NeuroImage xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood

Jack Grinband [a,*], Judith Savitsky [b], Tor D. Wager [c], Tobias Teichert [d], Vincent P. Ferrera [d], Joy Hirsch [b]

[a] Department of Radiology & Neuroscience, Columbia University, New York, NY 10032, USA
[b] Program for Imaging & Cognitive Sciences (PICS), Columbia University, New York, NY 10032, USA
[c] Department of Psychology, Muenzinger D261D, University of Colorado, Boulder, CO 80309, USA
[d] Department of Neuroscience, Columbia University, New York, NY 10032, USA

ARTICLE INFO

Article history:
Received 30 September 2010
Revised 23 November 2010
Accepted 6 December 2010
Available online xxxx

ABSTRACT

The dorsal medial frontal cortex (dMFC) is highly active during choice behavior. Though many models have been proposed to explain dMFC function, the *conflict monitoring model* is the most influential. It posits that dMFC is primarily involved in detecting interference between competing responses thus signaling the need for control. It accurately predicts increased neural activity and response time (RT) for incompatible (high-interference) vs. compatible (low-interference) decisions. However, it has been shown that neural activity can increase with time on task, even when no decisions are made. Thus, the greater dMFC activity on incompatible trials may stem from longer RTs rather than response conflict. This study shows that (1) the conflict monitoring model fails to predict the relationship between error likelihood and RT, and (2) the dMFC activity is not sensitive to congruency, error likelihood, or response conflict, but is monotonically related to time on task.

Introduction

The dorsal medial frontal cortex (dMFC), including the dorsal anterior cingulate cortex and the supplementary motor area, has been central to neural models of decision making (Mansouri et al., 2009). It has been proposed that its main role is the detection of internal response conflict during choice behavior (Botvinick et al., 1999). Though functional imaging studies have provided strong evidence in favor of conflict monitoring (Mansouri et al., 2009; Nachev et al., 2008), electrophysiology and lesion studies have been unable to provide supporting data (Ito et al., 2003; Nakamura et al., 2005). A key finding in conflict monitoring studies is that decisions involving high interference from multiple stimulus–response representations generate longer mean response latencies than decisions with low interference (Carter et al., 1998). However, recent data have suggested that the duration of a subject's decision process, or time on task, can have large effects on the size of the elicited hemodynamic response, independent of the nature of the decision (Grinband et al., 2008). Thus, it is unclear whether the activity in dMFC reflects the amount of response conflict or the longer processing time needed to choose the correct response. Our goal was to dissociate stimulus–response compatibility and error likelihood, two indicators for the presence of conflict, from RT, an indicator of time on task, and thus determine if dMFC activity is consistent with predictions of the conflict monitoring model.

The *conflict monitoring model* proposes that response conflict is the simultaneous activation of neuronal assemblies associated with incompatible behavioral responses (Botvinick et al., 2001; Brown and Braver, 2005) and that the dMFC detects changes in response conflict which require reallocation of attentional resources (Kerns et al., 2004). Functional imaging studies using the Stroop task (Botvinick et al., 1999; Carter et al., 1998; MacLeod and MacDonald, 2000), Ericksen flanker task (Kerns et al., 2004), go/no-go task (Brown and Braver, 2005), and other tasks that require cognitive control (Nee et al., 2007) have shown that activity in the dMFC increases as a function of response conflict. Because response conflict produces a cost in terms of the speed and accuracy of decisions, mean response time (RT) and error likelihood have been used as measures of conflict intensity (Botvinick et al., 2001; Carter et al., 1998). dMFC activity correlated with these variables has been interpreted as real-time monitoring for the presence of response conflict (Botvinick et al., 2001). A related model holds that error likelihood and conflict are dissociable, and suggests that the dMFC detects interference-related changes in error likelihood (Brown and Braver, 2005). Both models propose that the detected signal is sent to other brain regions (e.g. the dorsolateral prefrontal cortex) to regulate levels of cognitive control (Brown and Braver, 2005; Kerns et al., 2004).

However, some data has been difficult to incorporate into the conflict monitoring framework. Both permanent (di Pellegrino et al.,

* Corresponding author.
E-mail addresses: jackgrinband@gmail.com (J. Grinband), judysavit@gmail.com (J. Savitsky), tor.wager@colorado.edu (T.D. Wager), tt2288@columbia.edu (T. Teichert), vpf3@columbia.edu (V.P. Ferrera), joyhirsch@yahoo.com (J. Hirsch).

1053-8119/$ – see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.neuroimage.2010.12.027

2007; Mansouri et al., 2007; Pardo et al., 1990; Swick and Jovanovic, 2002; Turken and Swick, 1999; Vendrell et al., 1995) and temporary (Hayward et al., 2004) lesions of the dMFC produce minimal changes in performance during decisions involving response conflict. Furthermore, dMFC activity is present on most decision-making tasks (Ridderinkhof et al., 2004; Wager et al., 2004, 2009), even in the absence of response conflict (Bush et al., 2002; Milham and Banich, 2005; Roelofs et al., 2006), and has sometimes been shown to be unable to detect the presence of response conflict (Zhu et al., 2010). Electrophysiological studies in monkeys have found few dMFC neurons involved in conflict monitoring (Ito et al., 2003; Nakamura et al., 2005), and targeted dMFC lesions do not affect conflict-related increases in response time and error likelihood (Mansouri et al., 2009). These data present significant challenges for the conflict monitoring and related models.

Alternatively, dMFC activity may be unrelated to the detection of response conflict but instead may reflect non-specific sensory, attentional, working memory, and/or motor planning processes that are present for all decisions and that do not vary as a function of response conflict. In fact, neurons in the dMFC are strongly affected by spatial attention (Olson, 2003) and oculomotor control (Hayden and Platt, 2010; Schall, 1991; Stuphorn et al., 2010). Furthermore, a large percentage of these neurons show conflict-independent activity that begins at stimulus onset and terminates at the time of response execution (Ito et al., 2003; Nachev et al., 2008; Nakamura et al., 2005). Finally, imaging studies have shown that dMFC activity is common for most tasks that require attention (Wager et al., 2004) or working memory (Wager and Smith, 2003), and that it scales with RT in a wide range of conflict-free tasks (Grinband et al., 2006; Naito et al., 2000; Yarkoni et al., 2009).

The conflict monitoring model describes the dMFC as a region functionally specialized for the detection of interference between alternative responses, and thus predicts that high interference will generate greater neural activity per unit time. If, however, the dMFC reflects non-specific or conflict-independent processes such as spatial attention, then neural activity should scale with time on task or RT. Because high interference is associated with longer RTs, both interpretations predict larger BOLD responses on decisions with response conflict. However, they predict very different relationships between RT and MFC activity per unit time.

In the conflict monitoring model, the conflict detector receives input from neurons representing the mutually exclusive responses (Botvinick et al., 2001). Three variations of this model are consistent with the classic neuroimaging result suggesting dMFC involvement in conflict monitoring. In the first variant (Fig. 1A), the detector has a low firing threshold. It can detect input from a single active response neuron (congruent trials) or from multiple response neurons (incongruent trials) and will continue to fire as long as at least one response neuron is active. When detector activity is convolved with a hemodynamic response function, a monotonically increasing relationship between the BOLD signal and response duration is produced. Because activity per unit time in the detector is greater on incongruent trials, the slope of the BOLD vs. RT function is also greater. In the second variant (Fig. 1B), the detector is characterized by high activation thresholds and activity from a single response neuron (congruent trials) is unable to activate it. This results in a BOLD vs. RT function with zero slope. When both response neurons are active (incongruent trials) the detector continues to fire for the full response duration, resulting in a monotonically increasing relationship between BOLD signal and RT. A third variant of the conflict monitoring model (Fig. 1C) has high detection thresholds similar to variant 1B but with autoinhibitory connections or inhibitory feedback from other neurons that produce a refractory period on incongruent trials. This model is insensitive to firing duration of the response neurons, and is thus, a binary detector for the presence of conflict. In contrast, the dMFC may be insensitive to response conflict but still produce a larger BOLD response on incongruent trials: if RTs for the incongruent trials are, on average, longer than for the congruent trials, then the hemodynamic response will integrate the neuronal activity over a longer time-period to produce a larger response. In this case, the BOLD response will grow with RT but the BOLD vs RT functions will be identical for congruent and incongruent trials (Fig. 1D).

Results

To test these alternatives, normal subjects were scanned while performing a manual Stroop task (Stroop, 1935). In this task, subjects must name the ink color of the presented letters while ignoring the word spelled out by the letters. On congruent trials, the color of the ink matched the word (e.g. the word "red" written in red ink), whereas on incongruent trials, the color of the ink did not match the word (e.g. the word "red" written in green ink). Incongruent trials produced a state of high cognitive interference as indicated by higher mean error rates and higher median RTs across subjects (congruent error rate = 2.7%, s.d. = 2.7%; incongruent error rate = 4.8%, s.d. = 1.0%; paired t-test, $p = 0.022$, df = 22; congruent RT = 831 ms, s.d. = 104 ms; incongruent RT = 958 ms, s.d. = 133 ms; paired t-test, $p = 2 \times 10^{-8}$, df = 22; see Fig. S1 for RT distributions).

Standard fMRI multiple regression techniques were used to replicate previous results from the conflict detection literature (Botvinick et al., 1999; Carter et al., 1998; Kerns et al., 2004) showing increased activity in dMFC during incongruent, as compared to congruent, trials (Fig. 2A). To test whether differences in RT alone would produce a similar activation pattern, congruent trials with slow RTs (greater than the median, mean of subgroup = 1119 ms) were compared to congruent trials with fast RTs (less than the median, mean of subgroup = 711 ms). Slow RT trials produced greater activity in the dMFC (Fig. 2B) even when there was no difference in congruency. However, a lack of incompatible features may not necessarily eliminate interference; thus we confirmed that slow and fast congruent trials have equally low levels of conflict by measuring error likelihood, which, according to the conflict monitoring model, is proportional to conflict (Fig. S2). No significant difference in error likelihood existed between slow and fast trials (fast error: mean = 2.9%, s.d. = 3.8%; slow error: mean = 2.2%, s.d. = 2.0% paired t-test, $p = 0.41$, df = 22) confirming low conflict, independent of response duration. To further test whether the dMFC can be activated in the absence of response conflict, subjects were asked to view a flashing checkerboard and press a button when the stimulus disappeared. Since no choice decision was required and since only one response was possible, no response conflict could exist; nevertheless, activity in dMFC was proportional to time on task (Fig. S3).

These results demonstrate that response duration can affect dMFC activation, even in the absence of competing responses. But is RT a more powerful predictor of dMFC activity than response conflict? To test this, fast RT incongruent trials (mean of subgroup = 783 ms) were compared against slow RT congruent trials (mean of subgroup = 1190 ms). If response conflict drives the dMFC response, more activity should exist on fast RT incongruent trials due to interference from competing responses. On the contrary, dMFC activity was greater on slow congruent than fast incongruent trials (Fig. 2C), even though error rates were higher on fast incongruent trials (fast incongruent error = 4.4%, slow cong error = 2.2% paired t-test, $p = 0.033$, df = 22). Thus, dMFC activation tracks response duration, rather than the presence of incompatible stimulus–response features or increases in error likelihood.

Standard fMRI analysis methods rely on the general linear model, which makes assumptions about the intensity and duration of the underlying neuronal activity as well as the shape of the hemodynamic response function. Because incorrect assumptions can lead to invalid conclusions, the data was reanalyzed using event-related averaging, a

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

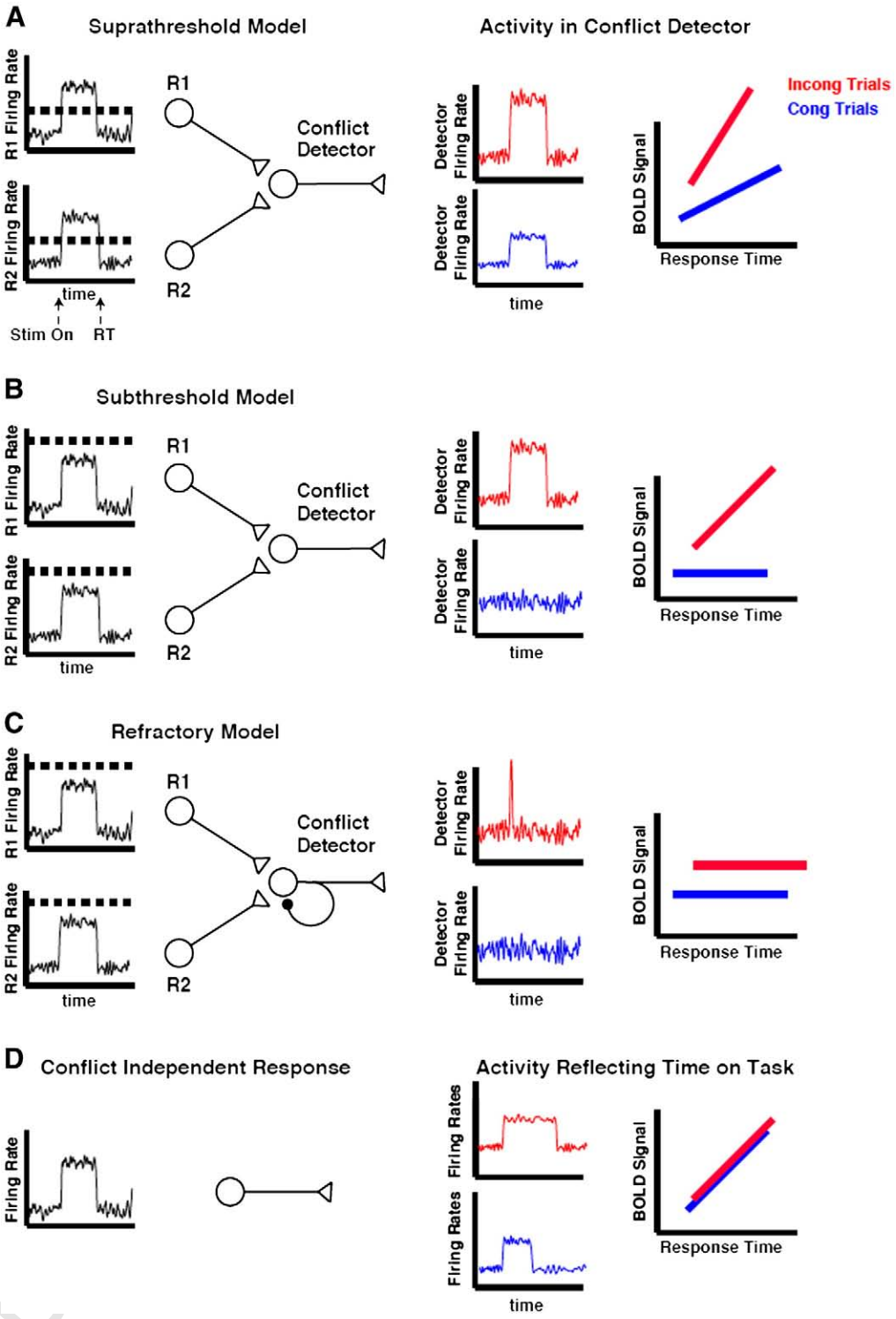

Fig. 1. Model predictions. Both the conflict monitoring and time on task accounts predict larger mean BOLD activity on incongruent trials. (A) In the "suprathreshold" model, activity in the response neurons (R1 and R2) activate the detector when threshold (horizontal dashed line) is exceeded. Thus, conflict can be present on both congruent (blue) and incongruent (red) trials, and depends on the firing duration of the response neurons. Because the input per unit time to the detector is greater on incongruent than congruent trials, the BOLD vs. RT functions have different slopes. (B) In the "subthreshold" model, activity from a single response neuron is not sufficient to activate the detector. Thus, the BOLD signal does not vary with duration of the response neuron on congruent trials. Activity from both response neurons is necessary to exceed threshold and cause a conflict-related response, which varies with response duration. (C) In the "refractory" model, activity from both response neurons is necessary to reach the detector's activation threshold. However, a refractory period created by autoinhibitory connections or inhibitory feedback from other neurons allows only a brief pulse of activity in the presence of conflict, resulting in activity that is independent of response duration (i.e. BOLD vs. RT functions with zero slope). (D) If activity in the MFC is determined only by time on task, then the BOLD vs. RT functions should be identical for congruent and incongruent trials. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

model-free analysis. When averaged across all RTs, voxels in the dMFC showed larger BOLD responses on incongruent than congruent trials (Fig. 3A), confirming the standard regression analysis result (Fig. 2A). We then tested for a relationship between dMFC activity and conflict in the absence of RT differences by comparing BOLD responses for trials with RTs within 100 ms of the global median. By comparing trials near

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

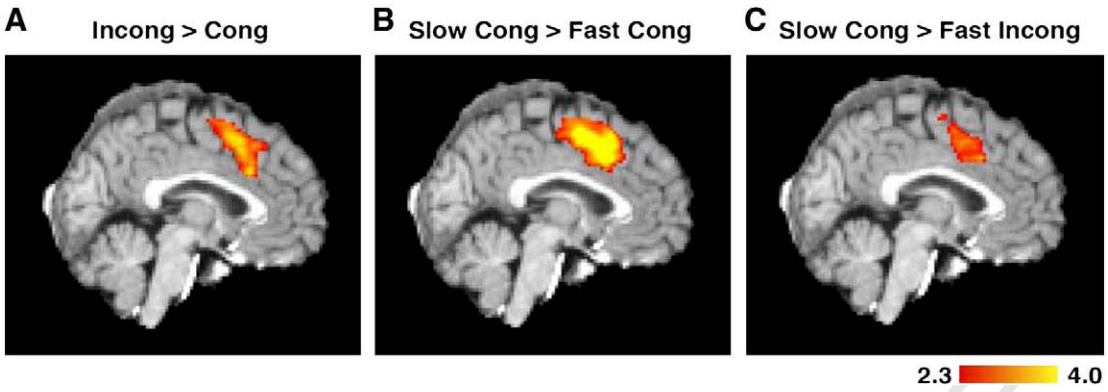

Fig. 2. Statistical parametric mapping. (A) Traditional GLM analysis comparing incongruent and congruent trials replicates previous results (Botvinick et al., 1999, 2001; Carter et al., 1998; Kerns et al., 2004) (peak activity = MNI152: 0/16/42). The activity was generated using only correct trials. (B) Congruent trials do not contain any incongruent features that could produce response interference. However, a comparison of slow vs fast congruent trials shows a pattern of activation in dMFC (MNI152: -4/16/40) that is similar to the "high conflict" pattern, indicating that dMFC activity is not specific to conflict. Fast and slow trials were defined as trials with RTs less than or greater than the median RT, respectively. (C) Slow congruent trials generate more activation than fast incongruent trials in dMFC (MNI152: -4/16/36), demonstrating that response time can better account for dMFC activation than the degree of response conflict. All activity is represented in Z-scores.

the median, the time needed to reach decision threshold is held roughly constant (mean RT of cong trials near median = 884 ms, s.d. = 117 ms, mean RT of incongruent trials near median = 891 ms, s.d. = 114 ms, paired *t*-test, $p = 0.856$, df = 44) and the resulting decisions vary only by the presence or absence of incompatible stimulus features. No significant differences in dMFC activity between congruent and incongruent trials were present when RT was held constant (Fig. 3B). Finally, a comparison of fast incongruent and slow congruent trials resulted in larger BOLD responses for the congruent condition (Fig. 3C; though error rates were significantly higher on fast incongruent trials — see earlier discussion). Again, BOLD activity was related to response duration, not conflict; taken together these data are inconsistent with the conflict monitoring model.

Computational models of conflict monitoring argue that both RT and error likelihood are determined by the degree of conflict in the decision (Botvinick et al., 2001; Siegle et al., 2004). The stronger the activation of the incorrect response, the greater is its interference with the correct response, leading to more errors and slower RTs. Moreover, for any given set of initialization parameters, the model produces a one-to-one relationship between the three variables (Fig. S2). Thus, the model predicts that error likelihood and RT can both be used as measures of conflict, and more importantly, that this relationship depends only on the input to the stimulus layer (i.e. the activation of the color and word units). If the three variables were not one-to-one, that is, if the relationship between error likelihood, RT, and conflict changed with context (e.g. congruency), this would provide evidence against the model.

We tested whether the relationship between error likelihood and RT remained constant (one-to-one). This was done by splitting each subject's RT distribution into ten equal quantiles. Error rates for congruent and incongruent trials were compared within each quantile to determine the degree of response conflict for each condition. A plot of error likelihood as a function of RT (Fig. 4A) shows that incongruent trials generate higher error rates than congruent trials for each RT (paired *t*-test of congruent quantiles vs. incongruent quantiles, $p = 0.033$, df = 9). In addition, incongruent trials have higher error likelihood than congruent trials at the majority of RT quantiles (Fig. 4B; paired *t*-test of congruent vs. incongruent trials within each quantile $p < 0.05$, df = 22). These data demonstrate that the frequency of conflict-induced errors is not uniquely related to RT as predicted by the conflict monitoring model (Fig. S2).

An analysis of the BOLD activity in dMFC demonstrated a monotonic increase as a function of trial-to-trial RT (Fig. 4C). No significant difference between congruency conditions (paired *t*-test of quantiles, $p = 0.90$, df = 9) was present. A comparison of congruent and incongruent trials within each quantile showed a significant difference only at a single point, consistent with a false positive rate of 0.05 (Fig. 4D paired *t*-test, $p < 0.05$, df = 22; mean difference between conditions = −0.0049 equivalent to a signal change of less than 10^{-5}%). Furthermore, if dMFC was a conflict detector, then dMFC activity should be proportional to the amount of conflict, as measured by error likelihood, at each RT. However, after controlling for RT, there was no relationship between BOLD activity and error likelihood: the difference in error likelihood (Fig. 4B) and the difference in BOLD activity (Fig. 4D) were not correlated (Pearson $r = 0.04$, $p = 0.78$). To further quantify this relationship, a sequential (or hierarchical) linear regression was performed on the BOLD data in Fig. 4C. For congruent trials, RT explained 43.4% of the variance; the addition of error likelihood to the model increased this value by 4.2%, but the increase was not significant ($p = 0.10$). For incongruent trials, RT explained 89.7% of the variance. The addition of error likelihood increased this value by less than 1×10^{-5}%. These data suggest that response conflict cannot explain a significant or physiologically relevant amount of variance in the dMFC.

The quantile analysis was repeated on a voxel-by-voxel basis to determine if any region of the dMFC showed activity consistent with the conflict monitoring model. No significant clusters were found (Fig. 5), even using an extremely lenient significance threshold of $p = 0.01$, uncorrected for multiple comparisons; this result indicates that the relationship between RT and dMFC activity does not depend on the exact position, shape, or extent of the region of interest tested. It is possible that methodological differences in our Stroop task may have produced brain activity that is uncharacteristic of previous studies. Voxels that showed greater activity on incongruent trials in our Stroop task were compared to those of previous Stroop (Fig. 6A) and non-Stroop (Fig. 6B) conflict studies. The position, shape, and extent of our region of interest were remarkably consistent with those of previous studies (Fig. 6C). To determine whether "conflict" voxels (from Fig. 2A) were equivalent to "RT" voxels, we performed a novel GLM analysis in which the design matrix consisted of a single RT regressor that did not differentiate between congruent and incongruent trials. The voxels identified by the Incong>Cong contrast from Fig. 2A were subsumed by those correlated with the RT-only regressor; this was also true for the majority of voxels from previous studies (Fig. 6D).

Finally, the quantile analysis was used to test whether any voxels outside the dMFC showed greater activity for incongruent than congruent trials (Fig. 7). Only the left inferior frontal gyrus, which includes Broca's area, expressed significantly greater BOLD activity per unit time on incongruent trials. This result demonstrates that the

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

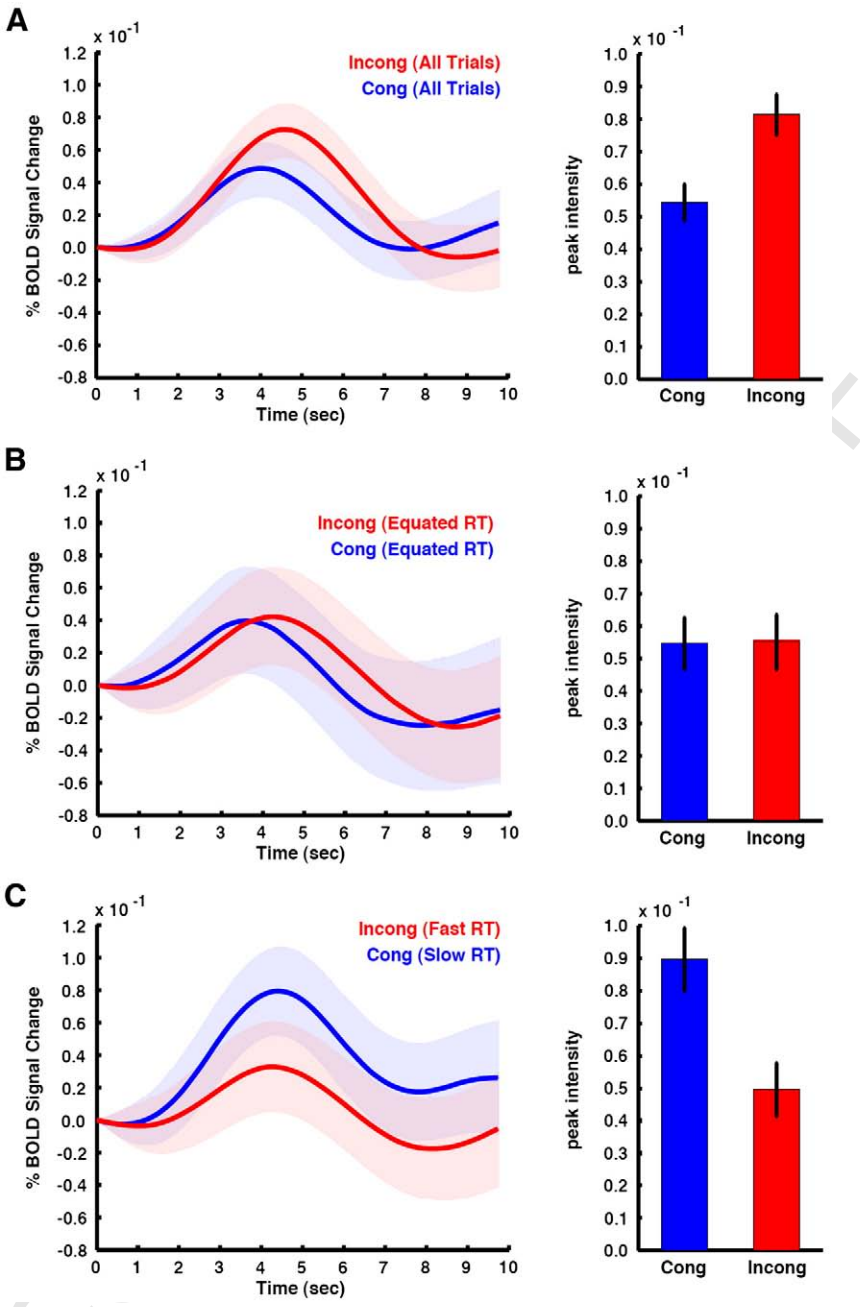

Fig. 3. Event-related averages. BOLD data was extracted from voxels active in the incongruent>congruent comparison (all comparisons used Fig. 2A as the region of interest). BOLD responses from correct trials were then averaged across subjects (shading represents standard error). (A) When all trials in the RT distribution are included in the analysis, average BOLD responses are larger for the incongruent condition. To quantify the differences between the two BOLD responses, the peak response for each subject was compared. Bar graphs show a significantly larger response for the incongruent than congruent condition, consistent with previous studies of conflict monitoring (error bars represent standard error across subjects). (B) To control for mean differences in RT between conditions, we compared only trials within 100 ms of each subject's median RT. No differences between congruent and incongruent trials were detected. (C) A comparison of fast incongruent and slow congruent trials produced a reversal in the relative size of the BOLD responses, demonstrating that slower RTs produce greater dMFC activity independent of stimulus congruency.

lack of significant voxels in dMFC is not due to insufficient statistical power, but rather, to the fact that interference between alternative semantic representations of color is localized to left IFG, not dMFC.

Discussion

In conclusion, we used three different approaches, a general linear model, event-triggered averaging, and RT quantile analysis, to demonstrate that dMFC activity is correlated with time on task, and not response conflict. Furthermore, the results showed that error likelihood was not monotonically related to either RT or dMFC activity, contrary to predictions of the conflict monitoring model. These data are inconsistent with a view of a dMFC specialized for conflict monitoring. Instead, dMFC activity is predicted by trial-to-trial differences in response time (Fig. 4C). Previous studies of conflict monitoring (Botvinick et al., 1999; Brown and Braver, 2005; Carter et al., 1998; Kerns et al., 2004; MacLeod and MacDonald, 2000; Nee et al., 2007) averaged trials with variable response durations when comparing neural activity from congruent and incongruent trials; however, incongruent trials have longer mean response times ensuring that they will produce greater dMFC activity, independent of interference effects. In countermanding tasks, differences in time on task between conditions can also explain correlations between dMFC activity and error likelihood. For example, in the study by

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

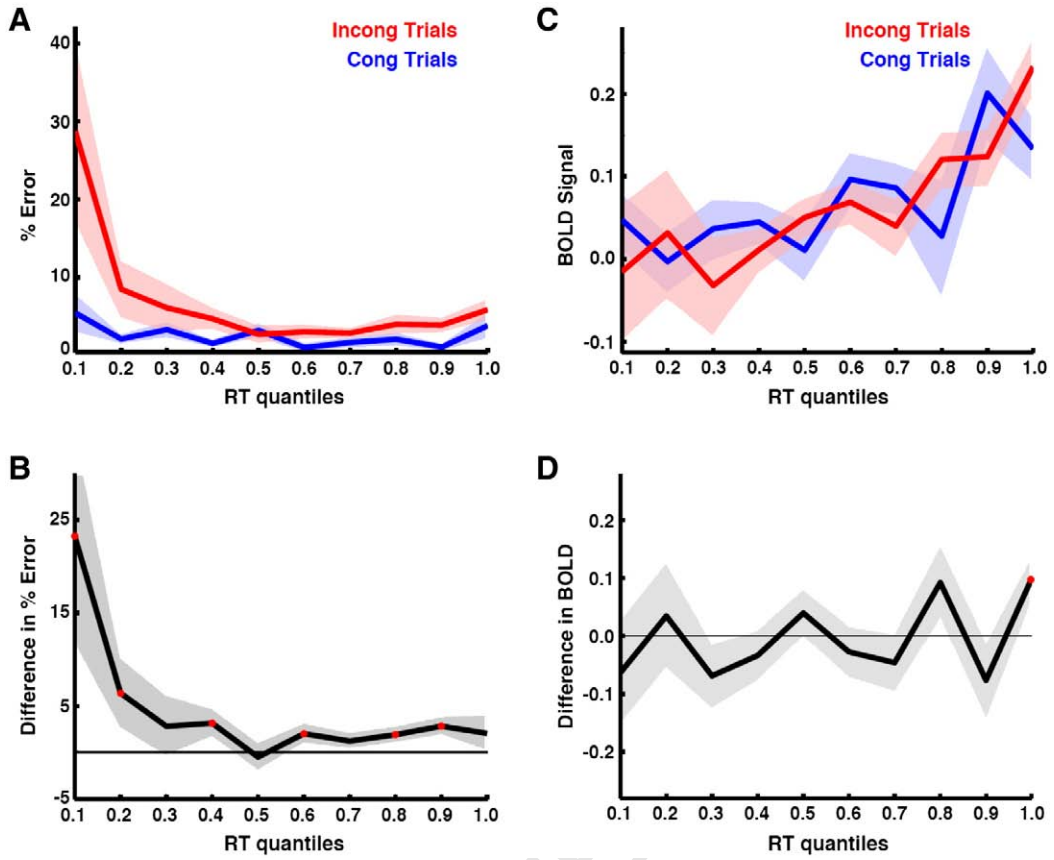

Fig. 4. Error likelihood and BOLD differences across RTs. (A) To determine the relationship between conflict and RT, the percent error across subjects was plotted for congruent and incongruent trials as a function of RT quantile. (B) The difference in error rates between the two conditions (i.e. incongruent–congruent trials) shows greater conflict on incongruent trials for most RT quantiles. Red points indicate quantiles for which a significant difference was present. (C) The mean BOLD response for correct trials was integrated over 10 s, averaged across subjects, and plotted as a function of RT quantile. The BOLD signal showed a systematic increase as a function of RT for both congruent and incongruent trials, consistent with the time on task account (Fig. 1D), but contrary to the conflict monitoring model (Figs. 1ABC). (D) The difference between the two BOLD responses was plotted for each RT quantile. Positive values indicate larger responses for incongruent trials; negatives values indicate larger responses for congruent trials. Resulting values were centered on zero, indicating that after controlling for RT, BOLD responses were not affected by conflict. This quantile analysis was repeated for two other masks defined using anatomical landmarks and functional imaging meta-analysis (Fig. S4), and demonstrated similar results.

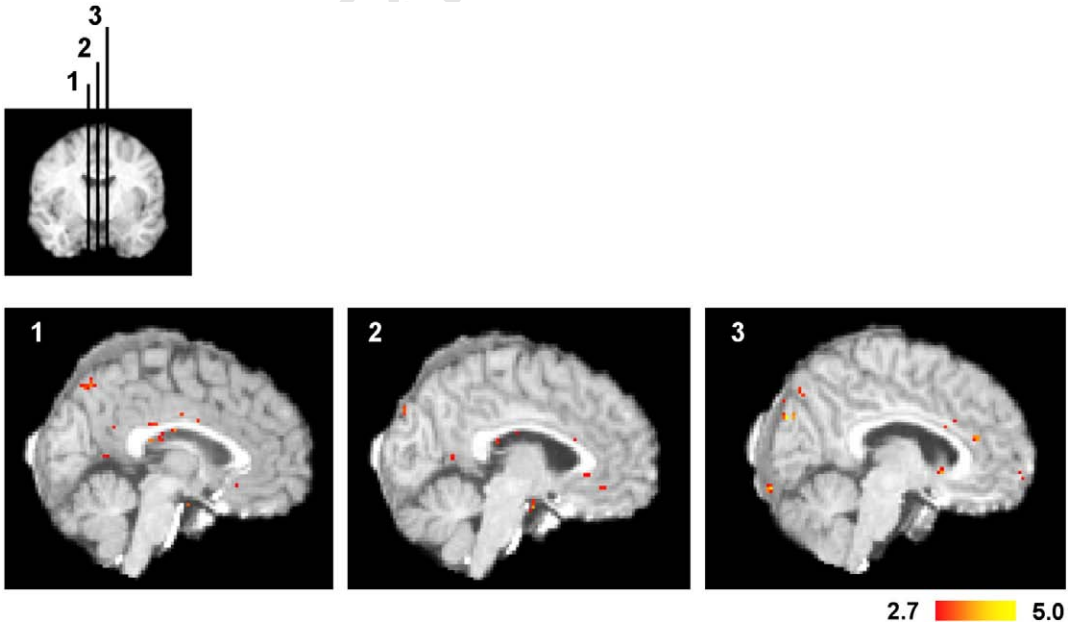

Fig. 5. Voxel-wise comparison of congruent and incongruent trials controlling for RT. For each subject, the BOLD responses were averaged across quantiles within each condition. A one-sided, paired *t*-test was performed between congruent and incongruent trials across subjects. Voxel *t*-values were thresholded at $p = 0.01$. To determine if any region of the dMFC might be involved in conflict monitoring, we performed the analysis without correcting for multiple comparisons. Despite this extremely lenient threshold, we found no regions in the dMFC that were consistent with the conflict monitoring hypothesis. This lack of significant regions in dMFC is not due to insufficient statistical power, since sufficient power was present to detect activity in Broca's area.

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

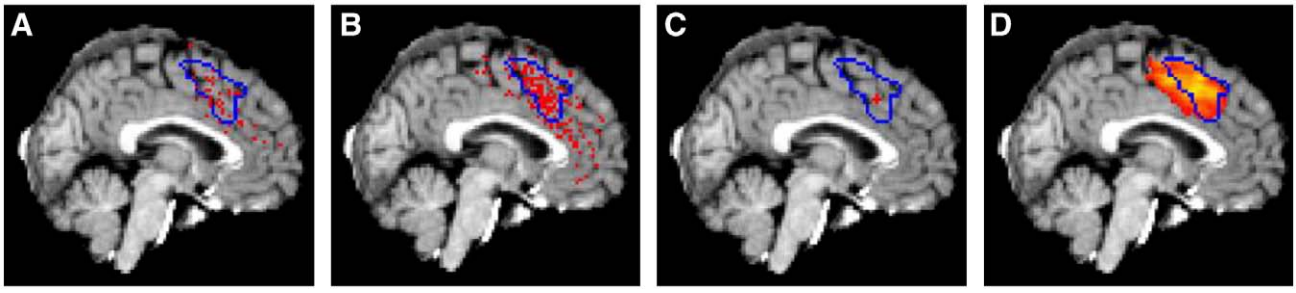

Fig. 6. Comparison against previous studies. (A) The region of interest tested for conflict-related activity was defined using a functional contrast of correct incongruent minus correct congruent trials (Fig. 2A; blue outline). To determine how similar this activity was to that of previous studies, we overlaid peak activation from 48 studies using the Stroop task (Nee et al., 2007). Each point represents the peak activation from an incongruent minus congruent comparison. The majority of previous studies were consistent with our activation. (B) We repeated the analysis for 200 non-Stroop studies (Nee et al., 2007) that identified conflict-related activity. The majority was consistent with our activation. (C) The mean locations (red cross) of both the Stroop studies (mean MNI152: 2/21/40, std: 7/13/13) and the non-Stroop studies (mean MNI152: 2/21/40, std: 6/14/16) were consistent with our activation. (D) To determine which of the voxels that show greater activity for the incongruent trials are also monotonically related to RT, we performed a GLM analysis using a single epoch regressor in which the duration of each epoch was equal to RT for that trial. Furthermore, this regressor included all correct trials, that is, it did not differentiate between congruent and incongruent epochs. All voxels that showed greater activity in the incongruent minus congruent contrast also showed a significant relationship to RT. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Brown and Braver (2005), both the period of focused attention toward the countermanding cue and the resulting RTs, were longer in the high error than the low error condition; thus, larger BOLD responses on high error likelihood trials are consistent with the time on task account. This study shows that once RT differences are accounted for, the correlation between error likelihood and dMFC activity is zero.

Though the dMFC is independent of response conflict and error likelihood, it has been shown to have a variety of other functional roles. For example, dMFC activity tracks reinforcement history and the occurrence of errors, which is critical for learning the value of response alternatives and making behavioral adjustments (Rushworth and Behrens, 2008; Rushworth et al., 2007). It has also been shown to be involved in motor planning, task switching, and inhibitory control (Nachev et al., 2008; Ridderinkhof et al., 2007) and it is modulated by decision uncertainty (Grinband et al., 2006) and complexity of stimulus–response associations (Nachev et al., 2008) suggesting its involvement in response selection. While it is not known whether these functions are localized to discrete subregions of the dMFC or whether the dMFC's role changes with context, this study suggests that determining the answer will require accurate control over response times.

Finally, activity in Broca's area was correlated with increased response competition. A meta-analysis of Stroop studies (Nee et al., 2007) found this region to be consistently more activated during incongruent trials. It has also been shown to be highly active in a linguistic emotional Stroop task (Engels et al., 2007) and, in fact, is associated more with linguistic, than visuospatial, interference (Banich et al., 2000). Interestingly, no differences in activity between conditions were found in any motor execution or motor planning regions (e.g. SMA or primary motor regions), even at extremely liberal thresholds ($p = 0.01$ with no correction for multiple comparisons; Fig. 5). These data suggest that the cognitive interference experienced during the Stroop task does not stem from competing motor response representations (i.e. left vs right button press) but rather from competing *linguistic* representations of the stimulus (i.e. red vs green). Moreover, it is possible that neural competition is localized to those brain regions specialized for the relevant stimulus modality, and a centralized brain module for detecting different types of conflict (sensory, linguistic, motor, etc.) may not be necessary for optimal response selection to occur.

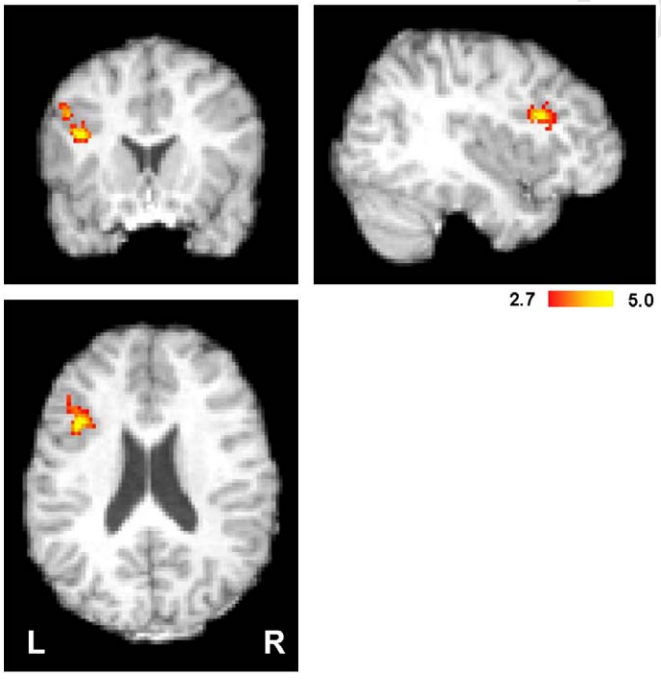

Fig. 7. Activity in left inferior frontal gyrus (BA44/45, including Broca's area) correlates with increased conflict after controlling for RT. For each subject, the BOLD responses were averaged across quantiles within each condition. A one-sided, paired *t*-test was performed between congruent and incongruent trials across subjects. Greater activity in Broca's area suggests increased competition between multiple linguistic representations of the stimulus on incongruent trials. Voxel *t*-values were thresholded at $p = 0.01$, clusters thresholded at $p = 0.01$ using Gaussian Random Field Theory. Peak activity was located at MNI152: -38/16/22.

Methods

Task

Twenty-three subjects (mean age = 24; 9 females) gave informed consent to participate in the study. Subjects were instructed to name the ink color of a stimulus by pressing one of four buttons using the fingers of the right hand and were asked to balance speed with accuracy. They practiced the task outside the scanner for at least 100 trials and continued practice until error rates fell below 10%. Four colors were used in the Stroop task – red, green, blue, and yellow – producing 16 possible ink color/word name combinations. To eliminate any confounding effects of negative priming on conflict, the ink color and word name were randomized such that no two colors were repeated on consecutive trials. The probability of incongruent and congruent trials was maintained at 50%. The Stroop stimuli were presented on screen until a response was made. Once a button was pressed, the stimulus was replaced by a fixation point.

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

Inter-trial intervals were randomly generated on each trial (range = 0.1–7.1 s, mean = 3.55 s, incremented in steps of 0.5 s). RT was used as an estimate of the neural time on task. We assumed that for most trials, subjects were awake and engaged, and that lapses of attention were equally distributed between congruent and incongruent trials, and accounted for only a small fraction of the overall RT variance. Furthermore, mean differences in RT between conditions were assumed to be modulated by changes in task difficulty related to the presence or absence of response interference.

Image acquisition

Imaging experiments were conducted using a 1.5 T GE TwinSpeed Scanner using a standard GE birdcage head coil. Structural scans were performed using the 3D SPGR sequence (124 slices; 256×256; FOV = 200 mm). Functional scans were performed using EPI-BOLD (TE = 60; TR = 2.0 s; 25 slices; 64×64; FOV = 200 mm; voxel size = 3 mm×3 mm×4.5 mm). All image analysis was performed using the FMRIB Software Library (FSL; http://www.fmrib.ox.ac.uk/fsl/) and Matlab (Mathworks, Natick, MA; http://www.mathworks.com). The data were motion corrected (FSL-MCFLIRT), high pass filtered (at 0.02 Hz), spatially smoothed (full width at half maximum = 5 mm), and slice time corrected.

Image analysis

Standard statistical parametric mapping techniques (Smith et al., 2004) (FSL-FEAT) were performed prior to registration to MNI152 space (linear template). Multiple linear regression was used to identify voxels that correlated with each of the four decisions types i.e. congruent, incongruent, error, and post-error trials. A primary statistical threshold for activation was set at $p = 0.05$ and cluster correction was set at $p = 0.05$ using Gaussian random field theory. Inter-subject group analyses were performed in standard MNI152 space by applying the FSL-FLIRT registration transformation matrices to the parameter estimates. For each run, the transformation matrices were created by registering via mutual information (1) the midpoint volume to the first volume using 6 degrees of freedom, (2) the first volume to the SPGR structural image using 6 degrees of freedom, and (3) the SPGR to the MNI152 template using 12 degrees of freedom. These three matrices were concatenated and applied to each statistical image.

Two regression models were used to generate activation maps (Fig. 2). To replicate previous studies, a regression model was created with three regressors, one for each trial type (congruent, incongruent, and error/post-error trials). Each regressor consisted of a series of impulse functions positioned at the onsets of the trials. The regressors were then convolved with each subject's custom hemodynamic response function (HRF) extracted from primary visual cortex (Grinband et al., 2008). The custom HRF estimate was computed using FLOBS (Woolrich et al., 2004), a three-function basis set that restricts the parameter estimates of each basis function to generate physiologically plausible results.

To contrast the effect of time on task with conflict, three contrasts were performed (1) incongruent vs congruent using all RTs, (2) slow congruent vs fast congruent and (3) slow congruent vs fast incongruent. In all three of these contrasts, only correct trials were used; error trials and trials immediately following error trials (i.e. post-error trials) were modeled explicitly using separate regressors. Using correct trials allowed us to study pre-response conflict. Due to the overall low frequency of errors in this task, it was not possible to study error-induced conflict or post-error processing on any of the subsequent analyses. For the incongruent vs congruent contrast, a regression model was created with four regressors: (1) congruent trials, (2) incongruent trials, (3) error trials, and (4) post-error trials. For the slow congruent vs fast congruent contrast, a regression model was created with six regressors: (1) congruent trials with RT less than the median RT of congruent trials, (2) congruent trials with RT greater than the median RT of congruent trials, (3) incongruent trials with RT less than the median RT of incongruent trials, (4) incongruent trials with RT greater than the median RT of incongruent trials, (5) error trials, and (6) post-error trials. For the slow congruent vs fast incongruent comparison six regressors were used: (1) congruent trials with RT less than the global median RT, (2) congruent trials with RT greater than the global median RT, (3) incongruent trials with RT less than the global median RT, (4) incongruent trials with RT greater than the global median RT, (5) error trials, and (6) post-error trials. Using the global median RT ensured no overlap between RT durations of congruent and incongruent trials. In all three regression models, trials with RTs less than 200 ms were excluded from the analysis. The median RT was computed independently for each subject and included only correct trials.

The event-related averaging (Fig. 3) was performed by extracting time series from the voxels identified using the traditional contrast of incongruent>congruent trials (Fig. 2A) using all correct trials with RTs greater than 200 ms. The mean time series from this mask was interpolated to 10 ms resolution and the responses for each event type were averaged for each subject. The mean response was then averaged across subjects and used in the statistical analysis. Paired t-tests were performed across subjects to test differences in the size of the BOLD response.

The quantile analysis (Fig. 4) was performed by generating an event-related average (similar to the data in Fig. 3) for each subject's ten RT quantiles. To summarize the steps: (1) an RT distribution for each subject was computed, (2) the deciles of each distribution were calculated, (3) the average error likelihood or BOLD response was computed for each decile, (4) the error rates or BOLD responses for each decile were averaged across subjects for the group analysis. Each subject's deciles were computed from an RT distribution that contained all of the subject's trials, both congruent and incongruent, thus ensuring that the same deciles were used for both trial types. Trials less than 200 ms in duration were excluded. In the group analysis, the congruent and incongruent mean error rates across subjects were calculated for each decile and a *t*-test on the difference was performed (Fig. 4A). A BOLD response magnitude was computed for each decile by integrating over the first 10 s of the BOLD response and averaging this value across subjects (Fig. 4B). Integrating over the first 10 s averages over the physiological noise, resulting in less variability than using the peak response.

To perform a whole brain analysis that is controlled for effects of response duration (Fig. 5), BOLD responses were averaged across RT quantiles for each subject. This effectively gives equal weight (1/10th) to each RT quantile when comparing the BOLD response between conditions i.e.

$$\mathrm{cHDR} = \left(\mathrm{cHDR}_{q1} + \mathrm{cHDR}_{q2} + \ldots + \mathrm{cHDR}_{q10}\right) / 10$$

where cHDR_{qx} is the estimated BOLD response for congruent trials within the corresponding quantile (qx). Thus, frequency differences in the RT distributions of congruent and incongruent trials no longer affect the shape or amplitude of the BOLD response. This analysis is equivalent to asking the question: Which voxels show greater activity per unit time for incongruent than congruent trials? For each voxel in the brain, a one-sided paired *t*-test was performed (iHDR>cHDR, $p<0.01$) across subjects to compare the integral of the BOLD response between conditions. The activation map was corrected for multiple comparisons using Gaussian Random Field theory ($p<0.01$).

Acknowledgments

We would like to thank the following people for providing very useful discussions and comments on the manuscripts: Jonathan D. Cohen, Tobias Egner, Peter Freed, Franco Pestilli, Ted Yanagihara, and

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

Tal Yarkoni. We also thank Greg Siegle and Matthew Botvinick for providing the Matlab code implementing the conflict monitoring model. This research was supported by NIH T32MH015174 (JG), NSF 0631637 (TDW), MH073821 (VPF), and fMRI Research Grant (JH).

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2010.12.027.

References

Banich, M.T., Milham, M.P., Atchley, R., Cohen, N.J., Webb, A., Wszalek, T., Kramer, A.F., Liang, Z.P., Wright, A., Shenker, J., Magin, R., 2000. fMri studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. J. Cogn. Neurosci. 12, 988–1000.

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402, 179–181.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. Psychol. Rev. 108, 624–652.

Brown, J.W., Braver, T.S., 2005. Learned predictions of error likelihood in the anterior cingulate cortex. Science 307, 1118–1121.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A., Rosen, B.R., 2002. Dorsal anterior cingulate cortex: a role in reward-based decision making. Proc. Natl Acad. Sci. USA 99, 523–528.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280, 747–749.

di Pellegrino, G., Ciaramelli, E., Ladavas, E., 2007. The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. J. Cogn. Neurosci. 19, 275–286.

Engels, A.S., Heller, W., Mohanty, A., Herrington, J.D., Banich, M.T., Webb, A.G., Miller, G.A., 2007. Specificity of regional brain activity in anxiety types during emotion processing. Psychophysiology 44, 352–363.

Grinband, J., Hirsch, J., Ferrera, V.P., 2006. A neural representation of categorization uncertainty in the human brain. Neuron 49, 757–763.

Grinband, J., Wager, T.D., Lindquist, M., Ferrera, V.P., Hirsch, J., 2008. Detection of time-varying signals in event-related fMRI designs. Neuroimage 43, 509–520.

Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. J. Neurosci. 30, 3339–3346.

Hayward, G., Goodwin, G.M., Harmer, C.J., 2004. The role of the anterior cingulate cortex in the counting Stroop task. Exp. Brain Res. 154, 355–358.

Ito, S., Stuphorn, V., Brown, J.W., Schall, J.D., 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 302, 120–122.

Kerns, J.G., Cohen, J.D., MacDonald III, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior cingulate conflict monitoring and adjustments in control. Science 303, 1023–1026.

MacLeod, C.M., MacDonald, P.A., 2000. Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. Trends Cogn. Sci. 4, 383–391.

Mansouri, F.A., Buckley, M.J., Tanaka, K., 2007. Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. Science 318, 987–990.

Mansouri, F.A., Tanaka, K., Buckley, M.J., 2009. Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. Nat. Rev. Neurosci. 10, 141–152.

Milham, M.P., Banich, M.T., 2005. Anterior cingulate cortex: an fMRI analysis of conflict specificity and functional differentiation. Hum. Brain Mapp. 25, 328–335.

Nachev, P., Kennard, C., Husain, M., 2008. Functional role of the supplementary and pre-supplementary motor areas. Nat. Rev. Neurosci. 9, 856–869.

Naito, E., Kinomura, S., Geyer, S., Kawashima, R., Roland, P.E., Zilles, K., 2000. Fast reaction to different sensory modalities activates common fields in the motor areas, but the anterior cingulate cortex is involved in the speed of reaction. J. Neurophysiol. 83, 1701–1709.

Nakamura, K., Roesch, M.R., Olson, C.R., 2005. Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. J. Neurophysiol. 93, 884–908.

Nee, D.E., Wager, T.D., Jonides, J., 2007. Interference resolution: insights from a meta-analysis of neuroimaging tasks. Cogn. Affect. Behav. Neurosci. 7, 1–17.

Olson, C.R., 2003. Brain representation of object-centered space in monkeys and humans. Annu. Rev. Neurosci. 26, 331–354.

Pardo, J.V., Pardo, P.J., Janer, K.W., Raichle, M.E., 1990. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. Proc. Natl Acad. Sci. USA 87, 256–259.

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004. The role of the medial frontal cortex in cognitive control. Science 306, 443–447.

Ridderinkhof, K.R., Nieuwenhuis, S., Braver, T.S., 2007. Medial frontal cortex function: an introduction and overview. Cogn. Affect. Behav. Neurosci. 7, 261–265.

Roelofs, A., van Turennout, M., Coles, M.G., 2006. Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. Proc. Natl Acad. Sci. USA 103, 13884–13889.

Rushworth, M.F., Behrens, T.E., 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat. Neurosci. 11, 389–397.

Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E., Bannerman, D.M., 2007. Functional organization of the medial frontal cortex. Curr. Opin. Neurobiol. 17, 220–227.

Schall, J.D., 1991. Neuronal activity related to visually guided saccadic eye movements in the supplementary motor area of rhesus monkeys. J. Neurophysiol. 66, 530–558.

Siegle, G.J., Steinhauer, S.R., Thase, M.E., 2004. Pupillary assessment and computational modeling of the Stroop task in depression. Int. J. Psychophysiol. 52, 63–76.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 (Suppl 1), S208–S219.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643–662.

Stuphorn, V., Brown, J.W., Schall, J.D., 2010. Role of supplementary eye field in saccade initiation: executive, not direct, control. J. Neurophysiol. 103, 801–816.

Swick, D., Jovanovic, J., 2002. Anterior cingulate cortex and the Stroop task: neuropsychological evidence for topographic specificity. Neuropsychologia 40, 1240–1253.

Turken, A.U., Swick, D., 1999. Response selection in the human anterior cingulate cortex. Nat. Neurosci. 2, 920–924.

Vendrell, P., Junque, C., Pujol, J., Jurado, M.A., Molet, J., Grafman, J., 1995. The role of prefrontal regions in the Stroop task. Neuropsychologia 33, 341–352.

Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory: a meta-analysis. Cogn. Affect. Behav. Neurosci. 3, 255–274.

Wager, T.D., Jonides, J., Reading, S., 2004. Neuroimaging studies of shifting attention: a meta-analysis. Neuroimage 22, 1679–1693.

Wager, T.D., Lindquist, M.A., Nichols, T.E., Kober, H., Van Snellenberg, J.X., 2009. Evaluating the consistency and specificity of neuroimaging data using meta-analysis. Neuroimage 45, S210–S221.

Woolrich, M.W., Behrens, T.E., Smith, S.M., 2004. Constrained linear basis sets for HRF modelling using Variational Bayes. Neuroimage 21, 1748–1761.

Yarkoni, T., Barch, D.M., Gray, J.R., Conturo, T.E., Braver, T.S., 2009. BOLD correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fMRI analysis. PLoS ONE 4, e4257.

Zhu, D.C., Zacks, R.T., Slade, J.M., 2010. Brain activation during interference resolution in young and older adults: an fMRI study. Neuroimage 50, 810–817.

Please cite this article as: Grinband, J., et al., The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.12.027

Supplementary Materials

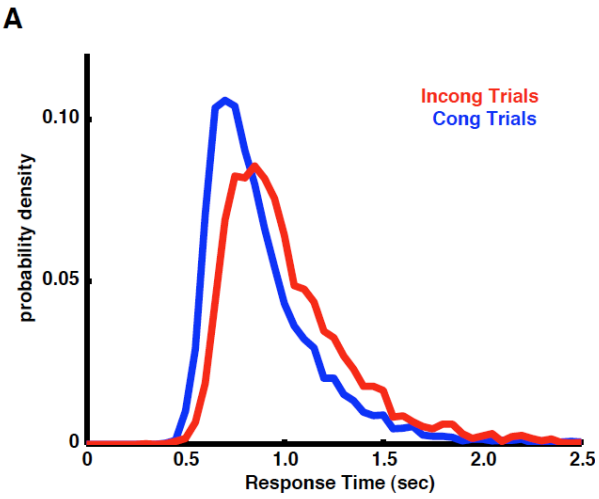

Figure S1 – RT probability density. High conflict (incongruent) trials have longer mean response times than low conflict (congruent) trials. When averaging across trial durations, both the conflict monitoring and the time on task accounts predict increased neural activity on incongruent trials, either from increased conflict or longer time on task.

To determine the relationship between conflict, error likelihood, and RT, we performed simulations using the conflict monitoring model (Botvinick et al., 2001; Siegle et al., 2004). Parameters were set to those used in Botvinick et al (2001): tau (rate of propagation through network) = 0.1; std dev of noise = 0.015; response threshold = 0.6; task input strength = 0.5; task unit gain = 1; bias vector = [0 0 -4 -4 -4 -4 -4 -4 0 0]. The unit weighting were a = -2, b = 4, c = 1.5, d = 2.5, e = 2, f = 2, m = -2, n = -2, where a = stim layer inhibition, b = color to color-task unit, c = color to color-response unit, d = word to word-response unit, e = from word to word-task unit, f = from colors to color-task unit, m = response layer inhibition, n = task layer inhibition. The network was primed with 1000 iterations.

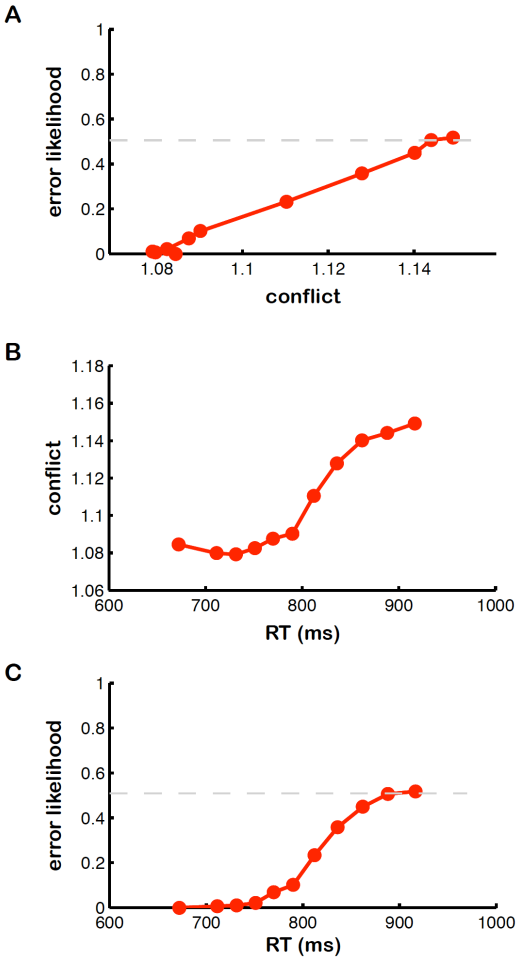

Figure S2– Conflict monitoring model. We executed 10,000 iterations of the conflict monitoring model to determine the relationship between conflict, error likelihood, and RT. (A) Error likelihood is linearly related to conflict. Thus, error likelihood can be used to determine if conflict is present after equating RT between congruent and incongruent trials. (B) Conflict is sigmoidally related to RT. (C) Error likelihood is sigmoidally related to RT. Our behavioral data does not show this sigmoidal relationship. Furthermore, contrary to predictions of the conflict monitoring model, the relationship between error likelihood and RT changes with congruency.

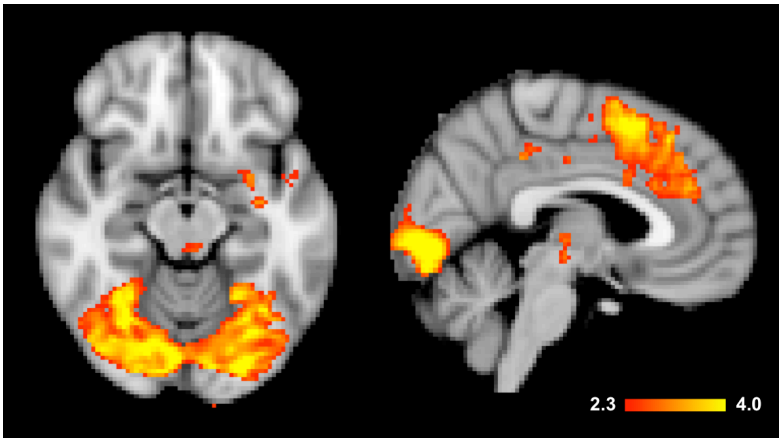

Figure S3 – Flashing checkerboard. To determine whether the dMFC can be activated in the absence of response conflict we asked a new group of subjects (n = 17) to view a flashing checkerboard and press a button whenever the stimulus disappeared. The stimuli were presented with random inter-trial intervals (0.1 – 7.1 sec) and randomly generated stimulus durations (gamma distributed, mean = 800 ms, min = 500 ms). In this task, no choice decision is made, no errors can occur, and only a single response is possible; thus, response conflict cannot exist. A GLM analysis was performed using a single regressor that modeled the checkerboards as a set of variable duration boxcars with onset and offset times that matched the actual stimuli. The gamma distribution simulated typical RTs found in the Stroop task. The task activated visual cortex and dMFC, consistent with the time on task account.

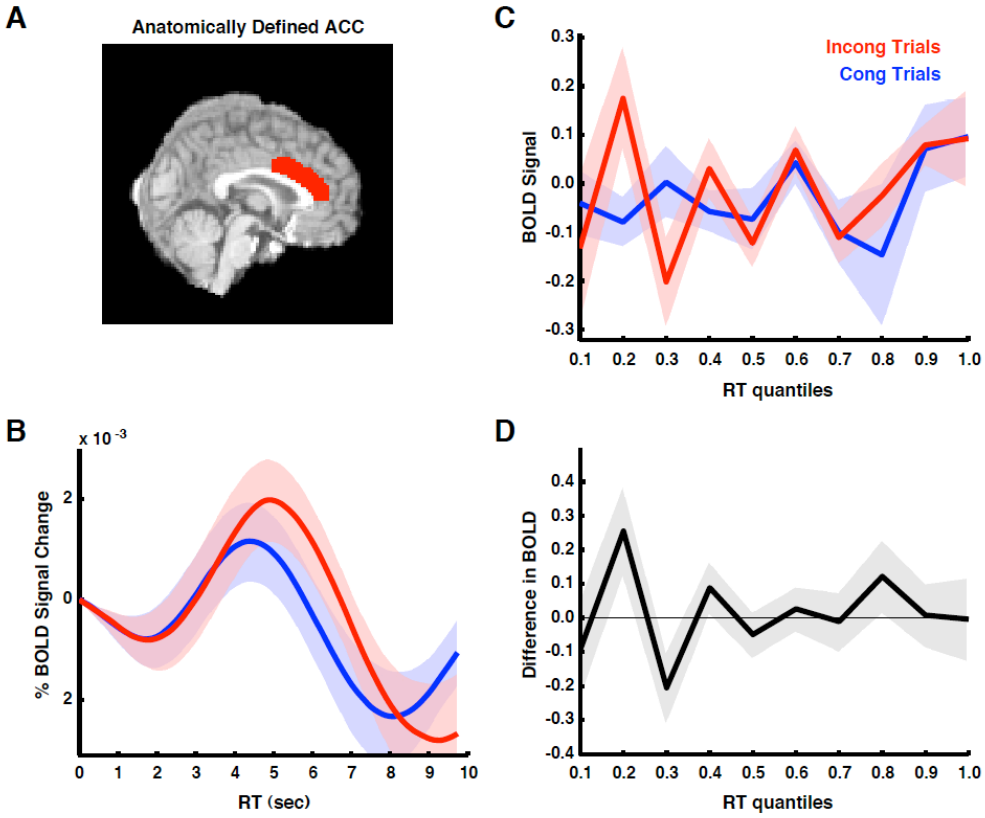

Figure S4 – BOLD activity for anatomically defined ACC. (A) The ACC was defined as the cortex between the cingulate sulcus and corpus callosum posterior to the genu and anterior to the collosal midpoint. (B) The hemodynamic responses for correct congruent and incongruent trials are about two orders of magnitude smaller in the anatomically-defined ACC than in the functionally defined dMFC mask. (C) The BOLD signal was integrated over 10 sec and plotted as a function of RT quantile. (D) No significant differences were present between congruent and incongruent trials ($p < 0.05$).

The voxels used in the model-free analysis (Fig 3) were identified using a functional contrast of the two conflict conditions (Fig 2A). This ROI is located above the cingulate sulcus. Since previous studies have focused on the anterior cingulate cortex (ACC) as the region involved in conflict monitoring, we repeated the quantile analysis using an anatomically defined ACC mask. The ACC was defined as the cortex between the corpus callosum and the cingulate sulcus, posterior to the genu of the corpus callosum and anterior to the rostro-caudal midpoint of the corpus callosum (Mansouri et al., 2009). There were no significant differences in ACC activity between congruent and incongruent trials (Fig S4). This analysis was repeated using an ACC mask derived from a meta-analysis of 47 conflict-related studies (mask centered on MNI152: 0/20/40) (Nee et al., 2007). This mask also showed a monotonic relationship with RT but no significant differences between conditions.