Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information

Tobias Egner & Joy Hirsch

A prominent model of how the brain regulates attention proposes that the anterior cingulate cortex monitors the occurrence of conflict between incompatible response tendencies and signals this information to a cognitive control system in dorsolateral prefrontal cortex. Cognitive control is thought to resolve conflict through the attentional biasing of perceptual processing, emphasizing task-relevant stimulus information. It is not known, however, whether conflict resolution is mediated by amplifying neural representations of task-relevant information, inhibiting representations of task-irrelevant information, or both. Here we manipulated trial-by-trial levels of conflict and control during a Stroop task using face stimuli, while recording hemodynamic responses from human visual cortex specialized for face processing. We show that, in response to high conflict, cognitive control mechanisms enhance performance by transiently amplifying cortical responses to task-relevant information rather than by inhibiting responses to task-irrelevant information. These results implicate attentional target-feature amplification as the primary mechanism for conflict resolution through cognitive control.

To meet changing environmental demands, humans make rapid, strategic adjustments to how they deploy their attentional resources[1,2], such that when we encounter increasing task difficulty, we tend to re-focus our attention on task-relevant aspects of our surroundings while ignoring less relevant ones. However, it is not fully understood how the human brain is able to swiftly adjust processing priorities in response to changing circumstances.

In the laboratory, the strategic control of attention to optimize performance is captured by 'conflict-adaptation' effects in classic selective-attention tasks[3,4]. For example, in the Stroop task[5,6], subjects are required to name the ink color of a printed color name (the 'target' dimension of the stimulus), while ignoring the word's meaning (the 'distracter' dimension of the stimulus). When target and distracter dimensions are incongruent (for example, the word RED printed in green ink), they induce conflicting response tendencies, and reaction times are slowed in comparison to trials where target and distracter information is congruent (for example, the word GREEN printed in green ink). However, this deleterious effect of incongruent distracters on the processing of target information is reduced (and often abolished entirely) after incongruent trials, as compared to after congruent trials[7,8]. This suggests that high conflict in an incongruent trial leads to a transient upregulation of selective attention in anticipation of the next trial, resulting in improved conflict resolution: that is, conflict adaptation[4,9].

It has been proposed that such context-sensitive regulation of attentional resources is mediated by a specialized conflict-monitoring system that gauges co-activation in the processing pathways associated with incompatible responses[9–12]. When conflict is detected, the conflict monitor triggers a 'cognitive control' system that is assumed to resolve the conflict through the attentional biasing of perceptual processes[7,9,13]. To illustrate, in the context of the Stroop task, an incongruent stimulus would elicit incompatible response tendencies, leading the conflict monitor to alert the cognitive control system to the need for conflict resolution. The cognitive control system would then deploy selective attention mechanisms to bias perceptual processing toward task-relevant stimulus properties and away from task-irrelevant, distracting stimulus properties, by modulating activity in the visual pathways involved in extracting target and distracter features of the stimulus.

To dissociate the neural correlates of the conflict-monitoring and cognitive control systems, human neuroimaging studies have exploited a particularly attractive feature of conflict adaptation: namely, the possibility of comparing identical incongruent trials on the basis of whether they are associated with low control (the incongruent trial follows a congruent one) or with high control (the incongruent trial follows an incongruent one)[7,8,12,14]. By using variants of conflict adaptation to separate conflict and control processes, a number of studies have identified neural correlates of conflict detection predominantly in medial prefrontal cortex, particularly in the anterior cingulate cortex (ACC)[7,10–15]. Also in the context of conflict-adaptation tasks, correlates of cognitive control processes have mostly been localized to sites in dorsolateral prefrontal cortex (DLPFC)[7,8,13,14]. However, it is not known how the putative loci of cognitive control in DLPFC rapidly bias perceptual processing in response to trial-by-trial fluctuation in conflict. Specifically, it remains an open question, whether in tasks such as the Stroop protocol, conflict is resolved through excitatory modulation (amplifying the processing of target information), inhibitory
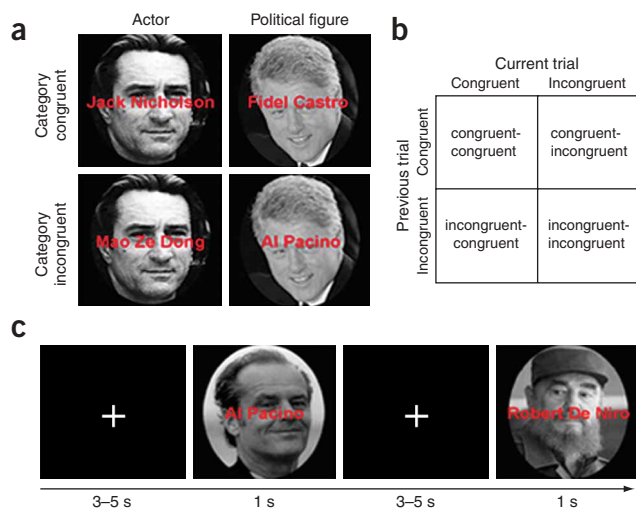
Figure 1 Experimental protocol. Subjects discriminated between actors and political figures, based on either the face stimulus (face-target condition), or the written name (face-distracter condition). (**a**) Stimuli could be either category-congruent (for example, an actor's face accompanied by an actor's name), or category-incongruent (an actor's face accompanied by a politician's name or vice versa). (**b**) Trials were presented pseudo-randomly so as to produce an equal number of each possible trial sequence with respect to previous-trial and current-trial congruency and were analyzed by trial type (congruent-congruent, congruent-incongruent, incongruent-congruent and incongruent-incongruent). (**c**) Stimuli were presented for 1,000 ms with a jittered interstimulus interval (ISI) of 3,000–5,000 ms (mean ISI = 4,000 ms). Shown is an example trial in which a congruent stimulus is followed by an incongruent stimulus, resulting in a low control–high conflict trial.

modulation (suppressing the processing of distracter information), or a combination of the two[16]. The current study aimed to complement our understanding of cognitive control by unraveling the nature of this perceptual biasing process.

Target amplification and distracter inhibition have both been described as feasible neural mechanisms of selective attention[17]. For example, neural responses are enhanced for attended spatial locations[18,19], stimulus features[20–22] or objects[23,24]. However, previous studies investigating attentional selection have typically required subjects explicitly, via external cues, to attend to a particular aspect of their visual environment. In contrast, the current study investigates which selection mechanism underlies rapid, 'online' performance adjustments arising endogenously from the subject interacting with the ongoing task. This is because conflict adaptation emerges from the task context, as defined by the stimulus history, and in the absence of any explicit external cues or instructions to the subjects to shift their focus of attention or improve their performance. Therefore, conflict adaptation serves as a model of attention regulation in many real-life situations where, typically, we lack explicit external guidance as to where we should attend or what we should attend to, for optimal task performance.

To probe the neural mechanism underlying conflict adaptation, we used functional magnetic resonance imaging (fMRI) while subjects performed a new variant of the Stroop task, involving face stimuli. The face stimuli were expected to elicit blood oxygen level–dependent (BOLD) signals in the fusiform face area (FFA), an extrastriate visual region responsible for face processing[25] that is known to be susceptible to attentional and contextual top-down modulation[26–29]. By using face stimuli as either target or distracter stimulus features in a Stroop-like task, we obtained FFA responses that provided a window into the perceptual processing of target and distracter dimensions, under

varying levels of cognitive control during conflict adaptation. We found that behavioral conflict-adaptation effects were exclusively associated with the amplified processing of task-relevant stimulus properties and not with the suppressed processing of task-irrelevant stimulus features. This enhanced cortical representation of target information in visual cortex was accompanied by an increased functional interaction with cognitive control loci in DLPFC.

RESULTS

Subjects categorized stimuli consisting of familiar faces of actors and politicians, with either category-congruent or category-incongruent names of other actors and politicians written across them (**Fig. 1a**). Trials were analyzed on the basis of previous- and current-trial congruency (**Fig. 1b**): the previous-trial congruency determined the level of 'control' on the current trial, and the current-trial congruency determined the level of 'conflict' on the current trial. Thus, successive congruent-congruent trials reflected low conflict under low control; congruent-incongruent trials reflected high conflict under low control; incongruent-congruent trials reflected low conflict under high control; and incongruent-incongruent trials reflected high conflict under high control conditions. Within this factorial design, conflict adaptation was represented by the interaction of previous- and current-trial congruency, where the effect of current-trial conflict was greater following congruent trials (that is, congruent-incongruent > congruent-congruent) than following incongruent trials (that is, incongruent-incongruent > incongruent-congruent)[4]. The critical comparison in this analysis lies in demonstrating reduced interference (conflict) from incongruent distracters under conditions of high control as compared to conditions of low control (incongruent-incongruent versus congruent-incongruent trials)[7,12].

Subjects discriminated actors from political figures (**Fig. 1c**) using a two-alternative forced-choice button press, in two experimental contexts. In one condition, they responded according to the identity of the face stimulus ('face-target' condition); in the other condition, they responded according to the written name while ignoring the face stimulus ('face-distracter' condition). This design allowed us to compare subjects' responses to identical face stimuli under conditions of low and high cognitive control, depending on whether faces represented the target or the distracter dimension of the task. If attention regulation during the Stroop task was mediated by the amplification of target processing, we would expect to observe increased FFA activation in the high-control condition as compared to the low-control condition, in the face-target condition. Conversely, if performance adjustments depended on the suppression of distracter processing, we would expect FFA activation to be inhibited under the high-control condition as compared to the low-control condition, during the face-distracter condition. Finally, if excitatory and inhibitory mechanisms contributed to optimize performance, both of the above predictions would hold.

Behavioral data: conflict adaptation

Reaction times (RTs) for correct trials showed Stroop-like interference effects from incongruent distracters, both in the face-target condition (congruent mean = 711 ms; incongruent mean = 725 ms; t_{21} = 4.0, $P < 0.001$) and in the face-distracter condition (congruent mean = 862 ms; incongruent mean = 903 ms; t_{21} = 4.1, $P < 0.001$). Furthermore, the data in both tasks bore out classic conflict-adaptation effects (**Table 1** and **Fig. 2a,b**). In the face-target task (**Fig. 2a**), RTs to incongruent stimuli were faster in the high-control condition than in the low-control condition (incongruent-incongruent < congruent-incongruent; t_{21} = 2.2, $P < 0.05$). This resulted in a previous-trial × current-trial interaction ($F_{1,21}$ = 6.7, $P < 0.02$), as current-trial conflict

under low control (congruent-incongruent > congruent-congruent; t_{21} = 3.8, P < 0.002) was abolished under high control (incongruent-incongruent ≈ incongruent-congruent; t_{21} = 0.1, P > 0.5). Similarly, in the face-distracter task (**Fig. 2b**) responses to incongruent trials were faster in the high-control condition than in the low-control condition (incongruent-incongruent < congruent-incongruent; t_{21} = 3.0, P < 0.009). This led to a reduction in current-trial conflict from the low-control condition (congruent-incongruent > congruent-congruent, t_{21} = 4.5, P < 0.001) to the high-control condition ($F_{1,21}$ = 5.5, P < 0.03), although conflict was not abolished entirely (incongruent-incongruent > incongruent-congruent; t_{21} = 2.4, P < 0.03). These effects were not related to speed-accuracy trade-offs, as the accuracy data reflected the same pattern of results (**Table 1**).

fMRI data: conflict adaptation in visual cortex

To assess how conflict adaptation was achieved at the level of perceptual processing in visual cortex, we analyzed task-related BOLD responses in individually defined regions of interest (ROIs) in the FFA (**Fig. 2c**; Methods). When faces served as target stimuli, conflict adaptation was evident in the FFA (**Fig. 2d**). FFA responses to incongruent trials were enhanced in the high-control condition as compared to the low-control condition (incongruent-incongruent > congruent-incongruent; t_{21} = 2.3, P < 0.04). This resulted in a previous-trial × current-trial interaction ($F_{1,21}$ = 7.3, P < 0.02), as a current-trial conflict effect in the low-control condition (congruent-congruent > congruent-incongruent; t_{21} = 2.8, P < 0.02) disappeared under high control (incongruent-incongruent ≈ incongruent-congruent; t_{21} = 1.5, P > 0.14). Note that FFA activation to face target stimuli showed the inverse pattern of the behavioral data, in that low behavioral interference from the name distracters was associated with high FFA activation and high behavioral interference with low FFA activation. Notably, conflict adaptation was associated with the amplification of neural perceptual responses to relevant target stimulus features.

When faces served as distracters, on the other hand, we found no effects of cognitive control on FFA responses ($F_{1,21}$ = 0.4, P > 0.5; **Fig. 2e**), suggesting that the behavioral conflict-adaptation effect obtained in this task (**Fig. 2b**) was not mediated by selective inhibition of distracter processing. To illustrate, more directly, the effects of

cognitive control on target versus distracter processing, FFA responses to only incongruent stimuli were compared under conditions of low and high control (congruent-incongruent versus incongruent-incongruent), depending on whether they constituted target or distracter features. A task × control interaction ($F_{1,21}$ = 4.9, P < 0.04; **Fig. 3a**) was characterized by an increase in target-related responses from low to high control (t_{21} = 2.5, P < 0.03), with no effects of control on distracter-related responses (t_{21} = 0.6, P > 0.5).

To corroborate that this cognitive control–related increase in activation during face-target processing was specific to the FFA and not a generic effect on high-level visual regions, we conducted a control analysis comparing responses to congruent-incongruent and incongruent-incongruent trials in the FFA to those in the parahippocampal place area (PPA), an extrastriate visual region selectively responsive to natural scenes[30] (Methods). An expected main effect of cortical region in the processing of the face stimuli (FFA > PPA, $F_{1,21}$ = 17.1, P < 0.001) was accompanied by an interaction effect ($F_{1,21}$ = 4.3,
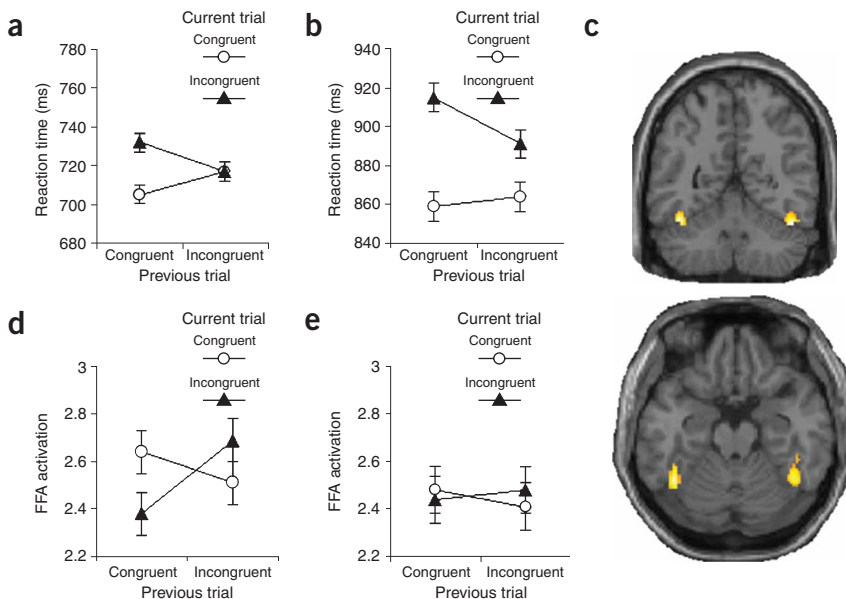
Table 1 Descriptive statistics of behavioral data

Face target condition			
	Reaction time (ms)	Standard deviation	Percentage accuracy	Standard deviation
Congruent-congruent	705	74	99.1	1.6
Congruent-incongruent	732	81	95.5	4.4
Incongruent-congruent	717	76	97.5	3.3
Incongruent-incongruent	717	77	97.2	3.1

Face distracter condition			
	Reaction time (ms)	Standard deviation	Percentage accuracy	Standard deviation
Congruent-congruent	859	116	97.6	4.6
Congruent-incongruent	915	119	94.2	4.5
Incongruent-congruent	864	122	97.5	3.4
Incongruent-incongruent	891	100	95.2	4.7

Figure 2 Conflict adaptation in behavioral and fMRI data. (**a,b**) Mean group reaction times (± s.e.m.) for current congruent and incongruent trials plotted as a function of previous-trial congruency (x-axis) for (**a**) the face-target condition and (**b**) the face-distracter condition. (**c**) Illustration of FFA activation on a rostral (top panel) and axial (bottom panel) brain slice, derived from a group analysis of the face-area localizer scan (MNI x = 46, y = –54, z = –24; 121 voxels, 968 mm[3]; x = –42, y = –54, z = –22; 85 voxels, 680 mm[3]) displayed at P < 0.05 (corrected). (**d,e**) Mean group activation values (betas ± s.e.m.) from subject-specific FFA ROIs for current congruent and incongruent trials plotted as a function of previous-trial congruency (x-axis) for (**d**) the face-target condition and (**e**) the face-distracter condition.
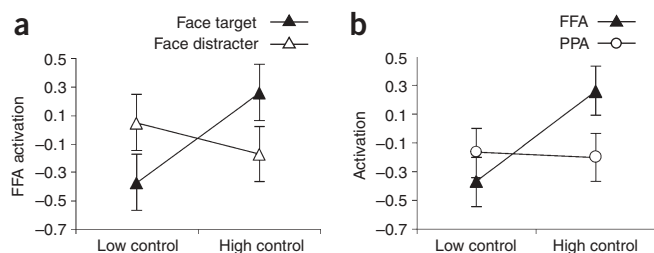
Figure 3 Task- and region-specificity of cognitive control effects. (**a**) FFA activation (normalized betas ± s.e.m.) for incongruent trials only are plotted as a function of control (low versus high) and task (face-target versus face-distracter). (**b**) Neural responses from FFA and PPA (normalized betas ± s.e.m.) are shown for incongruent trials under conditions of low versus high control during the face-target condition.

$P < 0.05$), as responses increased with control in the FFA ($t_{21} = 2.3$, $P < 0.04$) but not in the PPA ($t_{21} = 0.5$, $P > 0.5$; **Fig. 3b**). We probed the specificity of cognitive control effects further by assessing whether activity in early visual cortex (V1 and V2) was affected by task and control variables. To this end, we analyzed data from voxels in early occipital regions (all falling within V1 or V2) that were most highly activated by all visual stimuli during the face area localizer task (Methods). These early visual regions displayed no task ($F_{1,21} = 0.1$, $P > 0.7$), control ($F_{1,21} = 0.1$, $P > 0.7$), or task × control interaction effects ($F_{1,21} = 1.2$, $P > 0.2$). These results confirmed that the effect of cognitive control was specific to neuronal populations involved in the processing of task-relevant target stimulus features in the FFA and did not extend to other high level or to lower level visual processing.

fMRI data: top-down conflict resolution

Are the enhanced perceptual responses to face target stimuli under high control a result of top-down modulation from DLPFC? If this is the case, putative cognitive control regions should display increased connectivity with the FFA in the high-control condition as compared to the low-control condition; further, this effect should be limited to the face-target condition. To test these predictions, we first identified neural substrates of cognitive control during the face-target task with an incongruent-incongruent > congruent-incongruent contrast in a whole-brain group analysis. This analysis yielded clusters of activation in the right DLPFC (Brodmann's area 46), right middle temporal gyrus and left anterior insula (**Fig. 4a**). To assess the functional interaction, during conflict adaptation, between these cortical loci of cognitive control and the FFA, we then carried out a psychophysiologic interaction (PPI) analysis[31]. PPI represents a measure of context-dependent connectivity, explaining regionally specific responses in one brain area in terms of the interaction between input from another brain region and a cognitive or sensory process[31,32]. In the current study, PPI allowed us to assess if the FFA displayed context-sensitive increments in functional integration with these cognitive control ROIs when going from low- to high-control trials. We calculated the degree of functional interaction during high- versus low-control incongruent trials (incongruent-incongruent versus congruent-incongruent) for both the face-target and the face-distracter task; we then subjected the connectivity data to a task × control interaction analysis (Methods). This revealed a cluster of voxels in the DLPFC ROI (**Fig. 4b**) that showed task-specific (face target > face distracter) and control-specific (incongruent-incongruent > congruent-incongruent) increments in functional integration with the FFA. Thus, functional coupling between the right DLPFC and the FFA increased under high control in the

face-target condition, but not in the face-distracter condition—precisely as would be predicted for a region implementing conflict-sensitive top-down biasing of perceptual target-feature processing.

Conflict adaptation versus priming effects

It is controversial whether conflict-adaptation effects truly reflect conflict-driven cognitive control processes, or whether they can be accounted for by lower-level priming effects. Priming might arise from different proportions of stimulus-response repetitions (and alternations) between congruent-congruent and incongruent-incongruent trials on the one hand, and congruent-incongruent and incongruent-congruent trials on the other hand[7,8,33–35]. For instance, in the Eriksen flanker task[36], behavioral conflict-adaptation effects can be mediated entirely by particularly fast responses on congruent-congruent and incongruent-incongruent trials where identical stimulus-response pairings are repeated; this suggests that the effect may be due to repetition priming rather than to adjustments in cognitive control[33,35]. For this reason, we controlled for repetition-priming confounds by not including any direct repetitions of identical stimuli[7,8]. However, it could still be argued that the conflict-adaptation effects we obtained may have occurred because on 50% of congruent-congruent and incongruent-incongruent trials, both the target and distracter categories (that is, actor or politician) remained the same, possibly facilitating performance through some form of category priming. In contrast, on congruent-incongruent and incongruent-congruent trials, at least one category always changed. Thus, the behavioral conflict reduction (incongruent-incongruent < congruent-incongruent) and the accompanying neural target-feature amplification (incongruent-incongruent > congruent-incongruent) that we observed during face-target processing could, in theory, stem from incongruent-incongruent trials being subject to category-priming effects rather than top-down control influences. To test this alternative interpretation, we re-analyzed the data, splitting up incongruent-incongruent trials into those where target and distracter categories were repeated (repetition trials) and those where these categories alternated (alternation trials). If the previous results were driven by priming effects, repetition RTs should be faster than alternation RTs, and conflict adaptation should be observed exclusively for analyses including only repetition trials but not for analyses including only alternation trials[33,35].

Reaction times for repetition trials (mean ± s.d., 723 ± 88 ms) and alternation trials (713 ± 73 ms) did not differ ($t_{21} = 1.0$, $P > 0.3$). Note that, descriptively, responses for repetitions were actually slower than those for alternation trials. Accordingly, the conflict-adaptation
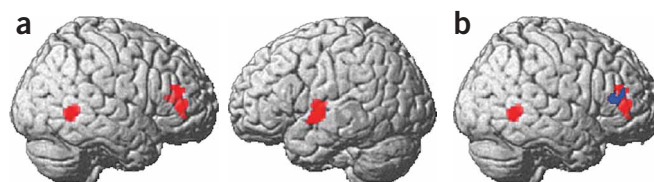

Figure 4 Regions associated with top-down control processes. (**a**) Brain areas implicated in cognitive control during conflict adaptation (incongruent-incongruent > congruent-incongruent) were identified in middle and inferior frontal gyri of the right DLPFC (MNI $x = 40$, $y = 38$, $z = 20$; 180 voxels, 1,440 mm³), right middle temporal gyrus (MNI $x = 48$, $y = -54$, $z = 0$; 153 voxels, 1,224 mm³) and left anterior insula (MNI $x = -46$, $y = -6$, $z = 4$; 216 voxels, 1,728 mm³), with a cluster threshold of $P < 0.05$ (corrected). (**b**) Within these cognitive control ROIs (red), right DLPFC exhibits voxels (blue) (MNI $x = 26$, $y = 30$, $z = 10$; 27 voxels, 216 mm³) that show a task-specific and context-specific increase in functional integration with the FFA, at a voxelwise threshold of $P < 0.05$.

1787

interaction effect (previous-trial × current-trial congruency) was evident to a similar degree when we included only repetition ($F_{1,21} = 4.2$, $P = 0.053$) or only alternation trials ($F_{1,21} = 6.2$, $P = 0.021$) in the analysis. Next, we conducted corresponding control analyses for fMRI data extracted from the FFA during face-target processing, modeling the data to include separate regressors for repetition and alternation trials. (Note that the inclusion of an additional regressor in the model inevitably alters parameter estimates for other regressors as well.) Again, FFA activation during repetition trials (mean beta ± s.d., 2.4 ± 1.6) and alternation trials (2.3 ± 1.4) did not differ ($t_{21} = 0.5$, $P > 0.6$). Conflict-adaptation effects were of similar magnitude when we included only repetition trials ($F_{1,21} = 3.6$, $P = 0.07$) or only alternation trials ($F_{1,21} = 3.1$, $P = 0.09$) in the analysis. These results show that incongruent-incongruent category-repetition and category-alternation trials contributed similarly to the overall conflict-adaptation effect in both the behavioral and fMRI data, and that these data thus cannot be explained by category priming effects, but rather can likely be attributed to conflict-driven adjustments in cognitive control.

Finally, the current results could also have been driven primarily by one of the two stimulus categories. For example, it could be argued that the actors in the stimulus set displayed a greater similarity with each other than the political figures, both in terms of their faces and their names. From the viewpoint of a priming account of conflict adaptation, higher within-category similarity would be expected to lead to stronger priming effects—that is, greater conflict adaptation. To test whether our results may have stemmed from effects within a particular subset of stimuli, we split up the data set according to whether a given trial presented the picture (or name) of an actor or a politician. If overall conflict adaptation was driven disproportionately by one stimulus category, a three-way interaction effect between stimulus category, previous-trial congruency and current-trial congruency would be observed. In the RT data, pictorial category did not interact with conflict adaptation ($F_{1,21} = 1.6$, $P > 0.2$) and neither did the name category ($F_{1,21} = 0.1$, $P > 0.8$). Next, we re-analyzed the fMRI ROI data, this time splitting up each trial-type regressor according to whether it contained the picture (or name) of an actor or a political figure. In accordance with the behavioral findings, there were no effects of stimulus category on conflict adaptation in terms of FFA activation: neither for the pictorial category ($F_{1,21} = 0.65$, $P > 0.4$), nor for the name category ($F_{1,21} = 0.1$, $P > 0.7$). These data demonstrate that the current results were not primarily driven by one particular stimulus category.

DISCUSSION

We obtained significant behavioral conflict and conflict-adaptation effects in a variant of the Stroop task that used face stimuli as either target or distracter stimulus features. In both versions of the task, responses were faster in incongruent trials that followed incongruent trials than in incongruent trials that followed congruent trials. This represented successful conflict resolution through cognitive control[7,9,12]. By simultaneously imaging BOLD responses in individuals' FFA, we showed that the strength of neural face representation varied with conflict and control under conditions where faces served as target stimuli, but not when they served as distracter stimuli. Specifically, FFA activation to face-target stimuli was increased in response to incongruent trials following incongruent trials, compared to when incongruent trials followed congruent ones. Thus, face processing was amplified when cognitive control was high (and conflict was reduced), compared to when the identical incongruent stimuli were processed under conditions of low control (and high conflict). Further, this effect was exclusive to the face-target condition. In addition, we contrasted the effect of cognitive control during face-target processing in the FFA to that in two other visual areas: a similar category-specific high-level visual area, the PPA, and early visual cortex. We showed that this amplification of neural target-feature representation was region-specific and did not represent a generic upregulation of activity in striate or extra-striate visual regions under conditions of high cognitive control.

The conflict-monitoring and cognitive control model would predict that this transient modulation of target-feature processing should be directly related to input from cognitive control loci in DLPFC[7,9,13]. We tested this hypothesis by measuring context-sensitive functional integration between the FFA and functionally defined cognitive control ROIs in right DLPFC, the left insula and right superior temporal cortex. In support of the notion that FFA modulation was mediated by a top-down biasing signal, we found that a subregion of the DLPFC ROI showed task- and control-dependent functional integration with the FFA. This subregion showed increased coupling with the FFA under conditions of high control, but only when the face stimuli served as targets for attentional selection. Therefore, along with previous studies documenting neural substrates of conflict monitoring in the ACC[7,10–15] and cognitive control in DLPFC[7,8,13,14], the current data complement an emergent neural model of cognitive control[9] by supplying the mechanism through which conflict resolution is implemented at the target site of attentional modulation: namely, through target facilitation rather than distracter inhibition. The exact way in which this neural modulation of target-feature processing is achieved during conflict adaptation raises important questions for future research. One possible mechanism for target-feature enhancement is that during the face-target condition, attentional top-down signals may lead to enhanced pre-stimulus baseline neural activity in the FFA[22,37], thus favoring this area in the competition for processing resources during subsequent stimulus processing[1,17].

Our results demonstrate that rapid, online performance adjustments in response to high conflict are mediated by amplified neural processing of task-relevant (target) stimulus features but not by inhibited processing of task-irrelevant (distracter) stimulus features. These data are in agreement with electrophysiological and neuroimaging studies that have reported enhancement of target processing with respect to cued attention shifts between different spatial locations[18,19], stimulus-features[20–22] and objects[23,24]. These previous investigations demonstrated neural target-feature enhancement in the context of explicitly (and exogenously) cued attention shifts; in contrast, the current study provides evidence suggesting that target-feature enhancement constitutes the main selection mechanism when attention regulation is driven endogenously so as to optimize performance—as is likely the case in many real-life situations. In keeping with computational models[9,16], our findings strongly suggest that performance on classic selective-attention tasks, such as the Stroop task, may be accounted for without invoking a mechanism that actively inhibits the perceptual processing of task-irrelevant stimulus features (see also ref. 38). Rather, our data are consistent with the proposal that target-feature amplification represents the primary top-down mechanism of selective attention. Both behavioral and neuroimaging data suggest that perceptual suppression of task-irrelevant (distracter) information may not be possible unless attentional resources are entirely bound up by the processing of task-relevant (target) information under highly demanding conditions[39,40].

The above interpretation of the current data hinges critically on whether conflict adaptation is truly a reflection of conflict-driven adjustments in cognitive control or results from priming effects within particular stimulus sequences[33,35]. Our study did not contain any direct stimulus repetitions and was therefore not confounded by repetition priming effects[7,8]. Higher-level priming effects, however, could feasibly

arise from repetitions of target- and distracter-feature categories, which occurred on 50% of congruent-congruent and incongruent-incongruent trials but never on congruent-incongruent and incongruent-congruent trials. In control analyses, we found no effects of category repetition on behavioral or fMRI data, and we can therefore reject this alternative interpretation of our data. Another potential confound in the conflict-adaptation effect is that, typically, on 50% of congruent-incongruent and incongruent-congruent trials, a response repetition is accompanied by a stimulus alternation[35,41]; these trials may be associated with slower responses than trials in which either the stimulus and the response both alternate or both stay the same[41]. This confound was also controlled for in the current study, as congruent-congruent and incongruent-incongruent trials contained no exact stimulus repetitions (that is, the actual stimulus always alternated), but entailed response alternations on 50% of the trials—exactly as was the case for congruent-incongruent and incongruent-congruent trials.

Although we are confident in concluding that the current data set reflects a demonstration of conflict-driven adjustments in cognitive control, the careful assessment of other factors that may contribute to such sequential trial effects remains important for gaining a better understanding of cognitive control mechanisms. For instance, it is not clear why conflict adaptation on the flanker task may be mediated entirely by priming effects[33,35], whereas in other conflict tasks— such as the Simon task[42], the color-naming Stroop task[7,8] or the current task— this does not appear to be the case. An additional question of interest is the respective contribution of (and potential interaction between) preparatory processes that arise from conflict and those that may arise from specific expectancies regarding the nature of an upcoming stimulus. Notably, the original report of the conflict-adaptation effect envisaged subjects' expectancies, rather than conflict, as the driving force behind attentional adjustments[4]. Such expectancy effects, as well as their potential interaction with conflict-driven cognitive control processes, have been successfully modeled in the context of non-conflict, two-alternative, forced-choice tasks[43,44], but they have yet to be explored within the context of the conflict-adaptation protocol.

In conclusion, we have shown that conflict adaptation, reflected in improved selective-attention performance following high-conflict trials, is mediated by the amplified neural representation of task-relevant stimulus features, but is not related to the perceptual inhibition of task-irrelevant features. We propose that attentional target-feature amplification is the neural mechanism by which cognitive control optimizes performance.

METHODS

Subjects. 22 (14 females) healthy volunteers (mean age = 28.7 years, range = 20–40 years) gave written informed consent in accordance with institutional guidelines to participate in this study.

Experimental protocols. Stimuli were presented with Presentation software (Neurobehavioral Systems, http://nbs.neuro-bs.com) and displayed with a back-projection screen that was viewed by the subjects via a mirror attached to the head-coil. The FFA localizer task was adopted from a previous study in our laboratory[29]: subjects passively viewed photographic face and house stimuli in 12 alternating blocks of 15 s, separated by 10 s resting (fixation) periods. Within each block, 15 faces or houses were presented for 750 ms, with an interstimulus interval (ISI) of 250 ms. Each run of the main task consisted of 148 presentations of photographic stimuli depicting the face of either an actor (Robert DeNiro, Al Pacino or Jack Nicholson) or a political figure (Fidel Castro, Bill Clinton or Mao Zedong), all of whom were readily identified by the subjects before the experiment. Faces were presented with congruent or incongruent names (**Fig. 1**) written across them in red letters. No face stimulus was paired with its own name. Stimuli were presented for 1,000 ms, with a varying ISI of 3,000–5,000 ms (mean ISI = 4,000 ms), in pseudo-random order (counter-

balanced for equal numbers of congruent-congruent, congruent-incongruent, incongruent-congruent and incongruent-incongruent stimulus sequences). Stimulus occurrences were counter-balanced across trial types and response buttons, and the stimulus sequence included neither immediate stimulus repetitions nor instances of 'negative priming' (where the distracter feature of one trial turns into the target feature of the next trial). Subjects were instructed to respond as fast as possible, while maintaining accuracy, by pushing response buttons corresponding to 'actor' (right index finger) or 'politican' (right middle finger). In the 'face-target' run, subjects responded according to the face dimension of the stimuli, and in the 'face-distracter' run, they responded according to the name dimension, with the order of runs counter-balanced across subjects. We analyzed the behavioral data in terms of reaction times (excluding error and post-error trials, and condition-specific outlier values of more than 2 standard deviations from the mean), and accuracy rates.

Image acquisition. Images were recorded with a GE 1.5–T scanner. Functional images were acquired parallel to the anterior commissure–posterior commissure (AC-PC) line with a T2*-weighted EPI sequence of 24 contiguous axial slices (TR = 2,000 ms; TE = 40 ms; flip angle = 60°; FoV = 190 × 190 mm, array size 64 × 64) of 4.5 mm thickness and 3 × 3 mm in-plane resolution. Structural images were acquired with a T1-weighted SPGR sequence (TR = 19 ms; TE = 5 ms; flip angle = 20°; FoV = 220 × 220 mm), recording 124 slices at a slice thickness of 1.5 mm and in-plane resolution of 0.86 × 0.86 mm.

Image analysis. All pre-processing and statistical analyses were done using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/spm2.html). Functional data were corrected for differences in slice-timing, spatially realigned to the first volume of the first run and smoothed with a Gaussian kernel of 8 mm³ FWHM. For whole-brain analyses, the realigned images were spatially normalized to the MNI template brain (re-sampled voxel size: 2 mm³) before smoothing, whereas for FFA ROI definition and related analyses, the data remained in native space. The first five volumes of each run were discarded before we built and estimated the statistical model. A 128-s temporal high-pass filter was applied to the data and models, and temporal autocorrelation in the fMRI time series was estimated (and corrected for) using a first-order autoregressive function.

For the FFA and PPA localizers, epochs of face and house stimuli were modeled with two box-car functions convolved with a canonical hemodynamic response function (HRF). Subject-specific FFA and PPA ROIs from non-normalized data were defined by voxels within the fusiform/parahippocampal gyri that displayed face > house or house > face selectivity at a voxel-wise threshold of $P < 0.0001$ (uncorrected), with a minimum cluster threshold of five contiguous voxels (202.5 mm³). Marsbar software (http://marsbar. sourceforge.net/) was used to convert these clusters into ROIs and to extract ROI data for the subsequent analyses. As a comparison region in early visual cortex, data were also extracted from the voxel (40.5 mm³) displaying peak activation to both face and house stimuli in the FFA localizer task. For the main task, regressors of stimulus events (convolved with a canonical HRF) were created for congruent-congruent, congruent-incongruent, incongruent-congruent and incongruent-incongruent trial types, with error and post-error trials modeled separately. Beta values for each regressor were extracted from individual FFA ROIs (from non-normalized data) using Marsbar and analyzed in analyses of variance (ANOVA), followed by planned comparison t-tests. For the purpose of control analyses (Results), these data were modeled once more with the incongruent-incongruent trials regressor split up into category (actor or political figure)- repetition and category-alternation trials. Also, the data were modeled another two times with each trial-type regressor split up into two regressors, separately modeling trials that contained pictures (or names) of actors versus politicians. For the whole-brain search of top-down control regions, the original model was applied to normalized data across subjects in a random-effects analysis for an incongruent-incongruent > congruent-incongruent contrast, and results were thresholded at a cluster-level $P < 0.05$ (corrected).

For the psychophysiologic interaction (PPI) analysis[31], we extracted the deconvolved time-course of FFA activity in each subject (from normalized data), based on a sphere of radius 5 mm around the peak-activation voxel from the group FFA analysis (MNI $x = 46$, $y = -54$, $z = -24$). We then calculated the product of this activation time-course and the vector of the psychological

variable of interest (incongruent-incongruent > congruent-incongruent) to create the psychophysiological interaction term. New SPMs were computed for each subject, including, as regressors, the interaction term, the physiological variable (that is, the FFA activation time course) and the psychological variable. We then identified areas where activation was predicted by the psychophysiological interaction term, with the FFA activity and the psychological regressor treated as confound variables. These analyses were carried out separately for both the face-target and face-distracter tasks. Individual PPI SPMs were then entered into a random-effects group analysis contrasting connectivity patterns between face-target and face-distracter conditions with a paired-samples *t*-test, within the cognitive control ROIs and thresholded at $P < 0.05$ (uncorrected) with a cluster size of >5 voxels (40 mm^3).

1. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
2. Miller, E.K. & Cohen, J.D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
3. Logan, G.D. & Zbrodoff, N.J. When it helps to be misled: facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Mem. Cognit.* **7**, 166–174 (1979).
4. Gratton, G., Coles, M.G. & Donchin, E. Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* **121**, 480–506 (1992).
5. Stroop, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643–662 (1935).
6. MacLeod, C.M. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* **109**, 163–203 (1991).
7. Kerns, J.G. *et al.* Anterior cingulate conflict monitoring and adjustments in control. *Science* **303**, 1023–1026 (2004).
8. Egner, T. & Hirsch, J. The neural correlates and functional integration of cognitive control in a Stroop task. *Neuroimage* **24**, 539–547 (2005).
9. Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
10. Carter, C.S. *et al.* Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* **280**, 747–749 (1998).
11. Carter, C.S. *et al.* Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proc. Natl. Acad. Sci. USA* **97**, 1944–1948 (2000).
12. Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S. & Cohen, J.D. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* **402**, 179–181 (1999).
13. MacDonald, A.W., III, Cohen, J.D., Stenger, V.A. & Carter, C.S. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* **288**, 1835–1838 (2000).
14. Durston, S. *et al.* Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *Neuroimage* **20**, 2135–2141 (2003).
15. Casey, B.J. *et al.* Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. *Proc. Natl. Acad. Sci. USA* **97**, 8728–8733 (2000).
16. Cohen, J.D., Dunbar, K. & McClelland, J.L. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* **97**, 332–361 (1990).
17. Kastner, S. & Ungerleider, L.G. Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* **23**, 315–341 (2000).
18. Heinze, H.J. *et al.* Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* **372**, 543–546 (1994).
19. Luck, S.J., Chelazzi, L., Hillyard, S.A. & Desimone, R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* **77**, 24–42 (1997).
20. Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L. & Petersen, S.E. Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J. Neurosci.* **11**, 2383–2402 (1991).
21. Treue, S. & Maunsell, J.H. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
22. Chawla, D., Rees, G. & Friston, K.J. The physiological basis of attentional modulation in extrastriate visual areas. *Nat. Neurosci.* **2**, 671–676 (1999).
23. O'Craven, K.M., Downing, P.E. & Kanwisher, N. fMRI evidence for objects as the units of attentional selection. *Nature* **401**, 584–587 (1999).
24. Serences, J.T., Schwarzbach, J., Courtney, S.M., Golay, X. & Yantis, S. Control of object-based attention in human cortex. *Cereb. Cortex* **14**, 1346–1357 (2004).
25. Kanwisher, N., McDermott, J. & Chun, M.M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
26. Wylie, G.R., Javitt, D.C. & Foxe, J.J. Don't think of a white bear: an fMRI investigation of the effects of sequential instructional sets on cortical activity in a task-switching paradigm. *Hum. Brain Mapp.* **21**, 279–297 (2004).
27. O'Craven, K.M. & Kanwisher, N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013–1023 (2000).
28. Cox, D., Meyers, E. & Sinha, P. Contextually evoked object-specific responses in human visual cortex. *Science* **304**, 115–117 (2004).
29. Summerfield, C., Egner, T., Mangels, J. & Hirsch, J. Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cereb. Cortex,* published online 13 July 2005 (doi:10.1093/cercor/bhi129).
30. Epstein, R. & Kanwisher, N. A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998).
31. Friston, K.J. *et al.* Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* **6**, 218–229 (1997).
32. Friston, K.J. Functional integration in the brain. in *Human Brain Function* 2nd edn. (eds. Frackowiak, R.S. *et al.*) 971–997 (Academic Press, San Diego, 2004).
33. Mayr, U., Awh, E. & Laurey, P. Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* **6**, 450–452 (2003).
34. Botvinick, M.M., Cohen, J.D. & Carter, C.S. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* **8**, 539–546 (2004).
35. Nieuwenhuis, S. *et al.* Accounting for sequential effects in the flanker task: Conflict adaptation or associative priming? *Mem. Cognit.* (in the press).
36. Eriksen, B.A. & Eriksen, C.W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).
37. Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R. & Ungerleider, L.G. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* **22**, 751–761 (1999).
38. Egner, T. & Hirsch, J. Where memory meets attention: neural substrates of negative priming. *J. Cogn. Neurosci.* **17**, 1774–1784 (2005).
39. Lavie, N. Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 451–468 (1995).
40. Rees, G., Frith, C.D. & Lavie, N. Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* **278**, 1616–1619 (1997).
41. Hommel, B., Proctor, R.W. & Vu, K.P. A feature-integration account of sequential effects in the Simon task. *Psychol. Res.* **68**, 1–17 (2004).
42. Sturmer, B., Leuthold, H., Soetens, E., Schroter, H. & Sommer, W. Control over location-based response activation in the Simon task: behavioral and electrophysiological evidence. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 1345–1363 (2002).
43. Cho, R.Y. *et al.* Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cogn. Affect. Behav. Neurosci.* **2**, 283–299 (2002).
44. Jones, A.D., Cho, R.Y., Nystrom, L.E., Cohen, J.D. & Braver, T.S. A computational model of anterior cingulate function in speeded response tasks: effects of frequency, sequence, and conflict. *Cogn. Affect. Behav. Neurosci.* **2**, 300–317 (2002).