Right hemispheric dominance of inhibitory control: An event-related functional MRI study

H. GARAVAN*, T. J. ROSS, AND E. A. STEIN

Department of Psychiatry, Medical College of Wisconsin, Milwaukee, WI 53226

Edited by Marcus J. Raichle, Washington University School of Medicine, St. Louis, MO, and approved May 11, 1999 (received for review February 22, 1999)

ABSTRACT Normal human behavior and cognition are reliant on a person's ability to inhibit inappropriate thoughts, impulses, and actions. The temporal and spatial advantages of event-related functional MRI (fMRI) were exploited to identify cortical regions that showed a transient change in fMRI signal after the withholding of a prepotent motor response. The temporal specificity of the event-related fMRI design also minimized possible contamination from response inhibition errors (i.e., commission errors) and other extraneous processes. Regions identified were strongly lateralized to the right hemisphere and included the middle and inferior frontal gyri, frontal limbic area, anterior insula, and inferior parietal lobe. Contrary to the prominence traditionally given to ventral frontal regions for response inhibition, the results suggest that response inhibition is accomplished by a distributed cortical network.

The ability to suppress irrelevant or interfering stimuli or impulses is a fundamental executive function essential for normal thinking processes and, ultimately, for successful living. Response-inhibition deficits have been implicated in such clinical syndromes as attention deficit hyperactivity disorder (1), Tourette's syndrome (2), obsessive-compulsive disorders (3), and assorted “disinhibition syndromes” (4). The ability to selectively attend to a subset of a complex environment, activate appropriate meanings during verbal and written language comprehension, and activate appropriate memories at encoding and retrieval may all be critically dependent on the ability to suppress interfering stimuli, interpretations, and memories, respectively (5). Onset of inhibitory abilities, as indexed by the A-not-B task, also marks an important milestone in cognitive development and is considered a characteristic of frontal lobe maturation (6). Understanding the neuroanatomical bases of inhibitory control, knowledge that could be used to understand its development and its contribution to normal cognitive functioning or pathological disruption, provided the impetus for the present study.

Neuroanatomical Basis of Inhibitory Control. Typically, inhibition in a laboratory setting is operationalized as the suppression of a prepotent response. By using such an approach, activity in ventral prefrontal cortices has been shown to be modulated by manipulating the probability of nontargets, thereby affecting response-inhibition demands (7). Activation in ventral prefrontal regions has also been shown to separate subjects with high false-alarm rates from those with low false-alarm rates on a discrimination task (8). However, consistent with the distributed activations that underlie most cognitive processes, neuroimaging studies have revealed cerebral activation beyond ventral frontal regions during response inhibition. To substantiate this point, Casey and colleagues (7) observed that dorsolateral prefrontal regions were responsive to target probability, whereas anterior cingulate activation also discriminated subjects with high and low false-alarm rates (8). The distributed network thought to underlie response inhibition, as observed with neuroimaging studies, includes the supplementary motor area (SMA) (9–11), dorsal and ventral frontal regions (7, 10–13), anterior cingulate (8, 11, 14, 15), and the occipital and parietal lobes (9, 14, 16).

This widespread activation is somewhat at odds with the animal lesion literature, which typically ascribes response inhibition to predominantly ventral frontal regions (16, 17). One hypothesis to reconcile these observations is that the additional regions detected in human neuroimaging experiments may not actually be necessary for response inhibition but may reflect additional cognitive processes recruited when performing inhibition tasks. This hypothesis may be especially germane when task activation is averaged over blocks of trials that may include both response errors and post-error processes in which subjects may engage. Alternatively, response inhibition may be subserved by a network of cortical areas, of which the ventral frontal region may be a necessary but not sufficient component. To determine whether response inhibition is indeed mediated by more than inferior frontal regions, we exploited the temporal (and cognitive) specificity afforded by event-related functional MRI (ER–fMRI).

ER–fMRI. To maintain response prepotency, a mixing of frequent responses and infrequent response inhibitions is ideal. The ER–fMRI design allows response inhibitions to be distributed throughout the testing session rather than grouped together, as in a block design. The desirable characteristic that response inhibition should be effortful can result in significant numbers of errors of commission. Including these errors in a block design adds both unwanted variance to the response inhibition condition and unknown post-error processes, as subjects are typically aware when they have committed errors. This contamination may be especially problematic if one is comparing populations that differ in performance (18). Block designs may also require separate stimulus and response controls for cognitive subtraction (14, 18). Most advantageously, the ER–fMRI design allows one to exclude errors from the functional analyses, thereby isolating activation associated solely with correct response inhibitions.

METHODS

Task. Fourteen right-handed subjects (6 female; mean age, 31 ± 8; range: 19–44) gave informed consent and participated in this study, which was approved by the institutional review board of the Medical College of Wisconsin. Subjects were presented with a stream of letters presented serially every 500 msec with 0-msec interstimulus interval. Subjects made a button response whenever certain target letters (X or Y) were

*To whom reprint requests should be addressed. e-mail: hgaravan@mcw.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.
presented. An additional rule stipulated that responses must alternate between the targets, such that if the most recent target was, for example, the letter X, then subjects should only respond to the next letter Y in the letter stream (see task schematic in Fig. 1). Consequently, subjects would have to inhibit responding if the next target letter was an X. These nonresponse targets (lures) were presented on average every 20 sec and valid targets every 3.5 sec. A minimum of 15 sec (30 letters) separated consecutive lures.

Prepotency to respond to the targets was established by prior training that contained few lures (two runs of 250 letters each with 75 targets and no lures, followed by one run of 250 letters with 38 targets and 7 lures). During the scanning session, response prepotency was maintained by including six times more valid targets than lures and through prior instructions that stressed fast responding (subjects were encouraged to respond to a target while it was still on screen, that is, within 500 msec). Subjects completed four runs with 250 letters per run. In total, there were 150 targets and 25 lures.

fMRI Parameters. Contiguous 7-mm sagittal slices covering the entire brain were collected by using a blipped gradient-echo, echo-planar pulse sequence (TE = 40 msec; TR = 2,000 msec; FOV = 24 cm; 64 × 64 matrix; 3.75 mm × 3.75 mm in-plane resolution). All scanning was conducted on a 1.5T GE Signa scanner equipped with a 30.5 cm internal diameter three-axis local gradient coil and an endcapped quadrature birdcage radio frequency head coil (19). High-resolution spoiled gradient recalled acquisition at steady state (GRASS) anatomic images were acquired before functional imaging to allow subsequent anatomical localization of functional activation. Foam padding was used to limit head movements within the coil. Stimuli were back-projected onto a screen at the subject’s feet and were viewed with the aid of prism glasses attached to the inside of the radio frequency head coil.

fMRI Analyses. All data processing was conducted with the software package AFNI version 2.2 (20). In-plane motion correction and edge-detection algorithms were first applied to the functional data. Subjects whose fMRI time series had perceptible, residual head movements, based on cinematic viewing, were excluded from further analysis. Functional images were time-locked to the lures and averaged to obtain a mean signal response for each voxel. Voxels to which subjects responded (errors of commission) were not included in this signal averaging. These time-series averages were modeled with a γ-variate function by using a nonlinear regression optimization procedure (21). The γ-variate function has been shown to effectively model the hemodynamic response (22). The exponential parameters of the γ-variate function were constrained to a range similar to previously published estimates (22), the onset times were constrained to occur within 4 sec of the lure event, and the magnitude parameter of the model was free to vary (see Fig. 2). The nonlinear regression procedure thus arrived at the best-fitting function for each voxel time series while retaining a hemodynamic waveform. Area under the curve for each voxel was expressed as a percentage of area under the baseline (the continuation of the linear fit component of the model; see Fig. 2) and was the measure used for the group analyses.

The percentage area-under-the-curve maps were converted to a standard stereotaxic coordinate system (23) and spatially blurred by using a 4.2-mm full-width-at-half-maximum isotropic Gaussian filter. A one-sample t test against the null hypothesis of no effect was performed voxel-by-voxel on the percentage area-under-the-curve measure of the 14 subjects. Finally, only clusters of significant voxels with a minimum size of 100 μl were preserved (roughly equivalent to the size of the originally acquired voxels). A false-positive statistical threshold level was established with an identical analysis sequence but with randomly selected time-series averaging points. That is, the average response for each voxel was not based on the lure locations but instead on an equal number of randomly selected locations. The images of the time series were time-locked and averaged as described above, thus preserving whatever autocorrelation and drift trends that may have existed in the data. A separate randomization was performed for each voxel of each subject (~10,000 voxels), and then t tests were performed on each of these voxels across the 14 subjects. This method of randomization is computationally much more efficient than performing multiple different randomizations, each of which are applied to all voxels, because it provides as many different randomizations as there are voxels. When combined with the cluster size criterion, a t value threshold was set such that no randomized voxels survived, establishing a false-positive level that approaches 0 and that accounts for the multiple comparisons that have been performed. This threshold value was 4.9 with an associated P value of 4.0 × 10⁻⁷ and was used as the threshold for the real data. The advantages of combining a voxel-based threshold with a minimum cluster size have been described elsewhere (24).

To determine the specificity of any activations observed to underlie response inhibition, a similar analysis was conducted for the successful response executions to target letters. Given the high frequency of the motor responses, a signal-
Behavioral Analyses. Errors of commission (responses to lures only) and omission were tallied, and average response times to targets were calculated for each subject. Responses were scored as omissions if subjects did not respond within 1 sec of the target presentation.

RESULTS

Behavioral Results. Subjects performed well on the task, making few commission errors (mean = 1.9; SD = 1.0; range = 0–3) and few omission errors (mean = 1.4; SD = 2.5; range = 0–9). No errors of commission were due to the subject missing the previous target (in which case, the subsequent lure would be considered a valid target by the subject). Response times to targets averaged 460 msec (SD = 46).

fMRI Results. A number of discrete areas of activation, predominantly right-lateralized, were observed to underlie response inhibition. These areas, plus the Talairach coordinates of their centers-of-mass are presented in Table 1.

Activation was most pronounced in frontal and parietal regions, specifically on the middle and inferior frontal gyr and in the inferior parietal lobule and angular gyrus. The insula, frontal limbic area, and fusiform gyri were also activated. Fig. 3 illustrates the main clusters of activation.

Correlational analyses between the mean response times to the targets and the mean cluster signal intensity were performed. Of the 14 clusters, 2, located on the right inferior frontal gyr [r = −0.63, P = 0.02] and in the left inferior parietal lobule [r = −0.67, P = 0.009], showed significant correlations. Fig. 4 shows that the faster a subject was in responding to the targets, the greater the signal intensity seen in these regions when inhibiting a prepotent response.

An analysis of activations associated with response executions to target letters identified an anticipated motor circuit. This included left primary motor (recall that all subjects were right-handed), SMA, striatal, and cerebellar regions. Critically, activation did not include those regions identified with the lure activations (see Fig. 5). To ensure that the lack of overlap was not due to a statistical thresholding artifact, a more liberal false-positive level of 0.001 with no minimum cluster size criterion was used. Even with this thresholding, only minimal overlap, in the area of the pre-SMA, was observed. In addition, mean signal intensity within each of the lure activation clusters was calculated for both the lures and targets, and these means were compared by using paired t tests. All clusters showed significantly greater activation in response to the lures than in response to the targets (P values ranged from 2.0 × 10⁻⁴ to 6.0 × 10⁻¹¹), further suggesting the specificity of the regions activated during response inhibition.

Table 1. Clusters of statistically significant contiguous activation larger than 100 μl associated with response inhibition

<table>
<thead>
<tr>
<th>Lobe</th>
<th>Side</th>
<th>Brodmann’s area</th>
<th>Volume, mm³</th>
<th>CM RL</th>
<th>AP</th>
<th>IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>Right</td>
<td>10</td>
<td>102</td>
<td>42</td>
<td>40</td>
<td>−1</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>Right</td>
<td>9</td>
<td>783</td>
<td>36</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>109</td>
<td>34</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>41</td>
<td>36</td>
<td>22</td>
</tr>
<tr>
<td>Insula</td>
<td>Right</td>
<td></td>
<td>504</td>
<td>33</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>108</td>
<td>36</td>
<td>8</td>
<td>−3</td>
</tr>
<tr>
<td>Limbic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal limbic area†</td>
<td>Bilateral</td>
<td>32</td>
<td>554</td>
<td>1</td>
<td>16</td>
<td>42</td>
</tr>
<tr>
<td>Occipital</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>Right</td>
<td>19</td>
<td>269</td>
<td>40</td>
<td>−79</td>
<td>0</td>
</tr>
<tr>
<td>Superior occipital gyrus</td>
<td>Right</td>
<td>19</td>
<td>132</td>
<td>32</td>
<td>−74</td>
<td>29</td>
</tr>
<tr>
<td>Parietal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angular gyrus</td>
<td>Right</td>
<td>39</td>
<td>151</td>
<td>33</td>
<td>−60</td>
<td>36</td>
</tr>
<tr>
<td>Inferior parietal lobe</td>
<td>Left</td>
<td>40</td>
<td>126</td>
<td>−44</td>
<td>−39</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Right</td>
<td>40</td>
<td>116</td>
<td>48</td>
<td>−46</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>399</td>
<td>41</td>
<td>−54</td>
<td>45</td>
</tr>
<tr>
<td>Temporal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>Left</td>
<td>37</td>
<td>233</td>
<td>−43</td>
<td>−63</td>
<td>−13</td>
</tr>
</tbody>
</table>

Coordinates of the centers of mass (CM) of each cluster, in mm relative to the anterior commissure, are provided [positive values denote right hemisphere, anterior and superior to the anterior commissure for right–left (RL) anterior–posterior (AP), and interior–superior (IS), respectively].

†Located inferior to the superior frontal gyrus and superior to the cingulate gyrus (25).
DISCUSSION

Response Inhibition Lateralization. The most notable result of the present study is the right-hemisphere dominance in response inhibition. This lateralization is consistent with previous neuroimaging studies in normal subjects (9–12, 15, 26). The importance of the right hemisphere for inhibitory functions has also been suggested in the clinical domain (for reviews, see refs. 4 and 27 implicating right-hemisphere dysfunction in various disinhibition syndromes). A recent MRI-based morphological comparison of attention deficit hyperactivity disorder children and age-matched controls reported significant relationships between response inhibitory abilities and right frontostriatal circuitry (28). Right prefrontal volume seemed especially important, correlating positively with inhibition accuracy on a sensory selection task for control subjects but correlating negatively with inhibition accuracy in a response execution task for attention deficit hyperactivity disorder subjects. Furthermore, both obsessive-compulsive disorder severity and clinical improvement have correlated with right prefrontal glucose metabolism (29). Why inhibitory control should be lateralized to the right hemisphere is uncertain. Functional hemispheric asymmetries are not uncommon, however, having been proposed for prefrontal involvement in modality-specific working-memory tasks (30) and in episodic memory, lateralized by content modality (31) or by encoding and retrieval function (32, 33).

Distributed Network Underlying Response Inhibition. As listed in Table 1, activation associated with response inhibition was observed in a distributed network of brain regions. Frontal lobe activity was observed in both ventral and dorsal areas. Although inferior frontal gyrus activation has been reported frequently (10, 12, 14, 18, 26), most of the prefrontal activation of the present study was located within the middle frontal gyrus (also reported in refs. 10, 14, and 18). In a comparison of inhibitory performance in children and adults, Casey and colleagues (14) reported that the volume of activation in the middle frontal gyrus was the only frontal region that correlated significantly with age, thus suggesting that this structure may be most important in the developmental improvement observed in inhibitory abilities. On a cautionary note, magnetic susceptibility artifact can diminish functional signal in orbitofrontal regions and did so in the more anterior and medial ventral brain areas in the present study. This loss of signal probably reduced the sensitivity to detect changes in functional signal in these areas.

Given its role in fine motor control, the observed parietal lobe activation may be because of the demand to coordinate the movements involved in retracting the motor response (34). Deiber et al. (35) have shown that right-handed movements activate bilateral parietal regions. These authors report activation in the right inferior parietal lobe in a comparison of internally generated movements with a fixed, repetitive movement sequence. Alternatively, the parietal activation may reflect a transient increase in attentional processes, although neuroimaging studies have typically found parietal activation during attentionally demanding tasks to be in the superior, rather than inferior, parietal lobule (36–38). The role of the right insula activation is difficult to interpret but does fall close (cluster centers are ≈11 mm apart) to a frontal operculum region reported to be activated by both an auditory and a visual Go/No-Go task (26). The insula is thought to play a role in sensory and motor functions, both somatic and visceral, but has also been implicated in cognitive functions such as language, working memory, and selective visual attention (39). It has connections with most of the areas activated in this study, sending efferents to pre-SMA, SMA proper, and inferior frontal gyrus and has reciprocal connections with prefrontal cortex, cingulate, and anterior inferior parietal cortex. We speculate that the insula’s role in the current task may be in the motor act, or resulting somatosensation, of retracting a response rather than in the more cognitive components of response inhibition, a speculation that will have to await further research.

The frontal limbic activity lay just superior to the cingulate gyrus and overlapped with a region that Humberstone et al. (9) identified as pre-SMA. In a comparison of Go and No-Go
trials, these authors identified the pre-SMA, but not the SMA proper, as critical for response inhibition. Consistent with the importance of this region is a case study (40) of a patient with a very specific inhibitory deficit, evidenced by extremely high commission error rates on a Go/No-Go task and impaired Trail-Making B performance. The patient’s performance on a battery of other neuropsychological tests that did not load heavily on inhibitory abilities was normal. After the removal of an 8-cm tumor located anteriorly above the corpus callosum on both sides of the falk, the patient’s impairment was dramatically reversed. The authors reported that bilateral medial aspects of the frontal lobes, especially on the right side, were affected by the tumor, which had also placed pressure on the surrounding tissue. This tumor may well have impinged on the frontal limbic area observed in the present study.

The lack of overlap between the activation maps for response inhibitions and response executions and the significant differences between the two underscore the specificity of the observed network as a map of response inhibition. Little contamination from potentially common mechanisms (e.g., identification of and increased attentiveness to the salient letters, X and Y) seemed to have occurred. Consistent with the conclusions of Humberstone et al. (9), we have identified the pre-SMA as critical for response inhibition and SMA proper for response execution. All of the medial wall activation during response inhibition was rostral to the anterior commissure, whereas almost all medial wall activation during response execution was caudal to the anterior commissure; the anterior commissure serves as a useful anatomical landmark for delineating pre-SMA and SMA proper (41). The pre-SMA receives inputs from frontal and cingulate regions, whereas the SMA proper projects to primary motor cortex and to the spinal cord. The pre-SMA is thus well situated to participate in the response decision and the SMA proper in the execution of the motor response.

Prefrontal Basis of Inhibitory Control. Animal lesion research has implicated ventral prefrontal regions in response inhibition (16, 17). In contrast, neuroimaging research, including the cognitive specificity afforded by the ER–fMRI design of the present study, has identified appreciable functional activation in other, more dorsal prefrontal regions during response inhibitions. Dorsolateral prefrontal regions have frequently been implicated to be essential for working memory processes (42, 43). Although the present task did include a working-memory component (identifying a letter as a lure depends on storing the identity of the last valid target), it is unlikely that the working memory demand should be transiently increased during lure presentation (rather, it is after the presentation of a target that working memory contents must be updated). Furthermore, the working memory demand of the current task is not present in other neuroimaging studies that have observed dorsolateral prefrontal activations during response inhibition. Why the same neuroanatomical substrate might be important for both working memory and inhibitory control functions may be explained if one appreciates the importance of inhibitory mechanisms for much of cognitive functioning. As others have suggested (14, 44), working memory and inhibition may reflect the same mechanism insofar as normal working memory functioning requires the suppression of interfering stimuli; intimately related to (or perhaps synonymous with) the activation of the current contents of working memory may be the suppression of other competing items. Inhibitory and working memory abilities have been observed to overlap psychologically (45), and activation associated with adding an inhibitory requirement to a working memory task has been localized to dorsolateral prefrontal regions (46). Our data would support this suggestion of an overlap between working memory and inhibitory abilities. Nonetheless, it remains to be demonstrated, perhaps in a within-subject comparison across different task types, if the inhibitory processes underlying the response suppression of the present task are indeed the same ones, with the same anatomical substrates, as those involved in working memory interference suppression.

An alternative interpretation for the dorsal prefrontal activation might invoke different working memory functions. Consistent with Baddeley’s conception of a central executive that performs attentional control of action (47) and Norman and Shallice’s model of an intervening supervisory attentional system that overrides and redirects ongoing behavior (48), an executive involvement that initiates the less frequent response-inhibition act in the face of a competing prepotent response may be driving this dorsolateral prefrontal activation. We have obtained other evidence implicating a critical role for this region of the right middle frontal gyrus in executive functions with a dynamic working memory task that parametrically manipulated executive demands while holding constant rehearsal and storage demands (49). This explanation suggests that the circuitry that we have observed to underlie response inhibition may be decomposable into distinct subprocesses. The correlational analysis between response times to targets and inferior frontal activation might suggest a critical role for this region in signaling a “stop” message if one interprets the correlation as meaning that those subjects that responded fastest required a larger inferior frontal signal to inhibit responding. The dorsal prefrontal involvement may then play a role in integrating this information and in executing the appropriate action commands. Although this functional interpretation is speculative and does not incorporate the significant correlation between target response times and the left inferior parietal activation, it may accommodate both the critical role typically ascribed to ventral prefrontal involvement in response inhibition and the frequent observation of dorsal prefrontal activation during working memory and response inhibition tasks. However, even if response inhibition can be decomposed, either along the lines outlined above or otherwise, it is still clear that the act of inhibiting a prepotent response requires the orchestrated participation of a number of distinct structures distributed throughout the brain.

**CONCLUSION**

The selective averaging used in the present study, in which only correct response inhibitions were included in the functional mapping, allows us to discount error-based contributions to the functional maps. This feature of the ER–fMRI design, combined with an analytic technique that was tailored to detect a characteristic waveform response at the time of the response inhibitions, suggests that response inhibition is performed by a distributed collection of primarily frontal and largely right hemisphere regions.

The assistance of B. Doug Ward and J. F. Kussman is gratefully acknowledged. This work was supported by National Institute on Drug Abuse Grants DA09465, DA10214, and CRC RR-00058.
