Published in final edited form as:

Neuroimage. 2011 July 15; 57(2): 316–319. doi:10.1016/j.neuroimage.2011.04.029.

Errors of interpretation and modeling: A reply to Grinband et al.

Nick Yeung^{a,*}, Jonathan D. Cohen^b, and Matthew M. Botvinick^b

^aDepartment of Experimental Psychology, University of Oxford, UK

^bDepartment of Psychology and Princeton Neuroscience Institute, Princeton University, USA

Abstract

Grinband et al., 2011 compare evidence that they have collected from a neuroimaging study of the Stroop task with a simulation model of performance and conflict in that task, and interpret the results as providing evidence against the theory that activity in dorsal medial frontal cortex (dmFC) reflects monitoring for conflict. Here, we discuss several errors in their methods and conclusions and show, contrary to their claims, that their findings are entirely consistent with previously published predictions of the conflict monitoring theory. Specifically, we point out that their argument rests on the assumption that conflict must be greater on all incongruent trials than on all congruent trials—an assumption that is theoretically and demonstrably incorrect. We also point out that their simulations are flawed and diverge substantially from previously published implementations of the conflict monitoring theory. When simulated appropriately, the conflict monitoring theory predicts precisely the patterns of results that Grinband et al. take to present serious challenges to the theory. Finally, we note that their proposal that dmFC activity reflects time on task is theoretically weak, pointing to a direct relationship between behavior (RT) and neural activity but failing to identify any intervening psychological construct to relate the two. The conflict monitoring theory provides such a construct, and a mechanistic implementation that continues to receive strong support from the neuroimaging literature, including the results reported by Grinband et al.

The conflict monitoring theory proposes that regions within dorsal medial frontal cortex (dmFC), including the anterior cingulate cortex, monitor for the occurrence of competition, or conflict, in action selection to detect and signal the need for increased cognitive control (Botvinick et al., 2001; Carter et al., 1998). This theory provides an account of perhaps the most replicated neuroimaging finding regarding dmFC: its increase in activation in conditions of high cognitive demand. A canonical example of this activation is seen in the Stroop task, in which greater dmFC activity is observed when color and word information are incongruent (e.g., RED in blue ink) than when they are congruent (e.g., BLUE in blue ink) (Bench et al., 1993). According to the theory, this activity reflects the increase in conflict generated by the co-activation of mutually incompatible actions in the incongruent condition (e.g., trying to say “red” and “blue” simultaneously) relative to the congruent condition. In the target article of this commentary, Grinband et al., 2011 present a novel critique of this proposed relationship between dmFC activity, conflict, and behavioral indices of task demand. Here, we discuss several significant weaknesses in their methods and conclusions, and show that their findings are entirely consistent with previously published predictions of the conflict monitoring theory.

The essence of Grinband et al.'s argument is that conflict is superfluous to explaining observed variations in dMFC activity across conditions: They suggest that conflict explains no residual variance in the data once the effects of RT (time on task) are partialled out. In support of this conclusion, Grinband et al. present three key findings from an analysis of fMRI data collected from the Stroop task. First, they show that dMFC activity does not differ for incongruent and congruent trials that are matched in terms of RT. Second, extending this analysis to dissociate conflict and RT, they show that dMFC activity is low on incongruent trials with fast RTs and high on congruent trials with slow RTs, reversing the typical effects of conflict. Finally, they show that whereas dMFC activity increases with RT for both congruent and incongruent trials, the likelihood of errors tends to decrease with RT—the ubiquitous speed–accuracy trade-off. On the assumption that error likelihood is a reasonable proxy for response conflict—an assumption they claim to validate with simulations of the conflict monitoring theory—Grinband et al. conclude that dMFC activity may even vary inversely with conflict. Thus, Grinband et al.'s data initially appear to seriously challenge the theory that dMFC activity reflects conflict.

We have three principal criticisms of Grinband et al.'s conclusions. First, their argument relies on an intuition that is superficially plausible—that conflict must be greater on all incongruent trials than on all congruent trials—but which is theoretically and demonstrably incorrect. We begin with an exploration of this point. Second, their simulations are flawed and diverge substantially from previously published implementations of the conflict monitoring theory. When simulated appropriately, the conflict monitoring theory predicts precisely the patterns of results that Grinband et al. take to present serious challenges to the theory. Finally, their proposal that dMFC activity reflects time on task is conceptually limited, since it implies a direct relationship between behavior (RT) and neural activity with no intervening psychological construct to relate the two.

Misleading intuitions

Two intuitive and uncontroversial statements about the conflict monitoring theory are (1) that conflict should be greater on incongruent trials than on congruent trials, and (2) that RT should increase as a function of experienced conflict. In most cases, these two statements run together and the predictions of the theory are clear; in particular, conflict and RT both typically increase on incongruent trials relative to congruent trials in the Stroop task. However, it is less obvious how conflict should vary when the effects of congruence and RT are set in opposition, as is the case when contrasting fast incongruent trials with slow congruent trials. Grinband et al. base their analysis and conclusions on the intuition that conflict should vary with congruence and dissociate from RT in this contrast, being greater for fast incongruent than slow congruent trials. But this intuition is incorrect. As demonstrated in previously published conflict model simulations (Yeung et al., 2004), conflict more closely tracks RT than congruence condition when the two are dissociated.

This important prediction of the conflict monitoring theory follows from two key principles. The first is that activity of the conflict monitoring system should scale with conflict regardless of its source: When the conflict monitor responds to co-activation of mutually incompatible actions, it does not distinguish between conflict caused by incongruent stimulus features and conflict caused by other sources of processing variability and noise. Second, and crucially, these additional sources of processing variability—such as trial-to-trial fluctuations in attentional focus, noise in stimulus processing, and idiosyncratically varying response biases—result in performance variability that dwarfs the variance caused by stimulus congruence (as illustrated in Fig. S1 of the target article). The consequence is that very slow RTs will be observed on some congruent trials—for example when participants incorrectly perceive the stimulus, or happen to be prepared to make the incorrect

response rather than the correct one—and, conversely, very fast RTs will be observed on some incongruent trials—for example when participants' attentional focus is high, or when they fortuitously anticipate which stimulus will be presented and thus pre-activate the corresponding response. Crucially, however, slow congruent trials are not slow *despite* having low conflict, and fast incongruent trials are not fast *despite* having high conflict. To the contrary, slow congruent trials are slow precisely *because* conflict is high—a consequence of failing to focus attention, misperceiving the stimulus, preparing the wrong response, etc.—whereas fast incongruent trials are fast precisely *because* conflict is low. Thus, slow RT congruent trials should be associated with greater conflict (and, hence, greater dMFC activity) than fast RT incongruent trials.

To illustrate these points, Fig. 1A replots data from a previously published simulation in which we compared conflict on congruent and incongruent trials as a function of simulated RT (Yeung et al., 2004). Two features of the data are immediately evident: First, within each RT-matched bin, there is little or no difference in conflict as a function of stimulus congruence. Second, simulated conflict is markedly higher on slow congruent trials than fast incongruent trials; for example, conflict is nearly 10 times as high for the slowest congruent trials (RT bin 10) as it is for the fastest incongruent trials (RT bin 1). Thus, for the reasons outlined above, the model actually predicts the pattern of results that Grinband et al. put forward as a challenge to the conflict monitoring theory. As described in our earlier article, “In the model, slow responses to congruent stimuli are marked by high conflict (because of noise in processing) in just the same way as are slow responses to incongruent stimuli. The overall difference in conflict between congruent and incongruent trials reflects the fact that a greater proportion of congruent trials fall in the faster RT bins (as a result of low conflict), whereas incongruent trials tend to have greater RTs (as a result of high conflict).” (Yeung et al., 2004; p. 948).

These simulation results are derived from a model of conflict monitoring in the Eriksen flanker task, a spatial-attention analog of the Stroop task that has been the focus of the majority of our work in the past decade. Nevertheless, corresponding patterns are evident in simulations of conflict monitoring in the Stroop task (Fig. 1B). These simulations were based on the same model employed by Grinband et al., with parameters modified to approximate the error rates of their subjects. (The original model used by Botvinick et al. (2001) was parameterized to simulate performance in the verbal Stroop task, in which errors are vanishingly rare. Indeed, the original parameterization yields no errors at all, a point which leaves us uncertain as to how Grinband et al. obtained their results.) The crucial pattern of results shown in Fig. 1B—with equivalent levels of conflict for RT-matched congruent and incongruent trials, and markedly greater conflict for slow congruent trials than fast incongruent trials—is observed in both the original and updated models. Thus, contrary to Grinband et al.'s motivating intuition, the conflict monitoring theory predicts rather precisely the first two key findings in their data.

Flawed modeling

Grinband et al.'s third key finding is that dMFC activity varies inversely with error likelihood when trials are sorted into RT quantiles within each congruence condition: Trials with the fastest RTs have the highest error likelihood (reflecting the ubiquitous speed–accuracy tradeoff) but the lowest dMFC activity; for trials with slower RTs, error likelihood decreases while dMFC activity increases. Grinband et al. present simulation results indicating that error likelihood, measured in this way, provides an effective marker of response conflict, and thus conclude that their findings provide a further challenge to the conflict monitoring theory.

Before we raise concerns with the validity of Grinband et al.'s simulations, we first note that their empirical results are in fact perfectly consistent with published conflict model simulations. In particular, we have previously shown that error likelihood and conflict dissociate as a function of RT, in an analysis that anticipated precisely the quantile RT plot used by Grinband et al. (Yeung and Nieuwenhuis, 2009). As shown in Figs. 1A and C, the conflict theory predicts that conflict and error likelihood should vary in opposite ways as a function of RT—with conflict increasing and error likelihood decreasing—such that error likelihood is a very poor proxy for conflict. This dissociation is marked for incongruent trials on which errors occur relatively frequently. For congruent trials, error likelihood is very low across all RT bins irrespective of the level of conflict, once again capturing the detail of Grinband et al.'s experimental data. Thus, like dmFC activity in Grinband et al.'s fMRI study, conflict is highest on trials associated with the lowest error likelihood; specifically, trials with the slowest RTs. As we noted in our earlier paper, “It may seem somewhat counterintuitive that trials with the highest conflict should produce the fewest errors. In our simulations, this feature follows from the fact that activity in the target stimulus unit and correct response unit tend to increase over time, as stimulus processing progresses under the influence of attention. As a consequence, although trials with long RTs tend to have high levels of conflict, the responses ultimately made tend to be correct ... Conflict and error likelihood therefore vary in opposing ways as a function of response speed.” (Yeung and Nieuwenhuis, 2009; p. 14509).

These simulation results contrast sharply with those presented by Grinband et al., but the discrepancy does not seem to reflect the specific task being simulated (flanker task vs. Stroop task): Once again, simulation data from our Stroop model replicate closely the patterns seen in our flanker task model (Fig. 1D). We have been unable to replicate Grinband et al.'s simulation results using this model, and we are unsure why their model should produce results that are inconsistent with basic features of all evidence accumulation models of human decision making. Specifically, the simulation data plotted in Fig. S2C of the target article indicate that their model exhibits error free performance at the fastest RTs, with error rates rising to chance levels at the slowest RTs. This perplexing pattern is inconsistent with a fundamental feature of human decision making—the ubiquitously observed speed–accuracy trade-off—that is evident in their empirical data (their Fig. 4A) and is replicated in all of our previous models of conflict monitoring (including those presented here). In the absence of an adequate model of this basic feature of human cognition, the value of Grinband et al.'s simulation results remains unclear.

Conceptual limitation

An important strength of the conflict monitoring theory is that it provides a mechanistic account of the observed relationship between dmFC activity and task demand (as reflected in behavioral measures of RT and accuracy) in terms of an intervening psychological construct (detection of response conflict). In contrast, Grinband et al. offer no corresponding psychological or mechanistic account of this relationship, and instead simply conclude that dmFC activity “is correlated with time on task” and “is predicted by trial-to-trial differences in response time”. These limited statements risk, in turn, conceptual confusion: Is the proposal that neural activity in some way directly relates to RT? Or that dmFC monitors RT and thus scales monotonically as RT increases? Or that dmFC serves another cognitive function that, in some unspecified manner, shares a consistent relationship with RT? Thus, while we agree with Grinband et al. that it is analytically useful to consider RT as an independent variable in analyses of behavioral and neuroimaging data—and are gratified that the results of their analyses concur with our own in validating key predictions of the conflict monitoring theory (Yeung et al., 2004; Yeung and Nieuwenhuis, 2009)—we would

argue that it is vital to provide some account of the relationship between RT and observed neural activity.

This theoretical story is important in linking the reported findings to the broader literature on dMFC function. For example, there is now substantial empirical support for a central claim of the conflict monitoring theory, that detection of conflict in dMFC should lead to increased recruitment of control in lateral frontal cortex (Kerns et al., 2004; Liston et al., 2006). The time-on-task view offers no ready explanation of this critical observation. Even more directly challenging to the time-on-task view are observed dissociations between dMFC activity and measured RT. For example, dMFC activity in the Stroop task is sometimes found to be greater on congruent trials (e.g., BLUE in blue ink) than on neutral trials (e.g., XXX in blue ink), even when RT is lower in the former condition (Kadosh et al., 2008). This finding is inconsistent with the time-on-task view, but can be explained in terms of conflict between color naming and word reading at the level of whole tasks, which is increased in the presence of irrelevant stimulus attributes (Botvinick et al., 2001; Herd et al., 2006). Activity in dMFC is also consistently increased on error trials, despite the fact that errors typically occur on trials with the fastest RTs. Again, the time-on-task hypothesis provides no obvious explanation of this finding, whereas the conflict monitoring theory has been shown to account for detailed aspects of this error-related activity such as its temporal dynamics (e.g., post-response stimulus processing) and sensitivity to key experimental manipulations (Botvinick et al., 2001; Yeung et al., 2004). Thus, whereas the conflict monitoring theory provides a theoretically articulated link from the observed association between task demand and dMFC activity to the wider literature on the functional role of this region, the time-on-task view offers only a conceptually empty restatement of the bare empirical facts.

Conclusion

Grinband et al.'s critique of the conflict monitoring theory of dMFC function has several important strengths: It presents a serious and thoughtful examination of the conflict monitoring view that is accompanied by detailed analyses of the relationship between behavioral and neuroimaging measures and by a formalization of the key predictions tested. However, as detailed above, the research also has critical weaknesses: The analyses are motivated by an intuition that is superficially plausible but that is theoretically and demonstrably incorrect, and all of the presented results are entirely consistent with the conflict monitoring theory (and, indeed, were anticipated by our previously published simulation results). Given that the conflict monitoring theory predicts all of the presented results, and provides a mechanistic account of these findings that encompasses other key features of observed dMFC activity, we conclude that this account continues to provide a productive and accurate account of the contributions of dMFC to decision making and cognitive control.

References

- Bench CJ, Frith CD, Grasby PM, Friston KJ, Paulesu E, Frackowiak RSJ, et al. Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*. 1993; 9:907–922. [PubMed: 8232848]
- Botvinick MM, Braver TS, Carter CS, Barch DM, Cohen JD. Conflict monitoring and cognitive control. *Psychol Rev*. 2001; 108:624–652. [PubMed: 11488380]
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*. 1998; 280:747–749. [PubMed: 9563953]

- Grinband J, Savitsky J, Wager TD, Teichert T, Ferrera VP, Hirsch J. The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. *NeuroImage*. 2011; 57:303–311. [PubMed: 21168515]
- Herd SA, Banich MT, O'Reilly RC. Neural mechanisms of cognitive control: an integrative model of Stroop task performance and fMRI data. *J Cogn Neurosci*. 2006; 18:22–32. [PubMed: 16417680]
- Kadosh RC, Kadosh KC, Henik A, Linden DEJ. Processing conflicting information: facilitation, interference, and functional connectivity. *Neuropsychologia*. 2008; 46:2872–2879. [PubMed: 18632120]
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS. Anterior cingulate conflict monitoring and adjustments in control. *Science*. 2004; 303:1023–1026. [PubMed: 14963333]
- Liston C, Matalon S, Hare TA, Davidson MC, Casey BJ. Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron*. 2006; 50:643–653. [PubMed: 16701213]
- Yeung N, Nieuwenhuis S. Dissociating response conflict and error likelihood in anterior cingulate cortex. *J Neurosci*. 2009; 29:14506–14510. [PubMed: 19923284]
- Yeung N, Botvinick MM, Cohen JD. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev*. 2004; 111:931–959. [PubMed: 15482068]

Fig. 1.

Conflict model simulation results. Simulated conflict and error likelihood are plotted separately for corresponding models of the flanker task (A,C) and Stroop task (B,D). Data for the flanker task model are reanalyzed from previously published simulations (Yeung et al., 2004), in which trials were divided into bins corresponding to stepwise increases in simulated RT (processing cycles). The Stroop model was based on an earlier implementation of the verbal response version of the task in which errors are very rare (Botvinick et al., 2001), with parameters modified to simulate the increased error rates observed empirically in Grinband et al.'s manual response task. For this simulation, trials were divided into decile bins according to RT. For both models, simulated conflict shows a monotonic increase with RT for both incongruent and congruent trials (A,B), whereas error likelihood shows a monotonic decrease for incongruent trials in which errors were relatively frequent, and no change for congruent trials in which errors were rare or absent (C,D). Replicating Grinband et al. and our previous analyses, simulated conflict is plotted for correct response trials only.