# Behavioral Conflict, Anterior Cingulate Cortex, and Experiment Duration: Implications of Diverging Data

**Kirk I. Erickson,[1]\* Michael P. Milham,[1] Stanley J. Colcombe,[1] Arthur F. Kramer,[1] Marie T. Banich,[2] Andrew Webb,[1] and Neal J. Cohen[1]**

[1]*Beckman Institute and University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois*
[2]*University of Colorado at Boulder, Boulder, Colorado*

**Abstract:** We investigated the relationship between behavioral measures of conflict and the degree of activity in the anterior cingulate cortex (ACC). We reanalyzed an existing data set that employed the Stroop task using functional magnetic resonance imaging [Milham et al., Brain Cogn 2002;49:277–296]. Although we found no changes in the behavioral measures of conflict from the first to the second half of task performance, we found a reliable reduction in the activity of the anterior cingulate cortex. This result suggests the lack of a strong relationship between behavioral measurements of conflict and anterior cingulate activity. A concomitant increase in dorsolateral prefrontal cortex activity was also found, which may reflect a tradeoff in the neural substrates involved in supporting conflict resolution, detection, or monitoring processes. A second analysis of the data revealed that the duration of an experiment can dramatically affect interpretations of the results, including the roles in which particular regions are thought to play in cognition. These results are discussed in relation to current conceptions of ACC's role in attentional control. In addition, we discuss the implication of our results with current conceptions of conflict and of its instantiation in the brain. *Hum. Brain Mapp. 21:98–107, 2004.* © 2003 **Wiley-Liss, Inc.**

**Key words:** conflict; anterior cingulate cortex; dorsolateral prefrontal cortex; neuroimaging; fMRI; Stroop

## INTRODUCTION

Throughout most of the past century of experimental psychology, inferences about cognitive function have relied upon the manipulation and measurement of some behavioral metric. For example, in the Stroop task one can measure the level of behavioral "conflict" elicited by mismatches between a color word and an incongruent ink color in which the word is printed (e.g., "RED" in blue ink) by measuring the time it takes to identify the ink color in these stimuli, compared to neutral (e.g., "HOUSE" in blue ink) or congruent (e.g., "RED" in red ink) stimuli. Subtle manipulations in the sorts of stimuli presented (e.g., colors, objects, words) induce changes in the behavioral metric (i.e., response time, accuracy), and researchers are able to make inferences about the way in which the human cognitive machinery resolves such conflict by examining the relationship between changes in the stimuli and changes in the behavioral metric.

Advances in neuroimaging methods such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) over the past decade have opened new windows into human cognitive function. These powerful tools have allowed researchers to identify the regions of cortex active during the performance of tasks such as the Stroop, and thus make inferences about the regions of cortex involved in cognitive processes like conflict resolution. This approach has consistently confirmed patient and lesion data linking regions of cortex to general cognitive functions (e.g.,

PFC and working memory, parietal areas and spatial processing, etc.), and has additionally provided new insight into more subtle differentiations in cortical functioning than is typically possible in human patient populations.

Along these lines, many neuroimaging studies have suggested functional roles for previously less well-studied regions such as the anterior cingulate cortex (ACC), found in the medial wall of the brain [e.g., Bush et al., 2000; Paus et al., 1993]. Initial studies of selective attention suggested that the dorsal ACC was responsible for the implementation of attentional control (e.g., the biasing of task-relevant representations over that of task-irrelevant representations in working memory) [Posner et al., 1988; Posner and Dehaene, 1994]. However, reports of variability in ACC activity during the Stroop and flanker tasks suggested a more limited involvement in attentional control. Specifically, this latter work suggested that the dorsal ACC was involved in response-related processes [Banich et al., 2000a; McDonald et al., 2000; Milham et al., 2001, 2003a, 2003b; Paus et al., 1993; Paus, 2001].

In addition, studies employing paradigms involving conflict, such as Stroop and flanker tasks, often report a positive correlation between increasing behavioral measures of conflict and the degree of ACC activity [Botvinick et al., 1999; Bush et al., 1998; Carter et al., 2000; Casey et al., 2000; Gruber et al., 2002], thereby supporting a conflict monitoring, detection, or resolution role for ACC [Barch et al., 2001; Carter et al., 1998; Carter et al., 2000; Cohen et al., 2000; MacDonald et al., 2000]. Furthermore, recent studies of the ACC [Gruber et al., 2002; Ruff et al., 2001; Van Veen et al., 2001] rely on positive correlations between behavioral interference and ACC activity to develop and extend current theories of ACC's role in conflict. In fact, ACC's involvement in conflict detection/resolution has been established by employing tasks in which behavioral interference is observed. Therefore, based on these observations of the relationship between behavioral measures of conflict and ACC, it is frequently predicted that ACC activity and behavioral measures of conflict parallel each other, so that a change in the magnitude of behavioral conflict parallels a change in the degree of ACC activity.

Given that the function of ACC has been related to behavioral interference in conflict paradigms, then evidence of a de-coupling of this relationship such as a disproportional change in one of the measures relative to the other, would be difficult to explain according to conflict monitoring hypotheses of ACC function. In fact, a disproportional relationship between ACC and behavioral measures of conflict may require a dramatic reconsideration of the role of ACC in conflict processes.

In support of a non-parallel, or disproportional, relationship, a number of studies incorporating Stroop and flanker tasks have failed to find ACC activity despite reporting significant behavioral interference effects [Bench et al., 1993; Hazeltine et al., 2002; Taylor et al., 1997]. One of these studies argued that methodological issues, such as stimulus presentation rate and interstimulus interval, might affect the ability to detect ACC activity [Bench et al., 1993]. Although these studies suggest that behavioral measures of conflict do not always covary with ACC activity, they do not provide confirmatory evidence for a non-parallel relationship between ACC and behavioral conflict. For example, if methodological differences in stimulus parameters are a cause for a failure to find ACC activity despite high levels of behavioral conflict [Bench et al., 1993], then a lack of ACC activity may be attributed to the methodological parameters rather than a property of the ACC per se.

In addition, a recent study by Milham et al. [2003a] examining the distinct roles of ACC and DLPFC in attentional control reported that with practice on a modified version of the Stroop task, participants experienced a significant reduction in ACC activity with a concomitant reduction in their degree of behavioral interference. Interestingly, although residual behavioral interference was found during the last two cycles of the task, Milham et al. [2003a] reported no activation in ACC activity during those same two cycles. Although both ACC activity and the degree of behavioral interference reduced with time, residual behavioral interference in the last two cycles with a lack of ACC activity supports the hypothesis that the two measures are not completely parallel.

In addition, recent studies suggest that the dorsal ACC is not involved in processing general conflict, but is primarily related to response-related processes [Milham et al., 2001, 2003a; Van Veen et al., 2001]. Therefore, conflict at other processing stages (e.g., semantic) may not result in heightened ACC activity. This is important because a disproportional relationship between ACC activity and behavioral measures of conflict may provide biological evidence for the processing stage at which the Stroop interference occurs.

There could be a number of possible explanations for the presence of a non-parallel relationship between ACC and behavioral conflict. For example, one explanation may be that the ACC's role in attentional control is limited to the initial detection of conflict under attentionally demanding conditions. According to this hypothesis, the ACC, although mainly involved in response-related aspects of attention, would appear to be sensitive to at least some non-response-related aspects of conflict but only during the initial stages of task performance. This hypothesis suggests that a new region becomes involved in later stages of the task [e.g., Raichle et al., 1994], or a region already active in task demands assumes responsibility for functions previously subsumed under the ACC [Milham et al., 2003a]. Another possibility is that better top-down control by DLPFC with practice reduces the degree of conflict at the response stage, which results in a reduction of ACC activity [Milham et al., 2003]. It is notable that an interplay between neural regions involved in attentional control might not be apparent by any observed changes in the behavioral metric. Furthermore, it would suggest the possibility of a dynamic interaction between neural regions involved in executive control [Cohen et al., 2000; Milham et al., 2003a].

Another related area of research has reported reductions in neural activity with increased experience on a task [Bush et al., 1998; Friston et al., 1992; Milham et al., 2003a; Paus et al., 1997; Peterson et al., 1998; Raichle et al., 1994; Van Mier et al, 1998]. These reductions are often complimented by increases in activity in other regions (e.g., greater parietal and temporal activity with less frontal activity) as well as complimentary changes in the behavioral metric [Bush et al., 1998; Raichle et al., 1994]. One way to investigate the effects of experience on neural activity is to divide the data into halves or thirds and examine the changes in the neural activity between the first, second, and/or third periods [e.g., Milham et al., 2003a]. This methodology provides a means by which the relationship between ACC and behavioral measures of conflict may be investigated.

Therefore, the current study was designed to investigate whether experience in selecting task-relevant information in the Stroop task disproportionately affects the degree of activity in the dorsal "cognitive" region of the ACC relative to the degree of behavioral measures of conflict. To investigate this question, we reanalyzed an existing and previously published data set that employed the color-word Stroop task to examine age-related changes in inhibitory functions [Milham et al., 2002]. The reanalysis consisted of only the young adult portion of the original sample and was implemented on the original time series data. Other neuroimaging studies have confirmed that reanalyses of existing data sets can result in informative answers to new questions [Peterson et al., 1998].

We performed two separate reanalyses on the neuroimaging data. The first reanalysis was conducted in order to investigate the coupling of behavioral measures of conflict and ACC activity by dividing the neuroimaging data in half and making comparisons between the degree of neural activity and behavioral interference between the halves. First, we verified that no differences existed in the degree of behavioral interference between the first and second halves of the task, thereby supporting previous research reporting that thousands of incongruent trials are necessary to reduce the Stroop interference effect [Dulaney and Rogers, 1994; MacLeod and Dunbar, 1988; Rogers and Fisk, 1991]. Therefore, we predicted that there would be a disproportional change in the degree of ACC activity relative to behavioral measures of conflict. This result would be disparate with current conceptions of ACC function with regards to its role in cognitive conflict.

Furthermore, this type of analysis (1st half vs. 2nd half) allows for an examination of temporal changes in the Blood Oxygen Level Dependent response (BOLD) in ACC within a time frame that is commonly used among other neuroimaging studies of attentional control. Importantly, dividing the task in half and making comparisons between the halves controls for methodological issues described by Bench et al. [1993] because both halves have exactly the same stimulus parameters. Therefore, any changes occurring in ACC activity could not be attributed to methodological issues such as differences in interstimulus interval or presentation rate.

If the first half and second half dramatically differ in their patterns of activity, then it is likely that a typical analysis of the neuroimaging data that included all experiment cycles would provide different results and interpretations compared to an analysis of only the first half. More specifically, the current experiment duration including all cycles was approximately 10 min. Therefore, if we treat the first half of the experiment from the first analysis as if it was a shorter but separate experiment (5 min) we could compare the results from a 10-min experiment (typical analysis of all cycles) with the results from a 5-min experiment (first half). If the pattern of activity dramatically differs between the different experiment durations, then it would suggest that the results and interpretations of a neuroimaging experiment are dependent upon the duration of the experiment. Moreover, this result would suggest that the lack of ACC activity in previous experiments might have been due to the duration of the study including practice on the task before data collection began. In order to examine this, we reanalyzed the entire data set in a typical fMRI analysis fashion that included all experiment cycles and compared it with the first half results obtained from the first analysis.

In summary, we investigated the relationship between ACC activity and behavioral indices of conflict. A non-parallel relationship would be incompatible with current conceptions of ACC's role in conflict monitoring. In addition, we performed a second analysis that examined how the duration of an experiment can affect the outcome and the interpretability of neuroimaging results.

## SUBJECTS AND METHODS

### Participants

Twelve right-handed, native English-speakers (7 men, 5 women; age range, 21–27 years, mean = 23 years) participated in the study. All participants were screened for a history of neurological insult and color-blindness and signed an informed consent prior to participation.

### Stimuli and Design

Previous behavioral research has determined that hundreds or thousands of incongruent trials are usually necessary in order to significantly reduce the amount of interference [Dulaney and Rogers, 1994; MacLeod and Dunbar, 1988; Rogers and Fisk, 1991]. Fewer than 100 incongruent trials were presented in the study published by Milham et al. [2002], and we verified that no significant changes in the degree of behavioral interference occurred. Therefore, variation in the BOLD response to conflicting stimuli throughout the performance of the task could be examined in the absence of a behavioral change in conflict. Specifically, an analysis and comparison between the first and second halves of the neuroimaging data should allow us to examine the parallel nature, or the lack thereof, of the ACC's activity and behavioral measures of conflict. Therefore, we felt that a reanalysis of the Milham et al. [2002] data set was well

suited for this investigation. In addition, the previously published data set was collected to address issues of attentional control in an aging population and, therefore, questions and hypotheses regarding the young adult sample were not thoroughly examined outside of an aging perspective. Therefore, a reanalysis using different statistical procedures could examine the relationship between the ACC and behavioral indices of conflict.

As described in Milham et al. [2002], the Stroop stimuli were programmed using Mel V2.0 and presented using an IBM-PC compatible computer. The Stroop task that was employed made use of three ink colors: red, orange, and green. Three trial types were included: (1) *incongruent trials* in which the word named a color incongruent with the ink color in which it was printed (e.g., the word "RED" in green ink); (2) *congruent trials* in which the word named the ink color in which it was printed (e.g., the word "RED" printed in red ink); (3) *neutral trials* in which the word was unrelated to color (e.g., the word "LOT" in red ink). The neutral words were matched with the color words for word frequency and length. Participants were told to respond to the ink color in which the word was printed. A three-button manual response box was used with each button corresponding to the three ink colors employed.

We made use of a blocked design, in which blocks of neutral trials alternated with blocks of incongruent and congruent trials (i.e., neutral-incongruent-neutral-congruent) such that participants were presented a total of 4 incongruent blocks, 4 congruent blocks, and 9 neutral blocks. Neutral trials were also intermixed within congruent and incongruent blocks (50:50 mix) to prevent the development of strategies regarding the degree to which the words are read (e.g., ignore them during the incongruent blocks, pay attention to them during the congruent blocks). Each block consisted of 18 trials at a rate of one trial every 2 sec. Each trial consisted of a 300-msec fixation cross, followed by a 1,200-msec presentation of the stimulus (colored word) and 500-msec pause. The order of incongruent and congruent block types was counterbalanced across participants to control for order effects.

## Procedure and Image Processing

Participants were placed in a 1.5 Tesla GE Signa scanner equipped for echo-planar imaging (EPI). Head position was stabilized using a bite-bar attached to the head coil to minimize motion during the session. The visual stimuli were presented using a goggle system by Magnetic Resonance Technologies. A total of 410 gradient echo EPI scans were acquired for each participant (TR = 1,517 msec, TE = 40 msec, flip angle = 90 degrees, 7-mm slice thickness, slices acquired interleaved). A field-of-view of 24 cm and data matrix size $64 \times 64$ results in a digital resolution of $3.75 \times 3.75$ mm. Each scan consisted of 15 contiguous slices parallel to the AC–PC line.

The first six volumes of each run were discarded to allow the MR signal to reach steady state. Functional data for each participant were then motion corrected, spatially convolved using a three-dimensional Gaussian kernel (FWHM = $8 \times 8 \times 8$ mm, kernel width = $7 \times 7$ mm), and intensity normalized. A matched bandpass filter was then applied with a low-pass cut-off of 2.8 and a high-pass cut-off of 216. MedX v. 3.2 was used for image processing and statistical analysis and MRIcro v. 1.35 was used for image visualization.

## Regression Analyses

### Analysis 1

We chose to analyze our data by first half vs. second half because others have had success with this approach and it provides a conservative estimate of temporal changes [Milham et al., 2003a]. The first regression analysis consisted of four regressors that investigated the functional changes occurring between the first half and the second half of task performance: (1) a *first half congruent regressor* (FHCR) to identify regions active for the congruent condition during the first half (first two cycles) of the task, (2) *a first half incongruent regressor* (FHIR) to identify regions active for the first half (first two cycles) of the incongruent condition, (3) a *second half congruent regressor* (SHCR) to identify regions active for the congruent condition during the last half (last two cycles) of the task, (4) a *second half incongruent regressor* (SHIR) to identify regions active for the second half (last two cycles) of the task for the incongruent condition. To maintain linear independence and orthogonality among the regressors, the lowest level condition (neutral) was excluded. Therefore, all analyses are compared to the neutral condition. The parameter estimates for each subject from each regressor were warped into a stereotaxic space and then tested against zero using a random effects single group *t*-test. T-scores were then converted to Z-scores and a cluster analysis was performed using a minimum cluster threshold of 15 contiguous voxels with a Z-score threshold of 2.58 [Forman et al., 1995].

In order to investigate whether any changes occurred in the pattern or degree of activity between the first half and second half of task performance for the incongruent condition, a direct comparison was conducted between the parameter estimates from the FHIR and the parameter estimates from the SHIR. In addition, to investigate the changes in the pattern of activity between the first half and second half of task performance for the congruent condition, a direct comparison was conducted between the FHCR parameter estimates and the SHCR parameter estimates. Both direct comparisons were performed using a random-effects paired *t*-test. In addition, following Milham et al. [2002], we also directly compared the FHIR vs. the FHCR as well as the SHIR vs. the SHCR in a random effects paired *t*-test. Based on Milham et al. [2002], these comparisons would allow us to examine the result of heightened conflict when two sources of color-information were present (i.e., color information contained in the word and color information contained in the ink color). The results from the direct comparisons were converted to Z-scores and a cluster analysis was performed using a minimum cluster threshold of 15 contig-

**TABLE I. Response times and accuracy rates for the first and second half of task performance**

| | First half | | | Second half | | |
|---|---|---|---|---|---|---|
| | Incong | Neut | Cong | Incong | Neut | Cong |
| Response time (msec) | 718 (85.7) | 600 (44.0) | 572 (46.2) | 736 (123.7) | 600 (36.4) | 585 (68.3) |
| Accuracy (% correct) | 0.88 (0.10) | 0.96 (0.04) | 0.97 (0.04) | 0.91 (0.10) | 0.97 (0.05) | 0.96 (0.09) |

Values in parentheses are standard deviations. Incong, incongruent condition; cong, congruent condition; neut, neutral condition.

uous voxels with a $Z$-score threshold of 2.58. Finally, centers of intensity within each of the clusters were identified [Mintun et al., 1989].

### Analysis 2

The second analysis investigated whether the inclusion of all cycles in a typical fMRI analysis would result in a different pattern of neural activity compared to the results from the first analysis, thereby leading to disparate conclusions and interpretations of the data. In order to examine this, we applied two regressors to all four cycles: (1) a *congruent regressor* (CR) to identify regions common to the congruent condition throughout all four cycles and (2) an *incongruent regressor* (IR) to identify regions unique to the incongruent condition throughout all four cycles. The parameter estimates from both regressors for each subject were then warped into a stereotaxic space [Talairach and Tournoux, 1988] and then tested against zero using a random effects single group $t$-test. T-scores were then converted to Z-scores and a cluster analysis was performed using a minimum cluster threshold of 15 contiguous voxels with a Z-score threshold of 2.58 [Forman et al., 1995].

## RESULTS

### Behavioral Results

The accuracy rates were coded as the percentage of correct responses. The accuracy rate for the incongruent condition (M = 0.89) was significantly poorer than the accuracy rate for the neutral condition (M = 0.96) (t(11) = 2.467; $P < 0.03$). The accuracy rates for the congruent and neutral stimuli did not differ (t(11) = 0.09; $P < 0.927$). A 3 (condition: incongruent, congruent, neutral) $\times$ 2 (half: 1st, 2nd) ANOVA was performed on the accuracy data. We found no significant changes in the accuracy rates between the first and second halves for either the incongruent condition F(1, 22) = 0.578, $P < 0.455$, congruent condition F(1, 22) = 0.245, $P < 0.626$, or neutral condition F(1, 22) = 0.085, $P < 0.773$. The mean accuracy rates for the first and second halves for all three trial types are presented in Table I.

Additionally, the response times for the incongruent condition (M = 727 msec) were significantly slower than the neutral condition (M = 600 msec) (t(11) = 5.62; $P < 0.000$) indicating a significant interference effect. Furthermore, there was a trend for the trials in the congruent (M = 579 msec) condition to be faster than the neutral condition (t(11)

= 2.09; $P < 0.060$) indicating a trend for a facilitation effect. A 3 (condition: incongruent, congruent, neutral) $\times$ 2 (half: 1st, 2nd) ANOVA was performed on the response time data. We found no significant changes in the response time rates between the first and second halves for either the incongruent condition F(1, 22) = 0.161, $P < 0.692$, congruent condition F(1, 22) = 0.281, $P < 0.602$, or neutral condition F(1, 22) = 0.000, $P < 0.997$. The mean response times for the first and second halves for all three trial types are presented in Table I.

Therefore, the degree of behavioral interference did not change as a function of increased experience on the task (from the first half to the second), thus providing support for previous research suggesting the need for experience on hundreds or thousands of incongruent trials to reliably reduce the degree of interference.

## Neuroimaging Results

### First analysis: comparison of first half vs. second half

The pattern of activity in the first half of the task for the incongruent condition resembled results from other neuroimaging studies of attentional control including observed activity in the ACC and right and left dorsolateral prefrontal cortex (DLPFC; inferior and middle frontal gyri) [e.g., Banich et al., 2000a]. We found a significant cluster of activation in the ACC (151 contiguous voxels [all voxels are 2 mm isotropic]; Z-score = 2.85) for the FHIR. In addition, we found a significant cluster of activity for the FHIR that covered the right inferior and middle frontal gyri (192 contiguous voxels; Z-score = 2.72) and another cluster in the left inferior and middle frontal gyri (563 contiguous voxels; Z-score = 2.91). Furthermore, we found a significant cluster of activity in the right middle frontal gyrus for the FHCR (1,620 contiguous voxels; Z-score = 2.98). Cluster sizes and centers of intensity for each of the clusters are reported in the first column of Table II.

The pattern of activity evident in the second half of the task differed dramatically from that seen in the first half of the task (see column 2 of Table II). The ACC activity evident in the first half of the task no longer exceeded a positive significance threshold, even at a more lenient level (Z-score threshold of 1.96; 0 voxel cluster threshold). In fact, the anterior cingulate showed a significant deactivation during the second half of the task (Z = −3.1) for the SHIR. In contrast, we found significant DLPFC activity for the SHIR. There were no significant clusters of activity for the SHCR.

**TABLE II. The number of voxels in the cluster for each region for the incongruent and congruent trials**

| Region | Cluster size | | | | |
| --- | --- | --- | --- | --- | --- |
| | First half of task | Second half of task | First > Second* | Second > First* | Aggregate (second analysis)** |
| Inconguent | | | | | |
| Anterior cingulate | 151 (2.85), 2, 22, 42 | NS | 279 (2.79), 2, 18, 40 | NS | NS |
| R inferior/middle frontal gyrus | 192 (2.72), 52, 22, 32 | 1231 (3.01), 46, 14, 46 | NS | 170 (2.85), 20, 44, 34 | 646 (2.87), 52, 18, 22 |
| L inferior/middle frontal gyrus | 563 (2.91), −42, 18, 32 | 787 (2.99), −44, 18, 32 | NS | 107 (3.12), −24, 44, 30 | 422 (3.04), −38, 18, 32 |
| Congruent | | | | | |
| R middle frontal gyrus | 1620 (2.98), 46, 18, 42 | NS | NS | NS | 97 (2.70), 48, 18, 28 |

Values are expressed as n (mean Z score), COI.
In order to convert the cluster sizes to mm³ multiply the number of voxels in the cluster by 8 (each $2 \times 2 \times 2$ mm isotropic voxel is 8 mm³).
Therefore, the ACC cluster for the first half converts from 151 contiguous voxels to 1208 mm³.
*Direct comparison in which the one half is greater than the other.
**Cluster size for the second analysis of all four cycles (i.e. the entire data set).
COI, center of interest in Talairach coordinates; NS, not significant.

Following Milham et al. [2002], we investigated the effect of heightened conflict in the presence of two sources of color information, in a comparison of incongruent vs. congruent trials. First, we found that the DLPFC activity evident for the incongruent trials during the first half failed to reach significance at a threshold of Z>2.58 in the direct comparisons between the FHIR and FHCR as well as the SHIR and SHCR. The activity in the ACC for the FHIR, however, was observed in the direct comparison between FHIR and FHCR (Z-score = 2.78; cluster size = 40 contiguous voxels), but the size of the cluster shrunk, suggesting that there was subthreshold activity in the ACC for the FHCR. In addition, the direct comparison between the SHIR vs. the SHCR revealed no activity in the ACC even at a more lenient threshold of Z>1.96, thereby supporting the results from the comparison of the FHIR vs. SHIR.

Importantly, although there was subthreshold activity for the FHCR, the direct comparison between the FHCR and SHCR revealed no significant differences even at a subthreshold level. Therefore, the temporal changes in ACC activity are specific to the incongruent condition and cannot be attributed to a non-specific time effect for all conditions.

Furthermore, a direct comparison of the neural activity in the first and second half of the task (FHIR vs. SHIR) revealed two regions of reliable change. Specifically, the response of the DLPFC (middle frontal gyrus) significantly increased bilaterally between the first and second half of the task (Table II, column 4). Conversely, as suggested above, activity in the ACC was reliably reduced, and even reversed, in the second half of the task (Table II, column 3), so that the direct comparison revealed a large region of ACC activity unique to the first half of task performance (Fig. 1) even though we found no reduction in the amount of behavioral conflict elicited by the incongruent stimuli between the first and second half of the task. This disjunction between changes in behavioral and neural measures of conflict will be addressed in greater detail in the Discussion.

### Second analysis: aggregate activity

The analysis of all four cycles for the IR is described in column 5 of Table II. We found that only the right and left DLPFC regions exceeded the statistical threshold in this analysis. However, even for the DLPFC, the clusters vary substantially in terms of peak magnitude and spatial extent in a comparison of the first analysis to the second analysis (see Table II). Moreover, the ACC was not identified as an active region when all time points were considered.

### Correlations

We performed a series of correlations on the neuroimaging data to investigate the relationship between ACC and DLPFC clusters. According to our predictions, since ACC decreases with experience on the task while DLPFC clusters increase, then this change should be evident by changes in the correlation coefficients between regions. These values are presented in Table III.

We found that for all comparisons, the relationship between the right and left DLPFC regions was positively correlated. In addition, for the FHIR, there was a positive correlation between the ACC and right and left DLPFC. However, the correlations between the ACC and DLPFC regions became non-significant by the second half (SHIR). Furthermore, the correlations for the second analysis between ACC and DLPFC regions were non-significant.

In addition, all correlation coefficients were transformed to Fisher's Z values to test for differences between correlations. All differences were non-significant at a $P < 0.05$ except for the difference between the FHIR and SHIR for the ACC and right DLPFC ($Z = 2.24$; $P < 0.05$).

### DISCUSSION

The results from the first analysis revealed that with experience in selecting task-relevant information over that of
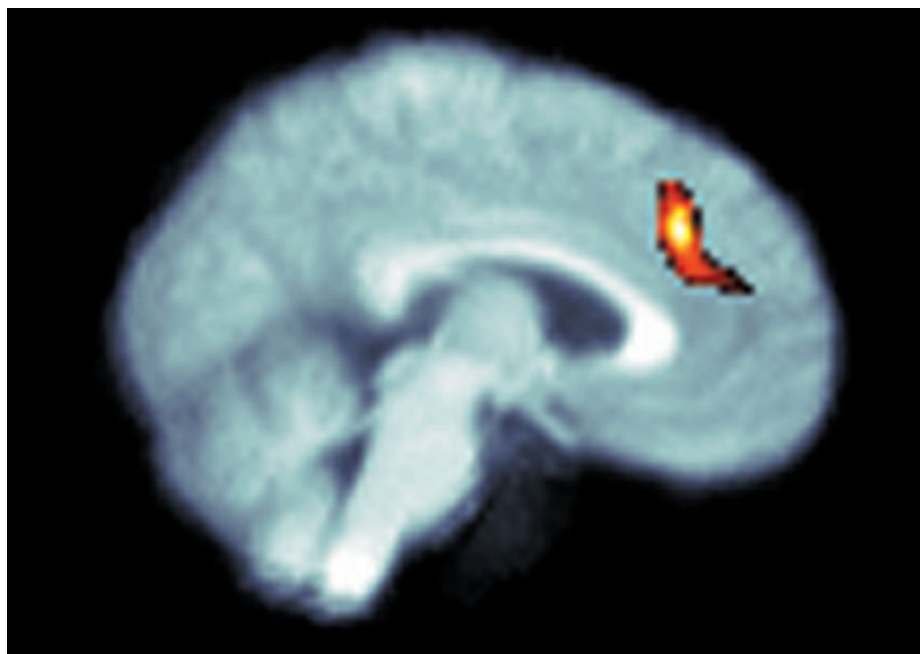
**Figure 1.**
Direct comparison between the first half incongruent regressor and second half incongruent regressor showing a significant cluster in anterior cingulate (279 contiguous voxels). The center of intensity is at 2, 18, 40.

task-irrelevant information in the Stroop task, the degree of activity in the dorsal ACC disproportionately changes in relation to the behavioral index of conflict. Most importantly, there were no apparent signs of any decreases in the behavioral indices of conflict even though there was no significant ACC activity by the second half.

Although others have reported a reduction in ACC activity with concomitant reductions in behavioral measures of Stroop interference [Bush et al., 1998; Milham et al., 2003a], our results report a reduction in ACC activity without any significant changes in the behavioral metric. In addition, our results compliment Milham et al. [2003a] who reported residual behavioral interference without ACC activity in the last cycles of the task. Therefore, our results, and the results from Milham et al. [2003a], suggest that ACC activity and behavioral indices of conflict are not completely parallel.

Although these results are inconsistent with conflict monitoring conceptions of ACC function, these results are in concordance with a hypothesis proposing that the activity of ACC is confined to the deployment of conflict processes during the early or initial period of task performance when information regarding the stimuli and task demands are unknown to participants. This would suggest that the ACC is more limited and transient in its role in attentional control and functions during the initial detection of conflict under attentionally demanding conditions. It may be that as another region (i.e., DLPFC) exerts more refined attentional control with practice in situations of conflict, less conflict needs to be resolved at the response stage, leading to decreased ACC activity. Although DLPFC may initially rely on ACC to detect response conflict, once engaged it is able to maintain its own involvement. This reduced engagement reduces the need for ACC [Milham et al., 2003a]. An increase in the size of the DLPFC clusters and the changing correla-

**TABLE III. Correlations between the ACC and DLPFC regions for all comparisons**

| | Correlation (r) | | |
|---|---|---|---|
| Regressor | ACC and RDLPFC | ACC and LDLPFC | RDLPFC and LDLPFC |
| FHCR | 0.31 | 0.14 | 0.73** |
| FHIR | 0.75** | 0.64* | 0.91** |
| SHCR | 0.10 | 0.06 | 0.65* |
| SHIR | −0.08 | 0.13 | 0.75** |
| CR | 0.19 | 0.03 | 0.68* |
| IR | 0.33 | 0.28 | 0.82** |

*$P < 0.01$
**$P < 0.005$.
RDLPFC, right dorsolateral prefrontal cortex; LDLPFC, left dorsolateral prefrontal cortex; ACC, anterior cingulate cortex; FHIC, first half incongruent regressor; FHCR, first half congruent regressor; SHIR, second half incongruent regressor; SHCR, second half congruent regressor; IR, incongruent regressor for second analysis; CR, congruent regressor for second analysis.

tions between DLPFC and ACC from the first half to the second half provides supporting evidence for this claim.

Additionally, we found that the increase in the size of DLPFC clusters in the second half appears to involve an extension of activity from posterior prefrontal regions to anterior prefrontal regions. Importantly, however, the DLPFC region that was active for the first half remained active (evident by the lack of any significant effects in DLPFC for the direct comparison between the first and second half). So, the increase in the size of the cluster was due to an increase in activation in a new adjoining region, anterior to the first half DLPFC cluster. This increase in

activation in a new region of DLPFC may suggest an underlying change in the mechanism by which the interference is resolved or the development of a new mechanism to handle conflict.

It is also possible that this initial activity in ACC may be related to either an anticipatory response [Murtha et al., 1998] or learning [Gabriel et al., 2002]. For example, ACC activity in the first half may be related to a greater need for attentional resources devoted to developing and learning task-related associations early in performance, especially at the response level (e.g., learning the association between buttons, color mappings, while concurrently inhibiting word reading) than is necessary later in the task. This hypothesis is supported by animal research reporting that the ACC is involved in associative learning [Gabriel et al., 2002]. More specifically, there are associations to be learned for all Stroop stimuli including the button mappings and learning to select one dimension of a multi-dimensional stimulus.

Another important point to consider is the possibility that our results provide important information regarding the level at which the behavioral interference effect in the Stroop task occurs. For example, our results may suggest that the behavioral conflict observed in the Stroop task occurs earlier in processing than the response stage since ACC activity has been recently related to the response level of conflict processing [Milham et al., 2001; Van Veen et al., 2001].The initial activity observed in ACC may be related specifically to response conflict when the button mappings for the color responses need to be learned and inhibited [Milham et al., 2001]. Once learned, the degree of response conflict is reduced, thereby reducing the necessity for ACC. However, since behavioral measures of conflict remain heightened without ACC involvement, the behavioral conflict observed in the Stroop task may be occurring at an earlier stage of processing (e.g., semantic) than the response stage, so that a reduction in the degree of ACC activity has no effect on the degree of behavioral conflict observed. The idea that interference occurs prior to a response is consistent with previous research indicating that activity in posterior brain regions is associated with processing the task-irrelevant dimension on incongruent as compared to neutral trials. For example, increased activity has been observed in word processing areas in a color-word Stroop task and object processing areas in a color-object Stroop task [Banich et al., 2000b, 2001].

Finally, it is possible that the brain measures are more sensitive to practice-related changes than the behavioral measures. However, if this is the case it does not detract from our finding that the behavioral and neuroimaging measures do not have a one-to-one correspondence with each other. Instead, it suggests that converging measures need to be employed in order to fully examine cognitive phenomena.

It is common in neuroscience to assign a function to a brain region based on some behavioral measure. This method has proven fruitful in developing maps of motor, visual, and sensorimotor cortices. A similar method is commonly utilized by cognitive neuroscientists to determine the brain regions involved in more complex processes and constructs. In this way, behavioral indices such as response times and accuracy rates are used to assign a cognitive process or mechanism to a particular brain region. However, if a behavioral metric is used to determine the function of a particular brain region, but the fluctuation of activity in that region does not correspond to any fluctuations in the behavioral metric, then the function that was first assigned to that region based on the behavioral metric is brought to question. This issue lies at the heart of epistemological problems in cognitive neuroscience and reemphasizes the need for a careful evaluation of the assumption of a direct relationship between behavior and the function of any brain region.

The second analysis revealed that our 10-min experiment would have resulted in dramatically different results and interpretations if it had been only 5 min or equivalent to the first half. Specifically, the number and size of the clusters dramatically changed between the first half of the first analysis (5-min experiment) and the typical MRI results of the second analysis (10-min experiment).

There are at least two implications of this result. First, the duration of an experiment can dramatically alter the outcome as well as the interpretations of the results. Since the total experiment length was only 10 min, we were working within a time frame similar to that of other neuroimaging experiments. The analysis of the 10-min experiment failed to reveal any ACC activity, similar to other studies such as Bench et al. [1993] and Taylor et al. [1997]. Therefore, a more in-depth investigation into the temporal effects of the BOLD response in neuroimaging studies and its effects on the interpretations granted to the results, is warranted.

A second implication is that even small amounts of practice prior to collection of data in the task could dramatically affect the results and interpretations of the neuroimaging data. Bench et al. [1993], Taylor et al. [1997], Hazeltine et al. [2001], Ruff et al. [2001], and Van Veen et al. [2001] reported an unspecified amount of practice before data collection on the Stroop and flanker tasks. Therefore, practice on the task, before data collection, may have resulted in a different pattern of activity, such as no ACC, or dramatically less ACC, associated with the task compared to studies that do not administer a practice session before data collection. Furthermore, studies investigating conflict may have found robust enough activity during the early portions of task performance that a statistical effect was found in ACC despite dramatic drops in the degree of activity during later stages of the task. Although speculative, it is one possible explanation why some studies of conflict report ACC activity while others fail to find ACC activity.

So, what is the optimum length of a neuroimaging experiment? This question becomes even more difficult, but more important, when considering that dramatic changes can occur in the neuroimaging data without concomitant changes behaviorally. How are researchers able to make valid claims about the neural components involved in cognitive processes if the pattern of activity is dependent on the duration

and independent of the behavioral results? Is there such a thing as an "optimum duration"? Future research and theory should help to establish answers to these and other similar questions.

Another set of issues that are not addressed by the current analysis is whether our results are an effect of a blocked design and whether our effect is generalizable to other conflict paradigms besides Stroop. Preliminary data in our laboratory suggest that the effect is generalizable to both other paradigms as well as slow event-related designs. Specifically, we are employing a slow event-related modified version of the flanker task and finding significant decreases in ACC activity within approximately a 5-min period with no concomitant decrease in the degree of behavioral conflict (Colcombe et al., unpublished data). Another related question is whether these effects generalize to other cognitive paradigms such as working memory, episodic memory, and spatial processing. Future research should examine these possibilities.

In summary, the results from our first reanalysis indicate that with increased experience on the Stroop task, the degree of ACC activity reduces, whereas indices of behavioral conflict remain unchanged. This result is disparate with current conceptions of ACC function in conflict detection and, therefore, has important implications for the role that ACC is thought to play in attentional control. We propose that the ACC may be involved in the initial stages of conflict resolution or monitoring and that DLPFC regions may increase control at later stages of performance. Another possibility is that the ACC is not involved in conflict per se, but rather is involved in learning how to manage conflicting situations or simply learning to associate and select task-relevant dimensions of the stimuli. Importantly, our result reemphasizes the need for a careful evaluation of the assumption of a strong relationship between behavior and brain function, which lies at the heart of the epistemology of cognitive neuroscience. Our second analysis suggests that the duration of a study can largely affect the outcome and interpretations of neuroimaging data, including the roles that particular regions are thought to play in cognition. This suggests an alternative explanation to that of Bench et al. [1993] for why some studies of conflict resolution and monitoring have failed to find ACC activity and questions the interpretability of results that vary in duration while also being independent of behavioral changes.

## ACKNOWLEDGMENTS

## REFERENCES

Banich M, Milham M, Cohen NJ, Wszalek T, Kramer A, Liang Z-P, Gullet D, Shah C, Brown C (2000a): Prefrontal regions play a predominant role in imposing an attentional "set": evidence from fMRI. Cogn Brain Res 10:1–9.

Banich MT, Milham MP, Atchley RA, Cohen NJ, Webb A, Wszalek T, Kramer AF, Liang Z-P, Wright A, Shenker J, Magin R, Barad V, Gullett D, Shah C, Brown C (2000b): fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. J Cogn Neurosci 12:988–1000,

Banich MT, Milham MP, Jacobson BL, Webb A, Wszalek T, Cohen NJ, Kramer AF (2001): Attentional selection and the processing of task-irrelevant information: Insights from fMRI examinations of the Stroop task. In CasanovaC, Ptito M, editors. Vision: from neurons to cognition. Progress in brain research, Vol. 134. Amsterdam: Elsevier Science. p 459–470.

Barch DM, Braver TS, Akbudak E, Conturo T, Ollinger J, Snyder A (2001): Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. Cereb Cortex 11:837–848.

Bench CJ, Frith CD, Grasby PM, Friston KJ, Paulesu E, Frackowiak RSJ, Dolan RJ (1993): Investigations of the functional anatomy of attention using the Stroop test. Neuropsychologia 31:907–922.

Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD (1999): Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402:179–181.

Bush G, Whalen PJ, Rosen BR, Jenike MA, McInerney SC, Rauch SL (1998): The counting Stroop: an interference task specialized for functional neuroimaging: validation study with functional MRI. Hum Brain Mapp 6:270–282.

Bush G, Luu P, Posner MI (2000): Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215–222.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998): Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280:747–749.

Carter C, Macdonald AM, Botvinick M, Ross LL, Stenger VA, Noll D, Cohen JD (2000): Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. Proc Natl Acad Sci 97:1944–1948.

Casey BJ, Thomas KM, Welsh TF, Badgaiyan RD, Eccard CH, Jennings JR, Crone EA (2000): Dissociation or response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. Proc Natl Acad.Sci USA 97:8728–8733.

Cohen JD, Botvinick M, Carter CS (2000): Anterior Cingulate and prefrontal cortex: who's in control? Nature-Neurosci 3:421–423.

Dulaney CL, Rogers WA (1994): Mechanisms underlying reduction in Stroop interference with practice for young and old adults. J Exp Psychol 20:470–484.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. Magnet Reson Med 33:636–647.

Friston KJ, Frith CD, Passingham RE, Liddle PF, Frackowiak RSJ (1992): Motor practice and neurophysiological adaptation in the cerebellum: A positron tomographic study. Proc R Soc Lond [Biol] 248:223–228.

Gabriel M, Burhans L, Talk A, Scalf P (2002): Cingulate cortex. Encyclopedia of the human brain, Vol. 1. New York: Elsevier Science. p 775–791.

Gruber S, Rogowska J, Holcomb P, Soraci S, Yurgelun-Todd D (2002): Stroop performance in normal control subjects-An fMRI study. Neuroimage 16:349–360.

Hazeltine E, Poldrack R, Gabrieli JDE (2000): Neural activation during response competition. J Cogn Neurosci 12(Suppl 2):118–129.

MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS (2000): Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835–1838.

MacLeod CM, Dunbar K (1988): Training and Stroop-like interference: Evidence for a continuum of automaticity. J Exp Psychol 10:304–315.

Milham MP, Banich MT, Webb A, Barad V, Cohen NJ, Wszalek T, Kramer AF (2001): The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. Cogn Brain Res 12:467–473.

Milham MP, Erickson KI, Banich MT, Kramer AF, Webb A, Wszalek T, Cohen NJ (2002): Attentional control in the aging brain: Insights from an fMRI study of the stroop task. Brain Cogn 49:277–296.

Milham MP, Banich MT, Claus E, Cohen NJ (2003a): Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. Neuroimage 18:483–493.

Milham MP, Banich MT, Barad V (2003b): Competition for priority in processing increases prefrontal cortex's involvement in top-down control: an event-related fMRI study of the Stroop task. Cogn Brain Res 17:212–222.

Mintun M, Fox P, Raichle M (1989): Highly accurate method of localizing regions of neuronal activation in the human brain with positron emission tomography. J Cereb Blood Flow Metab 1:96–103.

Murtha S, Chertkow H, Beauregard M, Dixon R, Evans A (1996): Anticipation causes increased blood flow to the anterior cingulate cortex. Hum Brain Mapp 4:103–112.

Paus T (2001): Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat Rev Neurosci 2:417–424.

Paus T, Petrides M, Evans C, Meyer E (1993): Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J Neurophysiol 70:453–469.

Paus T, Zatorre RJ, Hofle N, Caramanos Z, Gotman J, Petrides M, Evans AC (1997): Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. J Cogn Neurosci 9:392–408.

Peterson SE, Hanneke VM, Fiez JA, Raichle ME (1998): The effects of practice on the functional anatomy of task performance. Proc Natl Acad Sci USA 95:853–860.

Posner MI, Dehaene S (1994): Attentional networks. Trends Neurosci 17:75–79.

Posner M, Petersen S, Fox P, Raichle ME (1988): Localization of cognitive operations in the human brain. Science 240:1627–1631.

Raichle ME, Fiez JA, Videen TO, MacLeod AK, Pardo JV, Fox PT, Peterson SE (1994): Practice-related changes in human brain functional anatomy during nonmotor learning. Cereb Cortex 4:8–26.

Rogers WA, Fisk AD (1991): Age-related differences in the maintenance and modification automatic processes: arithmetic Stroop interference. Human Factors 33:45–56.

Ruff CC, Woodward TS, Laurens KR, Liddle PF (2001): The role of anterior cingulate cortex in conflict processing: evidence from reverse Stroop interference. Neuroimage 14:1150–1158.

Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain: 3-D proportional system: an approach to cerebral imaging. Stuttgart: Thieme.

Taylor S, Kornblum S, Lauber E (1997): Isolation of specific interference processing in the Stroop task: PET activation studies. Neuroimage 6:81–92.

Van Veen V, Cohen JD, Botvinick MM, Stenger VA, Carter CS (2001): Anterior cingulate cortex, conflict monitoring, and levels of processing. Neuroimage 14:1302–1308.

Van Mier H, Tempel LW, Perlmutter JS, Raichle ME, Peterson SE (1998): Changes in brain activity during motor learning measured with PET: effects of hand on performance and practice. J Neurophysiol 80:2177–2199.