The Counting Stroop: An Interference Task Specialized for Functional Neuroimaging—Validation Study With Functional MRI

George Bush,1–3* Paul J. Whalen,2 Bruce R. Rosen,1,3 Michael A. Jenike,2 Sean C. McInerney,2 and Scott L. Rauch1–3

1Nuclear Magnetic Resonance Center, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02129
2Department of Psychiatry, Psychiatric Neuroscience Program, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02129
3Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

Abstract: The anterior cingulate cortex has been activated by color Stroop tasks, supporting the hypothesis that it is recruited to mediate response selection or allocate attentional resources when confronted with competing information-processing streams. The current study used the newly developed “Counting Stroop” to identify the mediating neural substrate of cognitive interference. The Counting Stroop, a Stroop variant allowing on-line response time measurements while obviating speech, was created because speaking produces head movements that can exceed those tolerated by functional magnetic resonance imaging (fMRI), preventing the collection of vital performance data. During this task, subjects report by button-press the number of words (1–4) on the screen, regardless of word meaning. Interference trials contain number words that are incongruent with the correct response (e.g., “two” written three times), while neutral trials contain single semantic category common animals (e.g., “bird”). Nine normal right-handed adult volunteers underwent fMRI while performing the Counting Stroop. Group fMRI data revealed significant \( P \leq 10^{-4} \) activity in the cognitive division of anterior cingulate cortex when contrasting the interference vs. neutral conditions. On-line performance data showed 1) longer reaction times for interference blocks than for neutral ones, and 2) decreasing reaction times with practice during interference trials (diminished interference effects), indicating that learning occurred. The performance data proved to be a useful guide in analyzing the image data. The relative difference in anterior cingulate activity between the interference and neutral conditions decreased as subjects learned the task. These findings have ramifications for attentional, cognitive interference, learning, and motor control mechanism theories. Hum. Brain Mapping 6:270–282, 1998.

Key words: Stroop; cognitive interference; attention; functional magnetic resonance imaging; cingulate; learning; habituation; practice; response selection; motor

Contract grant sponsor: Clinical Investigator Training Program of Harvard/MIT Division of Health Sciences and Technology-Beth Israel Deaconess Medical Center, in collaboration with Pfizer, Inc.; Contract grant sponsor: NIMH; Contract grant number: MH-01215.

The results from this study were presented in preliminary form at the Fourth Annual Meeting of the Cognitive Neuroscience Society in Boston, MA. (March 23–25, 1997)

*Correspondence to: George Bush, M.D., M.M.Sc., Psychiatric Neuroscience Program, Massachusetts General Hospital-East, CNY-9, Building 149, Thirteenth Street, Charlestown, MA 02129. Phone: (617) 724–6142; Fax: (617) 726–4078; E-mail: geo@nmr.mgh.harvard.edu

Received for publication 28 July 1997; accepted 7 April 1998

© 1998 Wiley-Liss, Inc.
INTRODUCTION

Cognitive interference occurs when the processing of one stimulus feature impedes the simultaneous processing of a second stimulus attribute. Although many investigators have reported aspects of this topic, the essence of the second experiment described in the landmark article of Stroop [1935] (hereafter, “Color Stroop”) has become the prototypical interference task. Elegant in its simplicity and extremely reliable [Santos and Montgomery, 1962; Jensen, 1965; Smith and Nyman, 1974; Schubo and Hentschel, 1977], it possesses the ability to tell us much about the essential mechanisms of attention and cognition in both normal humans and those with neuropsychological impairments [Treisman and Fearnley, 1969; Dyer, 1973; MacLeod, 1991].

A number of excellent review articles [Jensen and Rohwer, 1966; Dyer, 1973; MacLeod, 1991] have detailed the historical background and subsequent significance of Stroop's work. Actually, the “Stroop interference effect” traditionally describes the second of three different tasks reported on in Stroop [1935]. Specifically, he reported that it took longer for subjects to name the color of the ink that color words were written in when the ink color and color word did not match (e.g., the word red written in blue ink, correct answer blue) than it did for them to simply name the color of colored squares.

The underlying neural substrate of cognitive interference has yet to be determined. We focused upon the anterior cingulate cortex in this study, hypothesizing that it plays a central role in interference and attentional tasks by mediating response selection and/or by allocating attentional resources when confronted with competing information-processing streams, consistent with the theories of Vogt et al. [1992], Posner et al. [1993], Posner and Dehaene [1994], and Devinsky et al. [1995]. Although the exact mechanistic role of cingulate cortex in distributed attentional networks is debated [Posner and Petersen, 1990; Mesulam, 1990; Colby, 1991], the anterior cingulate is incorporated into all these models of attention. Convergent evidence from positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies has implicated the anterior cingulate cortex as playing a pivotal role in attention. Specifically, anterior cingulate activation has been observed during Color Stroop tasks [Pardo et al., 1990; Bench et al., 1993; George et al., 1994; Carter et al., 1995]. The anterior cingulate has also been recruited during a divided attention task [Corbetta et al., 1991; Bush et al., 1995].

Seeking to develop an fMRI probe that might shed new light on attentional processing in both normals and neuropsychiatrically impaired populations, we initially considered the traditional Color Stroop. Despite its numerous positive features, however, some serious drawbacks limit its use in the fMRI environment. In any cognitive experiment, it is important to have an objective measure of task performance. Concomitantly collected performance data are vital, in that they 1) show that subjects are actually engaged in the task and 2) allow one to examine changes in neuronal activity as it relates to subjects' performance as they practice the task. Unfortunately, speaking produces head movements that can exceed those tolerated by fMRI (which can be seriously affected by motion of only 2–3 mm). An alternative strategy of arbitrarily labelling four buttons with color names is not optimal, as it adds an undesired layer of cognitive complexity and requires subjects to have training in order to learn the designated button-color combinations. Use of a Color Stroop with fewer colors might avoid the issue of cognitive complexity, but a review [MacLeod, 1991] of relevant studies that varied stimulus set sizes [Golden, 1974; Gholson and Hohle, 1968; Ray, 1974; McClain, 1983; Williams, 1977] and response set sizes [Ray, 1974; Nielssen, 1975; La Heij et al., 1985] could not provide a definitive conclusion on effects of manipulating either parameter. Therefore, to facilitate comparisons with existing Stroop task neuroimaging data, a stimulus/response set size of four is desirable. With regard to practice effects, it is true that subjects can improve Color Stroop performance with extensive practice, but this generally occurs over a period of hours to days, if at all [MacLeod, 1991]. Use of a task that subjects learn quickly (such as the Counting Stroop) allows one to examine practice-related neuronal modulation with fMRI.

The Counting Stroop was created to address these issues. It is a button-press Stroop interference variant that 1) allows on-line response time measurements without requiring speech, and 2) demonstrates significant behavioral improvement within 100 trials. The current study used the Counting Stroop to identify and characterize the mediating neural substrate of cognitive interference, with particular attention to the activity of the anterior cingulate cortex.

MATERIALS AND METHODS

Subjects

Informed consent was obtained following the established guidelines of the Massachusetts General Hospi-
tal Subcommittee on Human Subjects. All 9 subjects (5 males, 4 females, mean age 24.2 years (SD, 2.3 years) were strongly right-handed as assessed by the Edinburgh Handedness Inventory [Oldfield, 1971]. All had normal or corrected-to-normal vision and were native English speakers. No subject had a history of neurological, major medical, or psychiatric disorder; and none were taking medication. All specifically denied a history of attention deficit disorder or serious head injury. Subjects entered the study with knowledge that they would be paid for each session.

Psychophysical procedures

For the Counting Stroop (see Fig. 1), subjects were told that they would see sets of 1–4 identical words appear on the screen, and were instructed to report, via button-press, the number of words in each set, regardless of what the words were. During “neutral” blocks, the words were names of common animals (dog, cat, bird, or mouse). During “interference” blocks, the stimuli consisted of the number words “one,” “two,” “three,” or “four.” Thus, both sets of stimuli were common words within a single semantic category, balanced for length of word. Subjects were instructed that the keypad buttons represented one, two, three, and four from left to right, and subjects used the index and middle fingers of each hand to respond. Subjects were explicitly told that the sets would change every 1.5 sec. Furthermore, they were 1) instructed to “answer as quickly as possible, but since getting the correct answer is important, do not sacrifice accuracy for speed,” and 2) told, “Do not ‘blur your vision’ in an attempt to make the task easier—Keep the words in sharp focus.” After instructions were reviewed, and just prior to entering the scanner, subjects completed a 1-min computerized practice version of the task (20 neutral trials followed by 20 interference trials).

As illustrated in Figure 1, scans started and ended with 30 sec of fixation on a small dot. Eye movements were not monitored. Subjects completed two scans each of the Counting Stroop, where four 30-sec blocks of the neutral words alternated with four interference blocks. Given a fixed interstimulus interval of 1.5 sec, subjects completed 20 trials during each (neutral/interference) block, 80 trials of each type during a single scan, and 160 total trials of each type during the two-scan session. The order of presentation, regarding the neutral and interference blocks, was counterbalanced across runs and subjects.

Functional MRI scanning techniques and data analysis

Functional MRI scanning techniques developed by the Massachusetts General Hospital NMR Center were used. These methods have been extensively described previously [Cohen and Weisskoff, 1991; Kwong, 1995; Tootell et al., 1995]. Subjects were scanned in a General Electric Signa 1.5 Tesla high-speed imaging device (modified by Advanced NMR Systems, Wilmington, MA) using a quadrature head coil. The Instascan software is a variant of the echoplanar technique first described by Mansfield [1977]. Head stabilization was achieved using a plastic bite bar, molded to each subject’s dentition. The subjects lay on a padded scanner couch in a dimly-illuminated room, and wore foam ear plugs that attenuate high-intensity scanner sounds, but allow spoken instructions to be heard well.

Stimuli were generated on a Macintosh 100 MHz PowerPC(tm) computer using a Radius interface (model 0355, Videovision), and projected, via a Sharp XG-2000V color LCD projector, through a collimating lens onto a hemicycircular tangent screen. This rear-projection screen was secured vertically within the magnet bore at neck level after the subject had been positioned. Subjects viewed the images through a
tented mirror placed directly in front of their heads. Individual words subtended approximately 1° of the visual angle vertically, and a group of four words subtended a visual angle of approximately 6° vertically.

Initially, a sagittal localization scan was done to provide both a reference for slice selection in later scans and a high resolution scan for localization according to Talairach and Tournoux [1988] (SPGR, 60 slices, resolution 0.898 × 0.898 × 2.8 mm). Next, shimming was done to maximize field homogeneity [Reese et al., 1995]. In the third scan series, subjects had an MR angiogram, taken in the form of a spoiled gradient recall (SPGR, resolution 0.78125 × .78125 × 2.8 mm), to identify large- and medium-diameter blood vessels. The fourth series was a set of T1-weighted high-resolution axial anatomic scans (resolution 3.125 × 3.125 × 9 mm). For the functional series, asymmetric spin-echo (ASE) sequences (TE = 50 msec, TR = 2,000 msec, flip angle 90°, FOV = 40 × 20 cm, matrix = 64 × 64, in-plane resolution 3.125 × 3.125 mm, slice thickness = 9 mm, 150 images/slice) were used to minimize macrovascular signal contributions. The angiogram, T1 anatomical, and ASE functional slices (series 3–5) were collected using identical slice prescriptions. Twelve contiguous, interleaved slices, parallel to the anterior-posterior commissure line, were obtained for all studies.

All data sets had the amount of motion quantified, and were then motion-corrected. The mean displacement for all studies was 0.9 mm (SD, 1.2 mm). Motion-correction was accomplished using an algorithm developed by Jiang et al. [1995], based on that of Woods et al. [1992]. This process corrupts the images from the first and last slices of a stack; therefore, these were discarded from further analysis. The functional scans were transformed into a standardized anatomical space [Talairach and Tournoux, 1988].

Statistical analysis of functional images for regions of significant change was accomplished using a multistep process. Nonparametric statistical maps were calculated using the Kolmogorov-Smirnov (KS) statistic, and displayed in pseudocolor, scaled according to significance, and (after reslicing into coronal orientation) superimposed on (resliced) T1-weighted high-resolution sagittal localizer scans. As an objective measure of activated regions, an automated region-defining algorithm was used on smoothed KS maps [Jiang et al., 1996; Bush et al., 1996]. Smoothing was done using a Gaussian filter with a sigma of 1.1, giving an effective resolution of 8.1 × 8.1 mm full width half maximum (FWHM). Significance values of local maxima (P ≤ 10−4) within these identified regions are reported based upon the native (unsmoothed) statistical maps.

Based on off-line pilot work, in which we observed subjects demonstrating improved performance over a period of minutes, we predicted that the group reaction time data could be used as a guide in analyzing image data. Specifically, we sought to examine scan epochs during which, as a group, subjects showed significant interference effects. We predicted that this approach would permit us to cull out those epochs that did not show a difference in reaction time between interference and neutral block-pairs (and thus would presumably be contributing noise to the analysis), therefore narrowing our focus to epochs containing processes involved in cognitive interference.

**RESULTS**

**Behavioral results**

As has been stated, one of the main goals for the study was to use the reaction time (RT) results as a guide in analyzing the imaging data. In this manner we could have a measure of when significant interference was occurring and also observe how practice-related learning of the task modulated neuronal activity. Subjects showed an overall increase in RT during interference blocks (mean 709 msec, SD 119 msec) as compared to neutral blocks (mean 663 msec, SD 107 msec). All 9 subjects exhibited longer RTs during interference blocks, displaying a mean increase in RT of 46 msec (SD 20). A repeated-measures condition (interference vs. neutral) × epoch (8 time epochs across 2 scans) ANOVA demonstrated a significant main effect for condition (F = 40.16, df = 1, P < 0.001) and epoch (F = 4.48, df = 7, P < 0.001), and a significant condition × epoch interaction (F = 3.3, df = 7, P < 0.01). Figure 2 illustrates this interaction, as the early RT disadvantage during interference vs. neutral epochs diminished with practice, decreasing to nonsignificant levels over the course of two scans. Planned comparisons demonstrated significant increases in RT for interference vs. neutral conditions during the first five block-pair epochs. Importantly, while the RTs during interference blocks were observed to decrease (presumably due to practice effects), the neutral task RTs remained relatively constant from beginning to end of the scans.

Accuracy data (percentage correct) remained stable throughout the scanning sessions for all subjects (mean = 96.6%, SD = 2.3%, range = 88.8–100%). Mean accuracy scores were no different for interference trials during scan 1 (94.5%, SD 2.3%) and scan 2 (96.3%, SD 3.0%) (NS), and they were no different for neutral trials (scan 1 = 97.2%, SD 2.1%; scan 2 = 98.2%, SD 2.0%)

---

**Footnotes**

1. Talairach and Tournoux [1988] (SPGR, 60 slices, resolution 0.898 × 0.898 × 2.8 mm).
2. Motion-correction was accomplished using an algorithm developed by Jiang et al. [1995], based on that of Woods et al. [1992].
3. As has been stated, one of the main goals for the study was to use the reaction time (RT) results as a guide in analyzing the imaging data. In this manner we could have a measure of when significant interference was occurring and also observe how practice-related learning of the task modulated neuronal activity. Subjects showed an overall increase in RT during interference blocks (mean 709 msec, SD 119 msec) as compared to neutral blocks (mean 663 msec, SD 107 msec).
4. Accuracy data (percentage correct) remained stable throughout the scanning sessions for all subjects (mean = 96.6%, SD = 2.3%, range = 88.8–100%). Mean accuracy scores were no different for interference trials during scan 1 (94.5%, SD 2.3%) and scan 2 (96.3%, SD 3.0%) (NS), and they were no different for neutral trials (scan 1 = 97.2%, SD 2.1%; scan 2 = 98.2%, SD 2.0%)
Subjects showed a small but statistically significant decrement in accuracy during interference trials (95.4%, SD 2.1%) as compared to neutral trials (97.7%, SD 1.9%) (t = -2.44, P = 0.02).

**Functional MRI results**

As predicted (and illustrated in Fig. 3a), activation was seen in the anterior cingulate when comparing all interference blocks with all neutral blocks (i.e., 8 block-pairs per subject). This anterior cingulate activation was in close proximity to those found in other Stroop neuroimaging studies [Pardo et al., 1990; George et al., 1994, 1997; Carter et al., 1995; Derbyshire et al., 1998]. This activity reflects the overall differences in RT observed between the interference and neutral blocks seen accompanying Figure 3b.

The RT performance data proved useful when interpreting the fMRI data. Interestingly, while using scan results from all 8 interference-neutral block-pairs gave a P = 6.7 × 10^-6 in the anterior cingulate region, analysis of only the first 5 block-pairs (i.e., those which demonstrated a significant RT difference; see Fig. 2) yielded a “more significant” P value in this region of interest of 9.9 × 10^-7. Thus, the relative difference in fMRI signal during the first 5 block-pairs alone was statistically more significant than when using all 8 block-pairs, despite the smaller number of images compared (and hence the test’s reduced power to detect a difference).

Therefore, using the behavioral data as a guide, the Counting Stroop was seen to activate a network of brain regions involved in attention, response selection, motor planning, and motor output. These areas included the anterior cingulate cortex, middle frontal gyri, premotor and primary motor cortex, inferior temporal gyrus, and superior parietal lobule (see Table I).

**Decreasing anterior cingulate activity with practice**

Activity within the anterior cingulate region paralleled the observed practice-related interference block...
RT decrease. A three-block running average analysis of activation in the right anterior cingulate (Fig. 4) illustrates this relationship in two ways. First, it shows a decrease in relative fMRI activation with task practice between the interference and neutral conditions. Second, it reveals that high significance values were obtained when comparing scan data taken from the first three interference-neutral block-pairs ($P < 1.6 \times 10^{-5}$), whereas comparing scan data for the last three interference-neutral block-pairs gave a nonsignificant result. Such results, taken together, indicate that the largest fMRI signal effect occurred during the first few blocks (i.e., when interference effects were greatest).

Such a decrease over time in the significance values comparing interference and neutral epochs would suggest that anterior cingulate activity progressively decreased during successive interference blocks, but

**TABLE I. Regions activated during Counting Stroop:**

<table>
<thead>
<tr>
<th>Region</th>
<th>Talairach coordinates</th>
<th>$P$ value</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43 27 34</td>
<td>$9.4 \times 10^{-5}$</td>
<td>R. middle frontal gyrus (BA 9/46)</td>
</tr>
<tr>
<td>2</td>
<td>−46 18 25</td>
<td>$9.4 \times 10^{-6}$</td>
<td>L. middle frontal gyrus (BA 9)</td>
</tr>
<tr>
<td>3</td>
<td>12 9 34</td>
<td>$9.9 \times 10^{-7}$</td>
<td>R. anterior cingulate (BA 24/32)</td>
</tr>
<tr>
<td>4</td>
<td>−56 −3 18</td>
<td>$5.5 \times 10^{-7}$</td>
<td>L. precentral gyrus (BA 4)</td>
</tr>
<tr>
<td>5</td>
<td>−40 −6 37</td>
<td>$1.8 \times 10^{-6}$</td>
<td>L. premotor cortex (BA 6)</td>
</tr>
<tr>
<td>6</td>
<td>43 −48 −6</td>
<td>$5.8 \times 10^{-5}$</td>
<td>R. inferior temporal gyrus (BA 37)</td>
</tr>
<tr>
<td>7</td>
<td>−28 −66 40</td>
<td>$2.7 \times 10^{-5}$</td>
<td>L. superior parietal lobule (BA 7)</td>
</tr>
</tbody>
</table>

* Stereotactic coordinates and statistical significance values for local maxima meeting threshold and cluster size criteria (see Materials and Methods) are presented according to the conventions of Talairach and Tournoux [1988]. Coordinates are expressed in millimeter units. The origin (0, 0, 0) is the anterior commissure at the midsagittal plane, with $x > 0$ corresponding to right of midsagittal, $y > 0$ corresponding to anterior, and $z > 0$ corresponding to superior. Brodmann’s areas are indicated after the named structure (in parentheses and prefaced by “BA”).

Running average of anterior cingulate activation parallels practice effects. Running averages of intersubject group averaged fMRI signal in anterior cingulate cortex were obtained by selectively comparing activation during specific interference vs. neutral block-pairs (first block-pair 1, then pairs 1 + 2, 1 + 2 + 3, 2 + 3 + 4, etc.) in order to detect changes in regional activation over time. Despite the drastic reduction in the number of time points compared, significant activity was observed to occur in the anterior cingulate in the initial block-pairs. Paralleling the reaction time data (Fig. 2), this activity was reduced with practice to nonsignificance over the course of the two scans.
could alternatively have been produced by a progressive increase in variance without a change in mean signal intensity. Figure 5, in fact, shows the former to have been the case (i.e., that the absolute anterior cingulate activation (as measured by fMRI signal) decreased over time, and that the relative difference in fMRI signal between interference and neutral block-pairs steadily decreased in parallel with the observed decrease in RTs). Furthermore, reference to Figure 5, with a large increase in signal during the second neutral epoch (an increase driven by data from a single subject outlier), also helps to explain why an ANOVA of the anterior cingulate cortex scan data which found a significant main effect of condition ($F = 8.33, df = 1, P = 0.02$) failed to find a significant time effect or interaction.

The alternative explanation for such a relative decrease in significance with time (i.e., progressively increasing variability of the anterior cingulate cortical activity (fMRI signal) over time) was examined and found not to explain the data. While the normalized fMRI signal observed in the anterior cingulate activation varied between a range of 533–610, the standard deviations of the fMRI signal found for each block remained quite constant within a narrow range (6.8–8.6). More importantly, the mean standard deviation of fMRI signal in the anterior cingulate cortex activation was found to remain the same for interference epochs of scan 1 (7.80) and scan 2 (7.80), and actually decreased over time from scan 1 (8.25) to scan 2 (7.71) for neutral epochs, indicating that differential variability in fMRI signal cannot explain the observed progressive decrease in significance illustrated in the running average of anterior cingulate activation (Fig. 4).

None of the other activations (in prefrontal, primary motor, premotor, and parietal cortex) showed a significant, consistent trend with practice. While a running average of the left prefrontal activation tended towards a progressive increase over time, an ANOVA did not support a significant time effect.

**DISCUSSION**

Functional MRI was used to examine the neural substrate of the Counting Stroop, a cognitive interference task. There were six principal findings: 1) The Counting Stroop successfully produced a state of cognitive interference, as revealed by prolonged reaction times during interference blocks as compared to neutral blocks. 2) The anterior cingulate cortex, which has been hypothesized to allocate attentional resources when faced with competing information-processing streams or to mediate response selection, was indeed more active during interference blocks of the Counting Stroop. 3) The Counting Stroop successfully permitted collection of performance data as subjects were scanned. This RT data provided independent verification of subjects’ engagement in the task, and by showing a significant condition × time interaction, enabled a more refined analysis of regional cortical activity (i.e., one focused upon epochs with demonstrated cognitive interference effects). 4) In addition to the anterior cingulate locus, a network of attentional/motor control cortical regions was also activated during interference vs. neutral blocks. 5) Progressive relative decreases between interference and neutral conditions in both on-line RT data and fMRI signal in the anterior cingulate suggest that practice led to modulation of neuronal activity in this area. 6) Although implicit in the above findings, the most exciting finding should be stated explicitly, namely, that two different cognitive
interference tasks (i.e., the Color and Counting Stroops) both activated the same subterritory of the anterior cingulate cortex. We assert, therefore, as the Color Stroop is an excellent task with which to study interference during PET image acquisition, that we have demonstrated that the Counting Stroop is a valid homologue for fMRI.

**Relationship of the Counting Stroop to other interference tasks**

Before discussing the neurobiological findings, it is important to place the Counting Stroop in perspective as an interference task. Although incorporating specific modifications for use in the fMRI environment, it retains the essential aspects of a traditional cognitive interference task. Namely, it places two cognitive processes, reading and counting, into competition during the interference (incongruent) trials. In contrast, the word stimuli used in the neutral trials (also common words from a single semantic category) do not interfere significantly with the counting process. Cognitive interference is manifested by the observed increase in RT required to count the stimuli when the elements to be counted are themselves numerical words. Although we took the advice of MacLeod [1991] and did not mix trial types in this initial effort, experiments are underway examining facilitation effects (i.e., congruency between the stimulus word and number of words, such as “two” written twice).

Other investigators have used similar numerosity-based interference tasks in psychophysical studies. Using a between-subjects design and verbal response, Windes [1968] found that subjects took longer to report the quantity of numerals than to name them. Similar results have been reported using a variety of techniques. Morton [1969], using a card-sorting paradigm with both digits and digit names, Shor [1971] and Fox et al. [1971], using a list reading format and stimuli including digits on die faces and random patterns, and Reisberg et al. [1980], using list reading of grouped digits or digit names, have all reported interference effects. Additionally, Flowers et al. [1979] showed that response mode (oral naming, card sorting, and manual and oral “tapping”) could influence RT results when using numerosity stimuli.

While these previous studies serve to support the underlying model (i.e., that Stroop-like cognitive interference can be obtained using incongruent numerical stimuli), some subtle yet important distinctions differentiate the Counting Stroop from these other tasks. The Counting Stroop’s use of a computerized single-trial format and its employment of a button-press response mode are important differences in that they increase the flexibility of study design possibilities and the accuracy of data recording. For example, future studies may focus upon the neurobiological differences, if any, underlying manual vs. verbal responses, as discussed below.

**Anterior cingulate cortex: Role in cognitive interference, attention, and response selection**

Perhaps the most important finding of the present study is that the Counting Stroop activated a subregion of anterior cingulate cortex, the anterior cingulate-cognitive division (ACcd, caudal areas 24'/32'), that has often been reported to be active in neuroimaging studies using variations of the Color Stroop paradigm [Pardo et al., 1990; George et al., 1994, 1997; Carter et al., 1995; Derbyshire et al., 1998], in other cognitive interference tasks [Taylor et al., 1994], in divided attention tasks [Corbetta et al., 1991; Bush et al., 1995], in response selection/generation tasks [Petersen et al., 1988; Frith et al., 1991; Paus et al., 1993; Kawashima et al., 1996], in anticipation of cognitively demanding tasks [Murtha et al., 1996], and in error detection tasks [Dehaene et al., 1994]. This “cross-activation” of ACcd by different tasks, as depicted in Figure 6, lends support to the hypothesized involvement of the anterior cingulate cortex in tasks that require subjects to resolve processing conflicts between competing information streams by sensory and/or response selection. Figure 6 also shows that activation in ACcd, obtained in response to tasks emphasizing sensory/response selection, can be contrasted with activity seen in the anterior cingulate-affective division (ACad, rostral areas 24/25/32/33) during response to emotionally-laden stimuli. Notably, the ACad activation by Whalen et al. [1997, 1998] was obtained in the same group of subjects as the present study, as they performed a Counting Stroop variant that substitutes emotionally charged words for numerical ones. This ACcd/ACad distinction is supported by a large body of work [Vogt et al., 1992; Devinsky et al., 1995; Mega and Cummings, 1997a,b; Mayberg, 1997].

As seen in Figure 6, the neuroimaging studies using cognitive/motor tasks (i.e., those emphasizing attentional allocation/response selection) do show some variability, but rather consistently activate the ACcd. The ACcd activation in the current study is approximately 3 mm from that reported by Pardo et al. [1990], and only 9 mm from loci identified by George et al. [1994] and Carter et al. [1995]. Bench et al. [1993] did report anterior cingulate in their incongruent minus colored cross-naming subtraction (albeit in ACad,
approximately 30 mm anterior to that of the Counting Stroop ACcd activation), though they did not obtain any anterior cingulate activation in their incongruent minus congruent (facilitation) comparison. It must be noted that with the exception of the direct attempt of Bench et al. [1993] in their experiment 2 to replicate the task conditions of Pardo et al. [1990], none of the five studies were directly comparable. Pardo et al. [1990] looked at incongruent words [interference] vs. congruent words (e.g., “red” written in red ink) (facilitation), and used a fixed interstimulus interval. Bench et al. [1993] used multiple contrast conditions and two different interstimulus intervals, pointing out the importance of stimulus and timing issues. The study of George et al. [1994] examined incongruent color words vs. color-naming of hatch marks, and utilized a self-paced format rather than fixed-interval timing. While all three of these PET studies (and the current fMRI study) used homogeneous block designs (i.e., blocks contained only one type of trial, such as incongruent words), the PET study of Carter et al. [1995] employed a heterogeneous block design (mixing trial types within blocks), fixed interstimulus intervals, and words (animal names written in colored ink) as the neutral condition. Given these differences in imaging modalities, experimental design, task characteristics, subtractions used, and stimulus attributes, as well as the degree of intersubject anatomic and physiologic vari-
ability, the degree of convergence in results from these studies implicating anterior cingulate is impressive. Stroop tasks are not the only ones that support the role of ACCd as a mediator of competing information streams. The ACCd region identified by the Counting Stroop directly overlaps the ACCd locus Corbetta et al. [1991] identified \((x = 7, y = 11, z = 34 \text{ mm})\) in a PET study of divided attention, as well as the ACCd region \((x = 6, y = 6, z = 46 \text{ mm})\) localized in an fMRI replication using the same divided-attention paradigm [Bush et al., 1995]. In these studies subjects were asked to monitor paired visual stimulus arrays for changes in the shape, speed, or color of the elements. The stimuli remained the same throughout, but during selective attention blocks subjects were instructed to attend only to a single feature (e.g., shape), whereas during the divided-attention condition subjects were required to report changes in any of the three features. ACCd was active in both studies during the divided-attention task, consistent with its attention-directing/response selection role when faced with information from multiple rival input channels.

Of particular relevance to this hypothesis are the reported PET findings of Kawashima et al. [1996]. Using finger movements as the output in a visual GO/NO-GO task, they found a network of regions including, among others, ACCd, and prefrontal, premotor, precentral, and insular cortex. Casey et al. [1996] also reported anterior cingulate and prefrontal cortical fMRI activation during a GO/NO-GO task. The GO/NO-GO task, much like the Stroop tasks, requires the subject to utilize a top-down generated processing strategy (i.e., to modify behavior based upon a relevant instruction set), and a key element is that a motor response must be inhibited. Although existing GO/NO-GO work in primates [Kalaska and Crammond, 1995] and humans [Rosahl and Knight, 1995; Casey et al., 1996; Kawashima et al., 1996] has focused on prefrontal cortical function in this type of motor inhibitory task, it is entirely plausible that the ACCd may play a more important role than heretofore recognized, especially in light of the extensive reciprocal connections the ACCd maintains with both prefrontal cortex and lower motor areas in humans and other primates [Pandya et al., 1981; Vogt and Pandya, 1987; Vogt, 1993; Dum and Strick, 1993; Bates and Goldman-Rakic, 1993; Devinsky et al., 1995].

Similarly, it should be noted that while the current study focuses attention on the cognitive division of the anterior cingulate, it does not presume that this region is the only one relevant to performance of cognitive interference, response selection, or attentional tasks. The present results, as well as numerous previous studies, support the role of a network of prefrontal, motor, and parietal cortical structures recruited during such tasks. This should not be surprising, as a number of researchers have proposed parallel distributed models of attention that include these structures [Goldman-Rakic, 1988; Posner and Petersen, 1990; Mesulam, 1990; Colby, 1991].

**Practice effects and attenuation of the anterior cingulate response**

Interest in the effects of practice on Stroop task performance has been evident from the start. In fact, the third experiment in the classic paper of Stroop [1935] bore the title, “The Effects of Practice Upon Interference.” There, he found that although extensive practice on his color-naming interference task did lead to an improvement in performance (i.e., a decrease in RTs while naming incongruent color words), such practice did not completely eliminate interference effects. Similarly, the careful treatment by MacLeod and Dunbar [1988] of this issue reported that performance of the Color Stroop improved after many hours of practice, a finding that is further supported by the work of Feinstein et al. [1994] and Ogura [1980]. While most agree, however, not all groups have found that practice lessens interference in Stroop-like tasks [Shor et al., 1972].

Other investigators have looked at different facets of the practice issue. Reisberg et al. [1980] observed rapid learning of a numerosity-based interference task. DuLaney and Rogers [1994], based on differential patterns of practice-related performance improvement on a modified Color Stroop, suggested that different mechanisms might underlie the reduction in interference in groups of young and old subjects. Cohen et al. [1990] even developed a neural networks simulation model hypothesized mechanisms responsible for improvement with practice on the Color Stroop.

In a particularly germane analysis, Roe et al. [1980] compared manual (button-press) vs. verbal response modes. They reported that extensive practice improved with both, but that interference effects declined more rapidly when subjects responded manually as opposed to verbally. Nielsens [1975] also found that manual responding was affected more quickly by practice, and White [1969] reported that manual responding was generally faster than verbal. Our results indirectly support these conclusions, as we observed significant improvement over the course of 5–6 min of practice. Despite the convergence among these studies,
however, the issue is not settled, as Redding and Gerjets [1977] and Roe et al. [1980] failed to obtain a significant difference between the two response modes. Future studies may be able to use the Counting Stroop format to resolve the questions of which mode is faster (manual vs. oral) and which mode (if either) is more subject to practice effects, without contamination from an unnecessary layer of cognitive complexity (i.e., converting Color Stroop responses to button-presses).

Our fMRI finding that ACcd activity during interference blocks (relative to neutral ones) decreases with practice is bolstered by the work of Raichle et al. [1994]. In this PET study of practice-related changes during a verbal response selection task, they found that activation of the ACcd steadily decreased with practice, and that the original level of activation could be restored by showing subjects a novel (unpracticed) set of stimuli. Parenthetically, it is difficult to explain away the observed practice-related decrements in the ACcd in the current study as a pure “novelty effect,” as the task was not novel (all subjects completed a 40-trial practice block before scanning), and the block order was effectively counterbalanced. Friston et al. [1992] reported similar decrements in activation (in the cerebellum) with practice of a simple finger sequencing motor task. Certainly, the results of these studies, as well as those of the current study, are in line with the supposition expressed by Adams [1987] that the central level of energy expenditure should decrease with skill acquisition, presumably because the processing becomes more efficient.

CONCLUSIONS

Based on the performance data, the Counting Stroop was observed to produce cognitive interference akin to that of the Color Stroop. The button-press response mode enabled collection of vital on-line performance data, a feature that should make the Counting Stroop a useful tool in future studies of normals and neuropsychiatric impaired populations. Certainly, in the current study, the RT data proved to be an important guide in analysis of the fMRI data. The fMRI scan data, highlighting the role of the anterior cingulate cortex (specifically the ACcd), were found to agree with those of other Color Stroop and divided-attention imaging studies. Also, in keeping with other studies of learning, the fMRI response of the anterior cingulate paralleled observed practice effects (as measured by the on-line performance data).

Future studies might exploit the various similarities and differences between the Counting and Color Stroops in studies of attention, cognitive interference, learning, and motor control mechanisms. Certainly, the Counting Stroop gains immensely by its ability to draw upon the enormous Color Stroop literature. In return, the Counting Stroop (and word substitution versions [Whalen et al., 1998]) can serve as a useful fMRI probe in searching for the neural substrates of various neuropsychiatric disorders such as attention deficit disorder, schizophrenia, obsessive-compulsive disorder, depression, parkinsonism, and head injuries [Benes, 1993; Kingma et al., 1996]. Beyond these potential clinical applications, single-trial fMRI [Buckner et al., 1996; Dale and Buckner, 1997], and/or combined use of fMRI and event-related potentials, may help settle whether the anterior cingulate is activated pre-stimulus [Murtha et al., 1996] and/or poststimulus [Abdullaev and Posner, 1997; Dehaene et al., 1994]. Finally, subtle task manipulations (e.g., trial sequencing, stimulus word substitutions) can help us to better understand mechanisms of attention, response selection, practice and learning effects, manual vs. oral responding, and cognitive interference and facilitation.

ACKNOWLEDGMENTS

The authors thank Mary T. Foley, Terrance A. Campbell, Richard McNally, Robert Weisskoff, Kenneth Kwong, Cary Savage, Michael Vevea, Thomas J. Brady, and our intrepid volunteers for their generous assistance. Support for this work was provided by the Clinical Investigator Training Program of the Harvard/MIT Division of Health Sciences and Technology-Beth Israel Deaconess Medical Center, in collaboration with Pfizer, Inc. (G.B.) and grant MH-01215 from NIMH (S.L.R.).

REFERENCES


