produced its typical effect on RT. Overall, IG responses (506 msec) were about 21 msec slower than CG responses (485 msec) \([F(1, 23) = 18.23, p < .001]\). In line with the activation–suppression hypothesis, slope analysis revealed that fast IG responses were more error prone than slower IG responses \([Slope 1 vs. Slopes 2 and 3, F(1, 23) = 5.10, p = .034]\), reflecting location-driven response capture.
This slowing on IG trials cannot be explained in terms of a speed-accuracy tradeoff. Distributional analyses confirmed the specific prediction of the dual-process model that the interference effect decreases with RT. The interference effect reduced respectively from 33 msec and 32 msec in Segments 1 and 2 to 23 msec and 24 msec in Segments 3 and 4 [see Figure 4, top, $F(3, 69) = 16.33, p < .001$]. Overall, the delta slopes connecting the segments varied significantly [$F(2, 46) = 5.83, p = .009$]. Slope 2 ($-0.15$) was significantly more negative-going than Slope 1 ($0.00, p = .006$), whereas values of Slopes 2 and 3 ($-0.19$) did not differ significantly ($F < 1$).

**Neuroimaging Data**

The present study set out to investigate the neural correlates of response-processing dynamics using a Simon interference task. More specifically, we adopted a model-driven fMRI approach to test the response capture and selective response suppression hypothesis. A first analysis focused on the main effect of congruence, the main effect of IG versus neutral and CG versus neutral trials (Table 1). Direct comparison of IG versus CG trials yielded no significant main effect of congruence. The main effects between IG versus neutral trials and CG versus neutral trials revealed common activations in the middle occipital gyrus. Moreover, the main contrast of IG versus neutral trials showed activation in the right frontal, temporal, and parietal cortex as well as in the sulcus calcarinus. In a second analysis, we focused on the neural substrates of direct response capture. Hence, we entered individual slope values derived for each of the three segments of the CAFs for IG responses as covariates in the fMRI regression analysis (see red line in the upper part of Figure 3). The results revealed...
a significant activation in the right pre-SMA. Steeper positive-going CAF slope values for IG trials covaried with enhanced pre-SMA activation. Note that this relation was found only for the whole-brain analysis when entering the slopes derived from the first segment of the CAFs, not for the two latter segments. However, there was a marginal significant correlation between the percent signal change derived from the pre-SMA and the third CAF segment (see Figure 3, middle panel). Because the interindividual variance in the third CAF segment was very low, the whole-brain covariance analysis did not yield any significant result, which is in line with our hypothesis.

Our third analysis focused on selective response suppression as measured with RT distributions and plotted with the delta-plot technique (see Figure 4). Therefore, the slopes of three different segments of the delta plots derived for each participant were entered as covariates in the fMRI regression model. The only significant frontal activation was obtained in the right IFC (Brodmann’s area 44) bordering the right anterior insula (see Figure 4). Moreover, activations were found in the right cuneus \((x = 12, y = -86, z = 34)\), the inferior temporal gyrus \((x = -52, y = -62, z = -2)\), and the primary motor cortex \((x = -4, y = -20, z = 72)\). Because of our strong a priori hypothesis, in the following we will focus on the functional role of the right IFC. The activation in the right IFC covaried specifically with the negative-going delta slope of the slowest segment of the RT distribution for the contrast of IG trials compared to neutral trials. No such covariation was observed for the slopes of the two earlier delta-plot segments. To further verify the functional significance of the slow delta-plot slope as a specific indicator for selective response inhibition, the individual effect size of the overall interference effect was also entered as a covariate in the

---

**Figure 4.** Selective response inhibition. Covariate analysis with individual parameters derived from 24 participants for each segment of the delta plots (depicted in red). Top panel displays the delta slopes separated in fast, middle, and slow segments. Middle panel shows Pearson correlations between the % signal changes \(x\)-axis) derived from the right IFC and the delta slope \(y\)-axis) for each segment of the delta plot. Bottom panel depicts the averaged activation across 24 participants rendered onto a template brain \((z > 1.64, p < .05, \text{whole-brain corrected})\) of the covariance analysis with the three segments of the delta slopes. The only significant prefrontal activation is obtained for the slowest segment of the delta plot in the right IFC (BA 44; \(x = 38, y = 20, z = 4\)). Coordinates are given in MNI space.
fMRI regression model. Importantly, this analysis did not yield any significant effects \((F < 1)\).

**DISCUSSION**

In the present study, we used a novel combination of RT distribution-analytical techniques with model-based fMRI. With this approach we aimed to investigate the neural bases of direct response capture and selective response suppression in the Simon interference task. We predicted that individual differences in behavioral indices of response capture should covary with individual differences in activation in the pMFC. Likewise, we predicted that specific behavioral indices of selective response inhibition would covary with activation in the IFC. These specific behavioral indices were derived on an individual basis from RT distributional analyses (cf. Ridderinkhof et al., 2005).

In a first step, we computed main contrasts for congruence as well as IG versus neutral and CG versus neutral trials. Interestingly, the only frontal activation was found for IG trials compared to neutral trials in the right inferior precentral sulcus. When the threshold was lowered slightly \((z > 2.6)\), activation was also elicited in the pre-SMA. Given the latter threshold, the contrast of CG versus neutral trials revealed a comparable activation pattern as the IG versus neutral trials. The latter result highlights why the substantial congruence effect on behavior is not accompanied by activation patterns in the contrast between IG and CG conditions. One explanation for the missing congruence effect for the whole-brain analysis might relate to the equal probability of CG and IG trials that might have led to the recruitment of common brain areas in both conditions. A comparable discrepancy between behavioral effects and brain activation was also observed in task-switching studies using an equal probability of switch and repeat trials (for a discussion, see, e.g., Brass & von Cramon, 2004). However, previous fMRI studies with conflict tasks have often used equal probabilities and yet have observed activation in the IG–CG contrast. We speculate that, in the Simon task, response capture occurs in IG and CG trials alike, and that response inhibition is subsequently applied to control the location-driven response activation not only in IG trials but, to some extent, also in CG trials. As a consequence, the IG–CG contrast would not be very sensitive to the processes of response capture and inhibition, whereas the role of these processes is exposed unequivocally in the model-based analyses.

In general, the contrast for IG versus neutral trials revealed brain activations commonly reported in go/no-go tasks and interference tasks such as the Eriksen flanker, Stroop task, and SR compatibility tasks (see, e.g., Simmonds et al., 2008; Schumacher et al., 2007; Neumann et al., 2005; Hazeltine et al., 2003). This finding is noteworthy as it reveals a common neural network that is elicited when two competing responses are activated simultaneously. However, it still remains unclear which specific subprocesses are subserved by this network, as possible candidates range from action monitoring and conflict monitoring to mere signaling of conflicting situations (for a review, see Ridderinkhof et al., 2004). Hence, the aim of the present study was to investigate the neural bases of specific model-driven parameters reflecting response capture and response inhibition in the Simon task and to covary these with brain activation.

Evidence for brain areas activated in relation to direct response capture was found in the covariance analysis incorporating the accuracy slope of the fastest segment of the CAFs (see Figure 3, top). In line with our hypothesis,
and consistent with previous reports (Wittfoth et al., 2004; Peterson et al., 2002), activation was elicited in the pre-SMA on correct IG trials which were associated with elevated error rates. This finding was corroborated by a correlation analysis between the percent signal change derived from the pre-SMA and the error slope values derived for each participant, respectively (see Figure 3, middle).

In general, dorsal areas within the pMFC have been associated with goal-directed action selection processes in the context of multiple competing responses (cf. Rushworth et al., 2007). Here, activation of the pre-SMA appears to be involved in resolving the conflict among multiple responses that compete for activation (cf. Aron et al., 2007; Nachev et al., 2005; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999); the appropriate response (based on stimulus color) should be selected in favor of the response captured by the prepotent but task-irrelevant stimulus location. There is ongoing debate on whether action monitoring (including monitoring for conflicting responses) involves dorsal or somewhat more ventral areas of the pMFC (for a review, see Ridderinkhof et al., 2004). Although the RCZ encompasses mostly cingulate areas, it also extends into dorsal areas including the pre-SMA (Picard & Strick, 1996). The present data appear to corroborate the hypothesis that it is mainly the pre-SMA and not the RCZ that triggers conflict resolution (Wittfoth et al., 2006; Nachev et al., 2005; Ullsperger & von Cramon, 2001). Yet, it cannot be excluded that the presently observed pre-SMA activation also reflects, to some extent, the mere signaling of response conflict. It has been argued that the RCZ is activated when the context renders an elevated risk for conducting performance errors (Brown & Braver, 2005) or when it needs to work harder to avoid errors (Magnasco, Foxe, Molholm, Robertson, & Garavan, 2006). In the present study, however, overall error rates remained relatively low; further model-driven fMRI studies will be needed to further elucidate the role of pMFC areas in response capture.

Evidence for brain areas activated in relation to selective inhibition was found in the covariance analysis incorporating the slopes at the slow end of the delta plots (see Figure 4). Again in line with our hypothesis, and consistent with a previous finding (Peterson et al., 2002), the only prefrontal activation was found in the right IFC (BA 44) bordering the right anterior insula when entering the delta slope of the slowest segment. This finding was also corroborated by correlation analysis showing higher activation with negative-going slopes. The steepness of the slope has been referred to as reflecting the functional mechanism of selective suppression (see, e.g., Burle et al., 2002). In the present study, we can extend this notion by providing a linkage between the functional mechanism of selective suppression and a specific pattern of brain activation. There is broad neuroimaging literature as well as lesion data providing strong evidence for the right IFC to play a pivotal role in inhibitory processes (e.g., Aron et al., 2007; Aron & Poldrack, 2006; Casey et al., 2002; Garavan et al., 2002). Moreover, recent electrophysiological studies provide evidence for response inhibition to be a crucial mechanism in the Simon task (Burle et al., 2002). The present study contributes the first direct evidence that the right IFC is key in implementing selective response suppression when IG stimulus positions elicit response capture in the Simon task.

In order to test the specificity of the role of the right IFC in response inhibition, the model-based fMRI analyses were repeated with slopes for early and middle segments of the delta plots as covariates. No significant activations were observed in the right IFC or elsewhere. Our approach was validated further by entering the overall interference effect as a covariate in the analysis; this global interference measure indeed failed to reveal the specific brain areas associated with response inhibition. This outcome attests to the power of using specific behavioral parameters in model-based fMRI. In the present case, these behavioral indices were derived from RT distributional analyses that allow a focus on the efficiency of specific neurocognitive processes, typically not revealed in overall mean performance measures. By capitalizing on individual differences in these parameters (rather than pooling across them), this method sheds new light on the neural bases of response capture and selective response inhibition in the Simon task. In sum, we take the present results to indicate that, to the extent that the irrelevant position of the stimulus captures and activates the incorrect response, activation in the pre-SMA is elevated in order to resolve the conflict and select the appropriate response (as based on stimulus color), and activity in the right IFC is elevated in order to suppress the initial activation and execution of the incorrect response.

Acknowledgments

This work was supported by a VICI grant (B. U. F. and K. R. R.) and a VENI grant (W. P. M. vdW.) from the Netherlands Organization for Scientific Research (NWO). We thank Sanne Koemans for her help with data acquisition, Steven Scholte for technical support, and Uta Wolfensteller and Jan Derrfuss for helpful discussions.

Reprint requests should be sent to Birte U. Forstmann, University of Amsterdam, Department of Psychology, Amsterdam Center for the Study of Adaptive Control in Brain and Behavior (Acacia), Roetersstraat 15, 1018 WB Amsterdam, the Netherlands, or via e-mail: b.u.forstmann@uva.nl.

REFERENCES


Forstmann, van den Wildenberg, and Ridderinkhof 1865


