

**Defensive motivation increases conflict adaptation through local changes in
cognitive control: evidence from ERPs and mid-frontal theta**

Qian Yang^a, Katharina Paul^a, & Gilles Pourtois^a

*^aDepartment of Experimental Clinical and Health Psychology, Ghent University, Ghent,
Belgium*

Corresponding author:

Qian Yang

Department of Experimental Clinical & Health Psychology, Ghent University

Henri Dunantlaan 2

9000 Ghent

Belgium

Email address: qianyang.yang@ugent.be

Phone number: +32 9 2649144

Defensive motivation increases conflict adaptation through local changes in cognitive control: evidence from ERPs and mid-frontal theta

Abstract

Cognitive control is highly dynamic, and liable to variations in the affective state of participants. Recently, we found that defensive motivation, elicited by means of loss-related feedback contingent on task performance, actually increased conflict adaptation at the behavioral level, and hence tightened cognitive control. However, it remains unclear at which stage during stimulus processing this facilitatory effect takes place, and what his electrophysiological manifestation may be. To address this question, in the current study, we compared conflict adaptation between two conditions that differed in the amount of defensive motivation, and recorded 64-channel electroencephalography concurrently. Results showed that conflict adaptation was larger at the behavioral level when defensive motivation was elicited. Interestingly, event-related brain potentials showed that this effect was captured by a systematic amplitude modulation of the conflict-related N2 component, suggesting that defensive motivation could alter conflict processing locally and at an early stage following stimulus onset. In comparison, mid-frontal theta (MFT) power was globally augmented when defensive motivation was elicited, but did not co-vary with conflict adaptation however. Taken together, these neurophysiological results suggest that defensive motivation can exert specific facilitatory effects on cognitive control (N2), which can be dissociated from a more global alteration in information processing that likely reflects unspecific control or even motivational changes (MFT).

Keywords: defensive motivation; conflict processing; cognitive control; N2; P3; mid-frontal theta (MFT)

1. Introduction

At the behavioral level, conflict adaptation is reflected by faster response times on high-conflict trials when they follow another high-conflict trial compared to a low-conflict one. This reaction time facilitation translates enhanced cognitive control stemming from the prefrontal cortex (PFC; Egner, 2007). Following the dominant conflict monitoring theory (Botvinick et al., 2001), conflicting information deriving from the previous trial triggers a phasic boost in putative cognitive control mechanisms that allow to deal more efficiently and quickly with conflict information at the current trial level (Forster et al., 2011; Whitehead, Brewer, & Blais, 2017; Larson, Clayson, & Baldwin, 2012b). Recently, we found that defensive motivation increased conflict adaptation (Yang & Pourtois, 2018). Defensive motivation was elicited by means of loss-related feedback contingent on task performance (Lang & Bradley, 2013). More specifically, the conflict adaptation effect (CAE) was increased in blocks where incorrect or slow responses were punished with monetary loss, compared to a control condition where it was not. Subjective ratings and psychophysiological measures confirmed that defensive motivation was reliably elicited in those blocks where loss-related punishment was used. This finding is compatible with earlier behavioral studies showing that the CAE is usually increased by negative affect (either negative mood or defensive motivation) although this facilitatory effect appears to depend on how it is actually induced and specifically linked to cognitive control (Kuhbandner & Zehetleitner, 2011; van Steenbergen, Band, & Hommel, 2009, 2010; Schuch & Koch, 2015; Schuch et al., 2017; see Yang & Pourtois, 2018 for a systematic review of the existing literature on this topic).

However, it remains unclear through which mechanism negative affect, and more specifically defensive motivation, may facilitate conflict adaptation. A recent neurobiological model emphasizes close ties between negative affect, pain, anxiety and cognitive control, with a pivotal role of the dorsal anterior cingulate cortex therein (dACC, see Shackman et al.,

2011). More specifically, cognitive control could be heightened because of a common neurobiological ground in dACC with negative affect. Moreover, recent behavioral studies have shown that conflicting stimuli, such as Stroop stimuli, are perceived as being inherently aversive, even when implicit measures are used (Dreisbach & Fischer, 2015, 2016; Inzlicht, Bartholow, & Hirsh, 2015; Dreisbach & Fischer, 2012a; Schouppe, et al., 2015). For instance, compared to non-conflicting stimuli, conflicting stimuli are preferentially associated with negative affect during evaluative priming (Dreisbach & Fischer, 2015, 2016). Accordingly, the aversive nature of conflict could potentiate trial-by-trial adjustments, especially if negative affect (being also aversive by definition) is elicited concurrently.

Interestingly, recently, some authors have further suggested that the subjective experience associated with conflict processing is actually necessary to observe sequential effects during conflict processing, thereby confirming indirectly a close relationship between affect at the subjective level and cognitive control (Desender, van Opstal, & van den Bussche, 2014; Questienne et al., 2016; Fröber et al., 2017). Negative affect could swiftly influence cognitive control (Schuch & Koch, 2015), as conflicting stimuli share some similarities with negative affect in the sense that they are more salient and aversive. Therefore, negative affect could facilitate conflict processing by easing the encoding of incongruent trials. In agreement with this view, some studies found that loss-related feedback increased conflict adaptation (van Steenbergen et al., 2009; van Steenbergen, Band, & Hommel, 2012). From a motivational perspective, cognitive control could be influenced by specific, albeit indirect, evaluative processes that might be transiently reinforced by the encounter of negative affect (Inzlicht et al., 2015), for example by adopting a more controlled and analytic processing mode that helps to utilize and resolve conflict better (Gray, 2004; Schuch & Koch, 2015). Moreover, according to the adaptation-by-binding (ABB) model (Verguts & Notebaert, 2009; Abrahamse et al., 2016), motivational processes resulting from the encounter of punishment

(i.e. monetary loss) could facilitate conflict adaptation by strengthening the association between previous conflict processing and the currently active representation.

Although plausible, these explanations remain largely speculative and mostly based on behavioral results that do not allow to draw strong inferences about the locus of the facilitatory effect created by negative affect during conflict-driven adaptive control processes. The goal of the current study was to address this question by manipulating defensive motivation and capitalizing on the high temporal resolution provided by electroencephalography (EEG). More specifically, we harnessed well-defined components of conflict processing (in the time domain) as well as cognitive control (in the time-frequency domain) to gain insight into the facilitatory effect created by defensive motivation on the CAE.

At the event-related brain potential (ERP) level, conflict processing is usually related to the N2 and N450 components (Larson, Clayson, & Clawson, 2014; McKay, van den Berg, & Woldorff, 2017). Whereas the N2 reflects an early stage of conflict detection (Folstein & van Petten, 2008; Yeung, Botvinick, & Cohen, 2004; Yeung & Cohen, 2006), the N450 usually captures more elaborate stages of conflict processing, or even conflict resolution (Larson, Kaufman, & Perlstein, 2009). In agreement with this framework, if “simple” stimuli are used, such as provided by the Eriksen flanker or the Simon task where stimulus processing at the behavioral level is associated with shorter reaction times in general than more complex tasks (e.g. Stroop task), then N2 amplitude effects are usually observed for them, translating an early conflict processing. However, if more complex stimuli and task requirements are used such as a Stroop task, where stimulus processing at the behavioral level is usually associated with longer reaction times, then N450 effects are visible for conflicting processing, sometimes in the absence of earlier N2 effects (Larson et al., 2009). In both cases, it is assumed that a

larger (more negative) amplitude translates higher levels of conflict-related processing taking place in the dACC (Clayson & Larson, 2011a; Beste et al., 2012).

In this framework, conflict adaptation, which is usually associated with a reaction time facilitation for incongruent-Incongruent (iI) relative to congruent-Incongruent (cI) trial sequences, produces smaller (less negative) N2 and/or N450 amplitudes in iI compared to cI trials (Clayson & Larson, 2011b; Freitas, Banai, & Clark, 2009; Larson et al., 2012a, 2012b, 2013, 2015, 2016). Notably, this reduction in the amplitude of the ERP component during conflict adaptation has been reported more systematically at the N2 than N450 level. A reason for this asymmetry is that the N450 is usually more sensitive to complex monitoring effects occurring mostly at the current trial level, as opposed to lingering (control) effects from the previous trial that are more likely to influence the preceding N2 component (McKay et al., 2017; Larson et al., 2009, 2016). Hence, the N2 and N450 components likely reflect different cognitive control processes and are probably differently related to the CAE (Tillman & Wiens, 2011).

Besides ERPs that capture neural events time-locked and phased-locked to the onset of specific events (such as stimuli or response), time-frequency analyses of the EEG signal can provide additional information about the actual dynamics of information processing during conflict processing, and more specifically cognitive control (Cohen, 2014). Unlike ERPs, induced effects that slightly vary in phase across trials can be revealed with this method. In this context, mid-frontal theta (MFT) activity (4-8 Hz) has recently been related to cognitive control (Cavanagh & Shackman, 2015). Specifically, MFT power increases when encountering conflicting compared to non-conflicting trials (Sauseng et al., 2010; Cavanagh & Frank, 2014; Nigbur, Ivanova, & Stürmer, 2011; Cohen & Cavanagh, 2011), whereas the N2 and N450 ERP components are related to conflict processing per se. Moreover, MFT power is thought to reflect a more general index of cognitive control, for which conflict

processing is just one instance (Cohen & Donner, 2013). This explains why besides conflict, response errors, punishment, novel stimuli, unexpected or negative evaluative feedback usually all lead to enhanced MFT power (Cavanagh & Shackman, 2015). Given these properties, MFT has been proposed to coordinate brain activity in a distributed fronto-striatal network, which ultimately serves to monitor and adjust ongoing behavior in response to conflicts or challenges, and hence contributes to leverage cognitive control (Cavanagh, Cohen, & Allen, 2009; Cohen 2011).

In the current study, we sought to shed light on the neural mechanisms through which defensive motivation strengthens conflict adaptation, using these complementary electrophysiological measures (N2/N450 ERP components and MFT). To this end, we used a 4-alternative forced choice (AFC) Stroop task that was previously shown to be adequate to capture the facilitatory effect of defensive motivation on conflict adaptation at the behavioral level (see Yang & Pourtois, 2018). Hence, using a stringent within-subject design, we compared conflict adaptation at the behavioral and EEG levels between blocks where defensive motivation was induced (i.e. punishment blocks) or not (i.e. neutral blocks). At the behavioral level, we formulated the following hypotheses. First, we reckoned that reactive control responsible for the CAE should be blocked/reduced by the use of trial-by-trial evaluative feedback used in neutral blocks and devoid of a specific motivational value. In short, this intervening feedback should interfere with the short-lived reactive control process resulting from conflict processing happening at the previous trial level and influencing the current one (see Egner, Ely, & Grinband, 2010; Alpay, Goerke, & Stürmer, 2009). More importantly, we predicted that defensive motivation (by coupling negative feedback to monetary loss during punishment blocks) would help to overcome this interference and produces a significant CAE. Hence, we hypothesized no CAE in neutral blocks, but a significant one in punishment blocks; a result, if confirmed, that would replicate our previous

behavioral findings (Yang & Pourtois, 2018). At the ERP level, we hypothesized that either the N2 or N450's amplitude could capture this facilitatory effect of defensive motivation on conflict adaptation. More specifically, we surmised a smaller N2 and/or N450 component during conflict adaptation when participants were confronted with punishment compared to the control condition where this manipulation was omitted (i.e. neutral blocks).

Moreover, we also assessed whether conflict adaptation influenced MFT power. However, we predicted that this neurophysiological activity might be less sensitive to capture conflict adaptation in the current context where defensive motivation was elicited than standard ERP components (such as the N2), given that a previous study using intervening evaluative feedback (as used here) actually failed to link MFT to conflict adaptation (van Steenbergen et al., 2012). In addition, we also analyzed MFT power changes at the stimulus level (being either congruent or incongruent) depending on the immediately preceding feedback (being positive or negative, and punishment related or not), to better characterize the possible modulation of cognitive control by defensive motivation. More specifically, this second analysis allowed us to pit two alternative accounts against each other. Presumably, defensive motivation could enhance cognitive control “globally” (Pessoa, 2008), increasing MFT power following feedback associated with monetary loss, irrespective of congruency though. Alternatively, negative feedback could sharpen conflict processing selectively. In this case, MFT power should increase for incongruent compared to congruent trials, when they are delivered in punishment-related blocks selectively. Hence, this second analysis of MFT power enabled us to disentangle a general (block-wise) change in cognitive control elicited by defensive motivation from a local and condition specific effect on conflict processing per se.

In addition, we also extracted and compared brain responses to the evaluative feedback provided on a trial-by-trial basis, carrying presumably a different motivational value in punishment-related compared to neutral blocks. To this aim, we extracted standard ERP

components, namely the feedback-related negativity (FRN) and P3 (Santesso et al., 2011; Yeung & Sanfey, 2004), and compared their amplitudes between the two main conditions. We hypothesized that negative feedback provided in punishment blocks could lead to larger FRN and/or P3 amplitude compared to the one provided during neutral blocks, given the enhanced motivation significance in the former condition (Koban & Pourtois, 2014). Similarly, we also analyzed MFT power to the feedback, predicting a larger power for negative than positive feedback, especially when it was delivered during punishment-related compared to neutral blocks (Cavanagh & Shackman, 2015).

2. Materials and methods

2.1 Participants

Thirty-four participants (all native Dutch speakers) took part in this EEG experiment. Six participants had to be excluded from the analyses: four had a mean accuracy lower than 60% and two others had noisy EEG. These exclusions resulted in a final sample of twenty-eight participants (mean age = 21.36 years, SD = 2.9, 7 males) available for subsequent analyses. Participants lost between 3.4 and 4.9 Euro (average: 4.15 Euro, SD = 0.35) during the experiment (i.e. punishment blocks), leading to a compensation of 25-26.5 Euro (see procedure here below). All participants were right-handed and had normal or corrected-to-normal vision, and no history of psychiatric or neurological disorders. The study was approved by the local ethics committee (Faculty of Psychology and Educational Sciences, Ghent University).

2.2 Stimuli and task

Participants were seated in front of a 21-inch CRT computer monitor and performed a speeded Stroop task that was designed previously to control for possible confounding effects of stimulus repetition and contingency learning on conflict adaptation (Weissman, Jiang, &

Egner, 2014). On each and every trial, they were instructed to identify the color in which the word appeared on the screen as fast and accurately as possible by using four predefined keys tagged with corresponding colors (red, blue, green, yellow) of a response box. To do so, participants used their left middle finger to respond to red color, left index finger to blue color, right middle finger to yellow color, right index finger to green color.

The Stroop stimuli consisted of four words (in Dutch; “rood”/red: 3.81 °, “blauw”/blue: 4.66 °, “groen”/green: 4.58 °, or “geel”/yellow: 3.45 °) presented in one out of these four possible colors (red, RGB: 255, 0, 0; blue, RGB: 0, 176, 240; green, RGB: 0, 255, 0; yellow, RGB: 255, 255, 0). For a given participant, each word was presented in only two of the four possible hues however (see below). To rule out contingency learning, a four-alternative forced choice (4-AFC) task was used (Schmidt & Weissman, 2014; Weissman et al., 2014), where two pairs of S-R were created arbitrarily to balance congruent and incongruent trials. Each pair consisted of two words and two colors such that incongruent trials were created for the (incompatible) word-color association within each pair, but not across pairs however. According to this rule, 8 stimuli types were created in total (instead of 16 if all combinations were constructed), corresponding to 4 stimuli for congruent trials and 4 stimuli for incongruent trials. Each word was presented equally often in the congruent and incongruent color in each block within each mapping (Mordkoff, 2012). To rule out feature repetitions across successive trials, the stimuli were systematically alternated across successive trials to ensure that there was no stimulus (or response) repetition for both goal-relevant (color) and goal-irrelevant (meaning) dimensions.

Each trial started with a fixation cross that was used as ITI, with a mean duration of 500ms (range: 400-600ms). After the fixation cross, the Stroop stimulus was presented in the middle of the screen for 1000ms or until a response was given, followed by a blank screen shown for 700ms, before feedback was provided. Participants received either a negative

feedback (black cross) if the response was incorrect or too slow (i.e., above the response deadline corresponding to an arbitrary time limit; see here below), or a positive feedback (black tick mark) if the response was correct and fast enough (i.e., falling below this time limit; see Figure 1). With regard to the time limit, we used an algorithm validated previously that enforces fast responding (Vocat, Pourtois, & Vuilleumier, 2008; Aarts & Pourtois, 2010), and ensures a balanced proportion of positive and negative feedback on average without yielding excessive frustration (see Table 1 for the proportion of negative feedback for each condition separately; see also Yang & Pourtois, 2018). Unknown to participants, the reaction time (RT) cutoff was updated on a trial-by-trial basis to deal with unwanted fatigue or habituation effects throughout the experimental session.

Insert Fig 1 about here

Insert Table 1 about here

2.3 Procedure

The experimental session consisted of 10 blocks of 81 trials each. There were two different block types (5 per condition): punishment vs. neutral denoting the presence of negative feedback associated with monetary loss or negative feedback without any consequence, respectively. The order of block types was randomized across participants. The procedure was the same for the two block types, with the notable exception that punishment blocks included monetary losses in case of negative feedback encountered (i.e., errors or slow RTs). More specifically, participants were informed that unlike neutral blocks where incorrect

or slow responses had no specific consequences, each negative feedback received during punishment blocks would be converted into a small monetary loss. They were told that positive feedback would not/never be rewarded (with monetary gains), hence defensive motivation, but not approach motivation, was elicited during these five blocks. At the end of each punishment block (n=5), a general feedback was provided indicating ‘the number of trials associated with too slow RTs’, ‘the number of trials associated with response errors’, and thereby ‘the total amount of money lost’. For neutral blocks, no general feedback was provided. Prior to the start of the first block, they were informed that they could get a maximum payoff of 30 Euro in case they would perform flawlessly (i.e. not receiving any negative feedback during punishment blocks), but every negative feedback received during punishment blocks would reduce this total amount by 1.7 cent each time. At the start of each block, participants were encouraged to make accurate and fast responses. In between blocks, self-paced breaks were allowed. Stimuli were shown in a pseudo-random order within each block to lead on average to the same number of congruent-Congruent (cC), congruent-Incongruent (cI), incongruent-Congruent (iC), incongruent-Incongruent (iI) trials used to compute offline conflict adaptation at the behavioral and EEG levels. Stimuli presentation and data recording were controlled using E-Prime (Version 2.0; Psychology Software Tools Inc., Sharpsburg, PA).

After having signed the informed consent, participants were tested in a dimly lit, sound-attenuated, and electrically shielded cabin. They started with the practice session that consisted of four blocks comprising 12 trials each, and then completed the testing session.

2.4 Questionnaires

2.4.1 Positive and Negative Affect Schedule

A Dutch version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988; Engelen et al., 2006) was used to measure changes in affect between the two main block types. The PANAS thus served as main manipulation check regarding the expected increase in negative affect when encountering loss-related feedback (punishment). The scale consists of 20 words describing different feelings and emotions (10 – item for negative affect, 10 – item for positive affect). In order to measure participants' negative affect, participants were asked to report their subjective feelings at the end of each block, as well as at the end of the practice session for the first time (hence, 11 times in total), by rating the items on a 5-point scale ranging from 1 – *Very slightly or not at all* to 5 – *Extremely*. The order of these 20 affective words was changed across the 11 measurement points to avoid the use of a specific response strategy.

2.4.2 Stimulus ratings

To test if conflict-related trial (i.e. incongruent trials) were perceived as negative or aversive, compared to non-conflict ones (i.e. congruent trials), participants were asked to rate each visual stimulus they responded to during the main task by means of a Visual Analog Scale (VAS) ranging from 0 (not at all) to 100 (a lot) along a putative like-dislike continuum. There were eight stimulus types in total, divided into four congruent and four incongruent trials. These ratings were presented for the first time after the practice session and subsequently after each block (hence, 11 times in total). In analogy with the PANAS, the order of the eight stimuli was changed across the 11 measurements to minimize the use of a specific response strategy.

2.5 Behavioral data analysis

2.5.1 Manipulation check

The values of negative and positive affect were obtained from the sum of scores on negative and positive items, respectively. The resulting PANAS values were then submitted to a repeated measures analyses of variance (ANOVA) with Block type (punishment vs. neutral) and Affect (negative, positive) as within-subject factors. Mean values of subjective ratings for the congruent (n=4) and incongruent stimuli (n=4) were first computed. The resulting VAS values were then submitted to an ANOVA with Block type (punishment, neutral) and Congruency (congruent, incongruent) as within-subject factors.

2.5.2 Conflict adaptation

For each subject separately, the first trial of each block (1.23%), any subsequent error trials (punishment block: 22.41%, neutral block: 23.26%), post-error trials (punishment block: 16.73%, neutral block: 17.54%) and trials where RTs were larger than 3 SDs above or below the mean RT (punishment block: 0.65%, neutral block: 1.21%) were excluded from further analyses. These criteria led to the exclusion of 41.02% of the trials in punishment blocks and 43.24% of the trials in neutral blocks. These seemingly high percentages were explained by the fact that post-error trials were excluded too. Further, we ascertained that the number of trials retained for each condition was sufficient and balanced between conditions to compute conflict adaptation at the behavioral and EEG levels (see here below for actual number of trials included per condition in the statistical analyses). Next, mean RTs (and error rates) were submitted to an ANOVA with Block type (punishment, neutral), Previous congruency (congruent, incongruent), and Current congruency (congruent, incongruent) as within-subject factors.

2.6 EEG data recording

EEG was recorded at a sampling rate of 512 Hz from 64 Ag-AgCl (silver-silver chloride) electrodes mounted in an elastic cap, and using an ActiveTwo Biosemi system (<http://www.biosemi.com>). The EEG signal was referenced online to the Common Mode Sense (CMS)-Driven Right Leg (DRL) ground. The electrodes placement was made in agreement with the extended International 10-20 EEG system. The horizontal and vertical eye movements (electro-oculogram, EOG) were recorded by four additional electrodes, two positioned above and below the left eye; and two on each of the outer canthus. Two electrodes were also placed on the two mastoids (left and right).

2.7 EEG Data analyses

The EEG was preprocessed offline with EEGLAB 14.1.1 (Delorme & Makeig, 2004), implemented in Matlab R2014b, and included a 0.05/40 Hz high/low pass filter, re-referencing to the mastoids and the manual rejection of ICA components related to eye movements and blinks. Individual epochs were extracted from -400 to 1000ms around the stimulus onset (baseline corrected -200 to 0ms), and from -700 to 700ms around the feedback onset (baseline corrected -200 to 0ms). A semi-automatic artefact correction procedure was applied to eliminate trials with voltage values exceeding $\pm 90 \mu\text{V}$. For the stimulus-locked ERPs (N2 and N450), in punishment blocks, 61 trials for cC (SD = 11.82), 59 trials for cI (SD = 11.92), 59 trials for iC (SD = 11.34), and 57 trials for iI (SD = 12.01) were retained for averaging. In neutral blocks, 59 trials for cC (SD = 10.49), 55 trials for cI (SD = 12.56), 57 trials for iC (SD = 12.37), and 58 trials for iI (SD = 11.68) were included in the corresponding averages. For the feedback-locked ERPs analyses (FRN and P3), in punishment blocks, 147 trials (SD = 18.31) for negative and 156 trials (SD = 20.62) for positive feedback were used

for averaging. In neutral blocks, 149 trials (SD = 21.07) for negative and 148 trials (SD = 20.58) for positive feedback were retained for averaging.

2.7.1 Stimulus-locked ERPs: N2, N450

The N2 was quantified as the mean amplitude within the 280-320 ms post-stimulus onset interval over frontocentral electrodes (Fz and FCz pooled together)¹. The N450 was quantified as the mean amplitude within 400-500 ms post-stimulus onset interval at the same locations as the N2. For each subject, we extracted these two conflict-related ERP components for each condition separately. The N2 and N450 amplitudes were each submitted to an ANOVA with Block type (punishment, neutral), Previous congruency (congruent, incongruent), and Current congruency (congruent, incongruent) as within-subject factors.

2.7.2 Feedback-locked ERPs: P3, FRN

The FRN was quantified at FCz as the difference between the most negative peak (in the 250-350 ms post-feedback onset interval) and the average voltage of the preceding (P2) and the following positive peaks (P3), in order to control for possible confounding effects of the overlapping positive components surrounding the N2, as suggested in previous ERP studies (Sallet, Camille, & Procyk, 2013; see also Gheza et al., 2018 for a comparison of the different scoring methods of the FRN and their implications). The P3 was quantified at Pz as the mean amplitude within the 300-400 ms post-feedback onset interval (Severo et al., 2017). For each subject, we extracted FRN and P3 for each condition separately. FRN and P3 amplitudes were each submitted to an ANOVA with Block type (punishment, neutral) and Feedback valence (negative, positive) as within-subject factors.

¹ Time windows and channel locations for ERP analyses were based on the ERP-specific maximal amplitudes extracted from the grand average for all conditions collapsed (see Figures 3 and 4). As visible from Figure 3, the N2 peaked around 300 ms, while the N450 peaked around 450 ms following stimulus onset. As visible from Figure 4, the FRN peaked around 300 ms, and the P3 peaked around 350 ms following feedback onset.

2.8 Time frequency (TF) analysis

The time frequency decomposition was performed by convolving single-trial data with complex Morlet wavelets, defined as:

$$e^{i2\pi ft} e^{-t^2/(2\sigma^2)},$$

where t is time, f is frequency which ranged from 3 to 30 Hz in 50 logarithmically spaced steps, and σ is the width of each frequency band defined as $n/(2\pi f)$, where n is a number of wavelet cycles that varied from 4 to 8 in logarithmically spaced steps to obtain comparable frequency precision at low and high frequencies. Instantaneous power was estimated as the square of the complex convolution signal $Z(\text{power} = \text{real}[z(t)]^2 + \text{imag}[z(t)]^2)$ and averaged across trials. Power values at each time-frequency point were normalized by converting to the decibel scale to account for power-law scaling of oscillations in different frequency bands (amplitude increases when frequency decreases) by using the formula:

$$10\log_{10}(\text{power}/\text{baseline}),$$

power from -500 to -200 (-300 to -100) ms pre-feedback (pre-stimulus) served as the frequency band-specific baseline.

2.8.1 Conflict adaptation

Before applying the time frequency decomposition, we first removed the ERP at the single trial level to retain the induced component (i.e. non-phase locked) of MFT only, and this way ascertain that changes in MFT power across conditions were not merely confounded by the superposition of specific ERP effects (Cohen & Donner, 2013). The resulting induced MFT power was extracted within the 300-600 ms post-stimulus onset interval at FCz². For

² Time windows and channel locations for TF analyses were based on the band-specific maximal power extracted from the grand average for all conditions (see Figures 5-6). As visible from Figure 5/6, Theta power peaked around 450 ms following stimulus onset and around 300 ms following feedback onset.

each subject, MFT power was extracted for each condition separately, and was submitted to an ANOVA with Block type (punishment, neutral), Previous congruency (congruent, incongruent), and Current congruency (congruent, incongruent) as within-subject factors.

2.8.2 Influence of feedback on conflict processing

Induced MFT power was extracted at FCz within the 300-600 ms post-stimulus onset interval following negative and positive feedback separately. For each subject, MFT power was extracted for each condition separately, and was submitted to an ANOVA with Block type (punishment, neutral), Feedback valence (negative, positive), and Congruency (congruent, incongruent) as within-subject factors.

2.8.3 Feedback processing

MFT power (4-8Hz) was extracted within the 200-400 ms post-feedback onset interval at FCz. For each subject, MFT power was extracted for each condition separately, and was submitted to an ANOVA with Block type (punishment, neutral) and Feedback valence (negative, positive) as within-subject factors.

A standard alpha level of .05 was used for all statistical tests.

3. Results

3.1 Manipulation check

3.1.1 PANAS

The ANOVA showed a significant main effect of Affect, with higher subjective ratings for positive ($M = 24.29$, $SEM = 1.47$) than negative affect ($M = 17.65$, $SEM = 1.06$), $F(1, 27) = 12.499$, $p = .001$, $\eta_p^2 = 0.316$. Importantly, the interaction of Block type with Affect was significant, $F(1, 27) = 5.489$, $p = .027$, $\eta_p^2 = 0.169$. This interaction effect indicated lower

positive affect in punishment ($M = 23.51$, $SEM = 1.54$) compared to neutral blocks ($M = 24.77$, $SEM = 1.43$), $t(27) = -2.083$, $p = .047$, 95% CI [-2.570, -0.0193]; whereas there was no significant difference in negative affect between them, $t(27) = 1.584$, $p = .125$, 95% CI [-0.168, 1.308] (see Figure 2A). Planned paired t-tests showed that levels of positive affect during the practice session ($M = 31.68$, $SEM = 1.21$) were significantly higher than those of any other experimental condition ($M \leq 24.77$, $SEM \leq 1.54$), $ts \geq 7.213$, $ps \leq .001$; whereas levels of negative affect during the practice session were not significantly lower than any other experimental condition, $ts \leq 0.857$, $ps \geq .399$ (see Figure 2A).

3.1.2 Stimulus ratings

The ANOVA showed a significant main effect of Congruency, with higher dislike feelings for incongruent ($M = 60.48$, $SEM = 3.19$) than congruent trials ($M = 36.42$, $SEM = 3.31$), $F(1, 27) = 22.577$, $p < .001$, $\eta_p^2 = 0.455$. No other significant effects were found, $F_s \leq 1.367$, $ps \geq 0.253$, $\eta_p^2 \leq 0.048$. Planned paired t-tests showed that dislike feelings toward incongruent trials during the practice session ($M = 50$, $SEM = 3.01$) were significantly lower than any other experimental condition ($M \geq 60$, $SEM \leq 3.28$), $ts \geq 3.577$, $ps \leq .001$; whereas dislike levels toward congruent trials during the practice session were not significantly lower than any other experimental condition, $ts \leq 1.111$, $ps \geq .276$ (see Table 2).

Insert Table 2 about here

3.2 Behavioral results

3.2.1 RTs

The ANOVA showed that the main effect of Block type was significant, $F(1, 27) =$

13.447, $p = .001$, $\eta_p^2 = 0.332$, with faster RTs in punishment ($M = 445$ ms, $SEM = 8.32$) than neutral blocks ($M = 459$ ms, $SEM = 10.02$). The main effect of Current congruency was also significant, $F(1, 27) = 43.279$, $p < .001$, $\eta_p^2 = 0.616$, with faster RTs for congruent ($M = 444$ ms, $SEM = 8.50$) than incongruent trials ($M = 460$ ms, $SEM = 9.64$). Importantly, the three-way interaction between Block type, Previous congruency and Current congruency was significant, $F(1, 27) = 5.972$, $p = .021$, $\eta_p^2 = 0.181$. To explore the modulatory effect of defensive motivation on conflict adaptation, two ANOVAs were performed for Punishment and Neutral blocks separately. In punishment blocks (see Figure 2B), the Current congruency was significant, $F(1, 27) = 27.873$, $p < .001$, $\eta_p^2 = 0.508$, with faster RTs for congruent ($M = 437$ ms, $SEM = 7.94$) than incongruent trials ($M = 453$ ms, $SEM = 9.07$). The interaction of Previous congruency with Current congruency was also significant, $F(1, 27) = 6.224$, $p = .019$, $\eta_p^2 = 0.187$, indicating the presence of a reliable CAE in this condition. This interaction effect was explained by faster RTs in cC ($M = 433$ ms, $SEM = 8.38$) than in iC trials ($M = 441$ ms, $SEM = 7.83$), $t(27) = -2.055$, $p = .050$, 95% CI [-15.859, -0.113], whereas the difference between cI and iI trials did not reach significance, $t(27) = 1.564$, $p = .129$, 95% CI [-1.798, 13.334]. Although this analysis did not reveal a significant RT difference between cI and iI trials, the (iI-iC) ($M = 8.644$ ms, $SEM = 4.08$) RT difference was significantly smaller than the (cI-cC) ($M = 22.347$ ms, $SEM = 3.961$) RT however, $t(27) = -2.495$, $p = .019$, 95% CI [-24.972, -2.433], unambiguously translating a change (i.e. facilitation) in conflict processing at the current trial level depending on (in)congruency encountered during the previous trial. By comparison, in neutral blocks (see Figure 2C), the Current congruency was significant, $F(1, 27) = 22.836$, $p < .001$, $\eta_p^2 = 0.458$, with faster RTs for congruent ($M = 451$ ms, $SEM = 9.26$), than incongruent trials ($M = 467$ ms, $SEM = 10.96$), but the interaction between Previous congruency and Current congruency was not, $F(1, 27) = 0.115$, $p = 0.737$, $\eta_p^2 = 0.004$.

Insert Fig 2 about here

3.2.2 Error rates

The ANOVA revealed a significant main effect of Current congruency, $F(1, 27) = 4.279, p = .048, \eta_p^2 = 0.137$, with lower error rates for congruent ($M = 0.217, SEM = 0.01$) than incongruent trials ($M = 0.232, SEM = 0.01$). The interaction of Previous and Current congruency was significant, $F(1, 27) = 5.816, p = .023, \eta_p^2 = 0.177$. Error rates was similar for cC ($M = 0.211, SEM = 0.01$) and iC ($M = 0.223, SEM = 0.02$) trials, $t(27) = -1.406, p = .171$, 95% CI [-0.029, 0.005], while it was numerically lower, albeit not significantly for iI ($M = 0.224, SEM = 0.02$) relative to cI ($M = 0.240, SEM = 0.02$) trials, $t(27) = 1.822, p = .080$, 95% CI [-0.002, 0.338].

3.3 ERP results

3.3.1 Conflict adaptation: N2, N450

For the N2, the ANOVA revealed a significant main effect of Block type, $F(1, 27) = 9.691, p = .004, \eta_p^2 = 0.264$, with more negative amplitudes in neutral ($M = -0.27 \mu V, SEM = 0.85$) than punishment blocks ($M = 1.26 \mu V, SEM = 0.71$). This analysis also showed a marginally significant main effect of Previous congruency, $F(1, 27) = 3.804, p = .062, \eta_p^2 = 0.123$, with more negative amplitudes for incongruent ($M = 0.24 \mu V, SEM = 0.76$) than congruent trials ($M = 0.75 \mu V, SEM = 0.76$) at the previous trial level. Importantly, the three-way interaction between Block type, Previous congruency, and Current congruency was significant, $F(1, 27) = 6.508, p = .017, \eta_p^2 = 0.194$. No other significant effect was found, all $F_s \leq 0.834, p_s \geq .369, \eta_p^2_s \leq 0.030$. To explore the modulatory effect of defensive motivation on conflict adaptation at the N2 level, two separate ANOVAs were performed, for punishment

and neutral blocks separately. In punishment blocks (see Figures 3A & B), the Previous congruency was significant, $F(1, 27) = 4.426, p = .045, \eta_p^2 = 0.141$, with more negative amplitudes for incongruent ($M = 0.90 \mu\text{V}, SEM = 0.75$) than congruent ($M = 1.61 \mu\text{V}, SEM = 0.70$) trials at the previous trial level, as well as the interaction of Previous congruency with Current congruency, demonstrating that a reliable CAE was captured at the N2 level, $F(1, 27) = 5.334, p = .029, \eta_p^2 = 0.165$. This interaction effect was explained by a more negative N2 amplitude for iC ($M = 0.53 \mu\text{V}, SEM = 0.79$) than cC ($M = 1.89 \mu\text{V}, SEM = 0.72$) trials, $t(27) = 3.002, p = .006, 95\% \text{ CI } [0.431, 2.292]$, whereas the N2 amplitude was balanced between cI and iI trials, $t(27) = 0.163, p = .872, 95\% \text{ CI } [-0.805, 0.944]$, thus mimicking the result patterns found at the behavioral level (although we failed to find a significant correlation between the CAE at the behavioral level and the amplitude of the N2, $r = -0.192, p = .327$; see Figure 3B). In neutral blocks (see Figure 3D & E), no significant modulation of the N2 was observed, all $F_s \leq 0.765, p_s \geq .390, \eta_p^2_s \leq 0.028$.

For the N450 (see Figures 3A & D), the ANOVA revealed a significant main effect of Block type, $F(1, 27) = 9.091, p = .006, \eta_p^2 = 0.252$, with more negative amplitudes in neutral ($M = 2.98 \mu\text{V}, SEM = 1.51$) compared to punishment ($M = 5.58 \mu\text{V}, SEM = 1.26$) blocks (see Figures 3C & F). The effect of Current congruency was marginally significant, $F(1, 27) = 3.577, p = .069, \eta_p^2 = 0.117$, with more negative amplitudes for incongruent ($M = 3.95 \mu\text{V}, SEM = 1.41$) than congruent trials ($M = 4.61 \mu\text{V}, SEM = 1.26$) at the current trial level. The three-way interaction between Block type, Previous congruency, and Current congruency was not significant, $F(1, 27) = 3.098, p = .090, \eta_p^2 = 0.103$. No other significant effect was found, all $F_s \leq 2.460, p_s \geq .128, \eta_p^2_s \leq 0.083$.

Insert Fig 3 about here

3.3.2 Feedback processing: FRN, P3

For the FRN (see Figure 4A), the ANOVA revealed a significant main effect of Feedback valence $F(1, 27) = 9.833, p = .004, \eta_p^2 = 0.267$, with more negative amplitudes for negative ($M = -0.67 \mu\text{V}, SEM = 0.30$) compared to positive ($M = 0.28 \mu\text{V}, SEM = 0.24$) feedback. No other significant effect was found, all $F_s \leq 0.557, p_s \geq .462, \eta_p^2_s \leq 0.020$ (see Table 3). For the P3 (see Figure 4B & C), the ANOVA revealed a significant main effect of Block type, $F(1, 27) = 16.79, p < .001, \eta_p^2 = 0.384$, with more positive amplitudes for punishment ($M = 7.19 \mu\text{V}, SEM = 0.60$) compared to neutral ($M = 6.05 \mu\text{V}, SEM = 0.71$) blocks. The main effect of Feedback valence was marginally significant, $F(1, 27) = 3.854, p = .060, \eta_p^2 = 0.125$, with larger amplitudes for positive ($M = 7.04 \mu\text{V}, SEM = 0.61$) than negative ($M = 6.20 \mu\text{V}, SEM = 0.73$) feedbacks. Importantly, the interaction of Block type with Feedback valence was significant at the P3 level, $F(1, 27) = 5.541, p = .026, \eta_p^2 = 0.170$. This interaction effect was explained by larger P3 amplitude for positive ($M = 6.75 \mu\text{V}, SEM = 0.70$) than negative feedback ($M = 5.34 \mu\text{V}, SEM = 0.78$) in neutral blocks, $t(27) = 3.070, p = .005, 95\% \text{ CI } [0.468, -2.353]$, whereas in punishment blocks, the P3 had a similar amplitude for both, $t(27) = -0.537, p = .596, 95\% \text{ CI } [-1.359, 0.795]$. More specifically, in punishment blocks, the P3 amplitudes increased for negative feedback ($M = 7.05 \mu\text{V}, SEM = 0.74$), $t(27) = 4.141, p < .001, 95\% \text{ CI } [0.861, 2.551]$, selectively.

3.3.3 Relationship between P3 and CAE

Interestingly, in punishment blocks where a reliable CAE was found, its strength (as computed using a standard formula: $[cI - cC] - [iI - iC]$; see van Steenbergen et al., 2012) was positively correlated with the P3 amplitude following both negative ($r = 0.422, p = .025$) and positive feedback ($r = 0.418, p = .027$; see Figure 4D). No such significant relationship

was found for neutral blocks (following negative feedback: $r = 0.096$, $p = .628$; following positive feedback: $r = 0.214$, $p = .274$; see Figure 4E), although these two correlations were not significantly different from each other ($z = 0.93$, $p = .35$).

Insert Table 3 about here

Insert Fig 4 about here

3.4 MFT results

3.4.1 Conflict adaptation

This analysis showed a significant main effect of Block type, $F(1, 27) = 17.234$, $p < .001$, $\eta_p^2 = 0.390$, with larger MFT power for punishment ($M = 1.74$ dB, $SEM = 0.18$) relative to neutral ($M = 1.37$ dB, $SEM = 0.19$) blocks. Further, a marginally significant effect of Current congruency was found, $F(1, 27) = 4.007$, $p = .055$, $\eta_p^2 = 0.129$, with larger MFT power for incongruent ($M = 1.65$ dB, $SEM = 0.19$) relative to congruent ($M = 1.45$ dB, $SEM = 0.18$) trials. No other effect reached significance however, all $F_s \leq 1.148$, $p_s \geq .294$, $\eta_p^2_s \leq 0.041$ (see Table 4)³.

3.4.2 Influence of feedback on conflict processing

³ We also analyzed conflict adaptation using the total MFT power (i.e. without subtraction of evoked MFT power), instead of the induced component. This analysis showed a significant main effect of Block type, $F(1,27) = 8.504$, $p = 0.007$, $\eta_p^2 = 0.240$. The Current congruency was not significant, $F(1,27) = 0.177$, $p = 0.677$, $\eta_p^2 = 0.007$.

This analysis showed a significant main effect of Block type, $F(1, 27) = 14.067, p = .001, \eta_p^2 = 0.343$, with larger MFT power in punishment ($M = 1.80$ dB, $SEM = 0.17$) relative to neutral ($M = 1.49$ dB, $SEM = 0.18$) blocks. The main effect of Congruency was also significant, $F(1, 27) = 4.297, p = .048, \eta_p^2 = 0.137$, with larger MFT power for incongruent ($M = 1.74$ dB, $SEM = 0.18$) relative to congruent ($M = 1.54$ dB, $SEM = 0.17$) trials. The interaction effect of Block type with Congruency was marginally significant, $F(1, 27) = 3.558, p = .070, \eta_p^2 = 0.116$, while the three-way interaction was not significant, $F(1, 27) = 0.512, p = .481, \eta_p^2 = 0.019$. In order to test our a priori hypothesis (see Introduction) and because the two-way interaction between Block type and Congruency was trend significant, we performed two planned t-test comparisons to assess whether negative feedback (especially in punishment blocks) influenced incongruent trials selectively, or instead, boosted stimulus processing globally (i.e. irrespective of congruency) (see Figures 5A, B). The results showed that MFT power was significantly larger in punishment ($M = 1.74$ dB, $SEM = 0.21$) relative to neutral ($M = 1.36$ dB, $SEM = 0.19$) blocks, for incongruent trials only, $t(27) = 2.322, p = .028$, 95% CI [0.045, 0.727]. This difference was not significant for congruent trials, $t(27) = 0.887, p = .383$, 95% CI [-0.206, 0.520] (see Figures 5C).

Insert Table 4 about here

Insert Fig 5 about here

3.4.3 Feedback processing

This analysis showed a significant main effect of Feedback valence, $F(1, 27) = 21.856$, $p < .001$, $\eta_p^2 = 0.447$, with a larger MFT power for negative ($M = 2.04$ dB, $SEM = 0.31$) relative to positive ($M = 1.18$ dB, $SEM = 0.21$) feedback (see Figures 6A & B). Importantly, the interaction of Block type with Feedback valence was significant as well, $F(1, 27) = 5.096$, $p = .032$, $\eta_p^2 = 0.159$ (see Figures 6C). This interaction was explained by a larger MFT power for negative feedback in punishment ($M = 2.23$ dB, $SEM = 0.35$) compared to neutral ($M = 1.86$ dB, $SEM = 0.29$) blocks, $t(27) = 2.348$, $p = .026$, 95% CI [0.047, 0.695]; whereas MFT power extracted for positive feedback did not differ between the two block types, $t(27) = -0.194$, $p = .848$, 95% CI [-0.289, 0.239].

Insert Fig 6 about here

4. Discussion

The current study sought to explore the underlying mechanism responsible for the increase in cognitive control seen at the CAE level when defensive motivation is at play (Yang & Pourtois, 2018). Defensive motivation was elicited in specific blocks, where evaluative negative feedback contingent on task performance entailed monetary loss for the participants. In different blocks, no such consequence was coupled with negative feedback. Hence, defensive motivation differed between them, as confirmed by subjective ratings (PANAS) administered after each block. This effect was mostly expressed by a selective decrease of positive affect in punishment compared to neutral blocks, as opposed to an increase in negative affect in our previous study (Yang & Pourtois, 2018)⁴. Participants

⁴ A possible reason accounting for this discrepancy is that subjective feelings of participants were asked more frequently (11 times) in the current study than the previous one (i.e. 6 times for Experiments 1A and 1B, and 5 times for Experiment 2) (see Yang & Pourtois, 2018). Moreover, uncontrolled individual differences between these two studies might also explain this discrepancy. In order to test this prediction, we compared the PANAS scores during the baseline period (extracted from the practice session) between the current experiment and

performed a 4 AFC Stroop task used to explore conflict adaptation at the behavioral and EEG levels, when controlling for feature repetition and contingency learning (Weissman et al., 2014). At the behavioral level, we replicated our previous findings (Yang & Pourtois, 2018) and found that the CAE was larger when defensive motivation was high compared to low, suggesting that cognitive control was enhanced by this specific manipulation. Moreover, additional subjective ratings clearly showed that incongruent (Stroop) stimuli were eventually more disliked than congruent ones after the encounter of the task and equally strongly for the two blocks types, thereby confirming indirectly that conflict had a negative or aversive meaning for these participants (Dreisbach & Fischer, 2012a). Importantly, at the EEG level, we found that this trial-by-trial increase in cognitive control driven by defensive motivation was mostly traced at the N2 level (whose amplitude showed a significant interaction between Previous and Current congruency), hence at an early stage following conflict processing. Additional analyses performed at the feedback level confirmed that negative feedback was differentially processed at the P3 level in punishment compared to neutral blocks, suggesting that it likely acquired a larger motivational significance in this former compared to later context. Last, complementing time-frequency analyses focusing on changes in induced MFT power showed that defensive motivation led to a general gain in cognitive control visible across all stimuli conditions, thereby being dissociable from the local and trial-by-trial adjustment seen at the N2 level.

Behaviorally, conflict adaptation was present in punishment blocks, but absent in neutral ones, as we hypothesized. This pattern closely replicates our previous findings (Yang & Pourtois, 2018), confirming that defensive motivation fuels conflict-driven adaptive control processes (Inzlicht et al, 2015). At the ERP level, a smaller (i.e., less negative) N2 and N450

Experiment 2 in the previous study. Results showed that participants in the current experiment had higher scores for NA (without differences in PA) during the baseline period compared to the participants in our previous experiment.

amplitudes were found in punishment compared to neutral blocks. The early gain in cognitive control captured at the N2 level by defensive motivation accords well with previous ERP studies showing that this early conflict-related component is a reliable index of conflict processing as well as conflict-driven adaptive control processes, and hence the CAE (Forster et al., 2011; Larson et al., 2013, 2014, 2015; Clayson & Larson, 2011b). The amplitude of the N2 is sensitive to variations in cognitive control depending on the previous trial history and more specifically, the encounter of conflict during the immediately preceding trial. However, this gain was confined to the N2 component in our study, and did not encompass the subsequent N450 component, for which the amplitude was larger for incongruent than congruent stimuli (at the current trial level only). This dissociation lends support to previous ERP studies suggesting that the N2 and N450 likely reflect different stages of information processing during conflict monitoring (Larson et al., 2016; McKay et al., 2017). More specifically, whereas the N450 is mostly influenced by conflict processing occurring at the current trial level, the N2 is more sensitive to subtle changes in levels of cognitive control occurring across successive trials, and therefore reflected by conflict adaptation (Larson et al., 2009; Larson et al., 2012b).

Although our ERP results are broadly consistent with this framework, they also extend this literature by showing that the N2 ERP component and MFT power changes (induced activity) can capture different modulations of defensive motivation on conflict monitoring, and more generally cognitive control. Unlike the N2, MFT power did not vary with conflict adaptation, although it was larger for incongruent compared to congruent trials. Moreover, MFT power was larger in punishment than neutral blocks and for negative than positive feedback, thereby confirming that the power in this low frequency band did change with defensive motivation, and hence when levels of cognitive control had to be adjusted. MFT, given its peculiar function and electrophysiological properties (Cavanagh & Frank, 2014), is

likely involved in another, probably more complex control process than the phasic N2 ERP component, eventually making it a less suitable electrophysiological marker to capture conflict adaptation (see also van Steenbergen et al., 2012). However, this conclusion does not undermine the possibility to find striking similarities between the N2 and MFT in some cases, in particular when conflict adaptation is considered (Pastötter, Dreisbach, & Bäuml, 2013; Cohen & Cavanagh, 2011). A main difference between the current and these previous studies is that they typically used conflict-related paradigms without the use of (negative) feedback at the single trial level, which can interfere with dynamic changes in MFT power as a function of conflict adaptation. Hence, purer MFT power modulations depending on the recent trial history were likely obtained in those previous studies (Pastötter et al., 2013; Cohen & Cavanagh, 2011). However, it should be emphasized that the use of negative feedback contingent on task performance and shown on a trial-by-trial basis was needed here to create a specific affective state compatible with enhanced defensive motivation, and being susceptible to ease conflict-driven adaptive control processes.

As our new results suggest, changes in cognitive control occurring across trials probably involve rapid and transient processes following stimulus onset. These are therefore best captured by the N2 given its specific electrophysiological properties as well as temporal proximity with the previous trial, and the lingering processing of conflict resulting from it. In comparison, MFT is probably less phasic than the N2, and captures more complex and global changes in cognitive control or even motivation (such as the need for enhanced control or effort; see Cavanagh & Frank, 2014). In agreement with this view, MFT power did vary clearly with the main experimental condition manipulated in the current study (i.e. the use of loss-related evaluative feedback in some blocks): it was larger in punishment than neutral blocks, yet without precisely following more subtle changes in cognitive control during stimulus processing, as visible at the CAE level. Hence, in light of this dissociation, it appears

parsimonious to conclude that MFT power probably reflects enhanced coordination in a distributed network during defensive motivation, interference or uncertainty (Shackman et al., 2011; Mansouri, Tanaka, & Buckley, 2009; Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012; Hobson et al., 2014). However, this selective change does not necessarily influence conflict-driven adaptive control processes operating at a faster or different scale however.

According to the conflict-induced control model (Mansouri et al., 2009), conflicting information experienced during the previous trial may be held during the inter-trial interval, which is then used to adjust the control level needed at the current trial. Conflict processing is thus improved at the behavioral level if a similar conflicting situation is encountered twice in a row (e.g. in II trials; Mansouri et al., 2009; Jiang, Zhang, & van Gaal, 2015). Our new results suggest that defensive motivation could facilitate this specific carry-over effect, probably because conflict information and negative affect resulting from defensive motivation can easily be integrated with each other (Kanske & Kotz, 2010). Presumably, the inclusion of punishment-related feedback in between consecutive trials enabled to actively hold and timely reinforce such control signal extracted from the preceding trial to influence the subsequent one (Lindström et al., 2013). Alternatively, this event could serve to reinforce the binding between previous conflict processing and the currently active (conflict-related) representation (Dreisbach & Fischer, 2012b; Abrahamse et al., 2016).

By comparison, in neutral blocks, the intervening evaluative feedback probably blocked, rather than facilitated, trial-by-trial adjustments in cognitive control, thereby eliminating the CAE at the behavioral level (as well as N2 level). However, the conclusion that the absence of CAE in neutral blocks is caused by the intervening feedback awaits confirmation at the empirical level. Additional studies comparing the CAE in neutral blocks with and without evaluative feedback are needed to ascertain that reactive control is attenuated by this

intervening feedback, as we surmise here. Notwithstanding this limitation, it appears parsimonious to conclude that the intervening evaluative feedback was probably instrumental in facilitating conflict-driven adaptive control processes, if and only if it had a distinctive motivational significance, as achieved here using loss-related feedback contingent on task performance (see also Yang & Pourtois, 2018).

Interestingly, we also gained additional insight into the dynamic modulation of cognitive control by defensive motivation when analyzing feedback-locked ERP effects. Because the P3, but not preceding FRN for negative feedback was influenced by it, we have good reasons to believe that this manipulation actually altered the motivational significance of the feedback (P3 effect) per se, but not merely its value or expectation, for which a systematic modulation of the FRN would have been found (Ullsperger, Danielmeier, & Jocham, 2014; Koban & Pourtois, 2014). More specifically, our results showed that the P3 was larger for positive than negative feedback in neutral blocks, but this valence effect was erased in punishment blocks. Because the P3 has been repeatedly reported to be larger for positive compared to negative feedback (see van den Berg et al., 2012; Severo et al., 2017, 2018), the lack of valence effect on the P3 in punishment blocks deviates from these previous findings, whereas its present in neutral blocks appears compatible with them. Accordingly, we can assume a selective increase of feedback processing at the P3 level for negative feedback with defensive motivation. Moreover, as the feedback-locked P3 is thought to reflect the motivational significance or impact assigned to this event (Steinhauser & Yeung, 2010; n  nez Castellar et al., 2010; Severo et al., 2017, 2018), we can conclude that negative feedback selectively acquired a larger motivational significance or impact in punishment compared to neutral blocks. Moreover, the observation of a larger P3 for positive than negative feedback in neutral blocks is not odd (see van den Berg et al., 2012; Severo et al., 2017, 2018), but probably reflects the differential motivational value assigned by the participants to this goal conducive

event in this safe situation. Hence, we suggest that negative feedback was rapidly processed as a negative event in both conditions equally strongly at the FRN level, but later on, at the P3 level, selectively evaluated as distinctive or important if encountered in blocks where monetary loss resulted from its encounter.

For punishment blocks only, we also found that the amplitude of the P3 component (irrespective of the valence of the feedback though) was positively related with the CAE at the behavioral level. This correlation confirmed indirectly that the CAE was therefore influenced by specific motivational processes taking place at the feedback level. By comparison, we failed to observe a significant relation between the amplitude of N2 and the CAE at the behavioral level, although these two variables behaved in a strikingly similar way. Tentatively, we suggest that important motivational processes active at the feedback level during punishment blocks (e.g. P3 component) may have potentially obscured the observation of a direct link between this early ERP component, and conflict-driven adaptive control processes at the behavioral level, as seen using the CAE (Clayson & Larson, 2011a).

Further, at the feedback level, negative feedback was associated with a larger MFT power when comparing punishment to neutral blocks, without any similar modulation seen for positive feedback however. Presumably, negative feedback was selectively associated with the need for enhanced control when encountered in these specific blocks. It seems plausible to assume enhanced control reflected by MFT power in this situation since the negative evaluative feedback signaled an undesirable outcome and acted therefore as a potent cue triggering the need for behavioral adjustment, usually achieved through an active upregulation of behavior and/or optimization of task performance (Luft, 2014; Hayden & Platt, 2010). In our study, this MFT power change was mostly visible for negative feedback delivered in punishment blocks, as the feedback in this condition did not only inform about the task performance, but was also coupled with a specific monetary loss, and therefore, it likely

unlocked an transient increase in current levels of control to deal with this challenge (Marco-Pallarés et al., 2008). However, because of the use of a subject and trial specific calibration for the response deadline, this transient increase in cognitive control did not result in a specific improvement at the behavioral level visible for the trial immediately following negative feedback in punishment blocks. Nonetheless, this adaptive and strict response deadline ensured that the proportion of negative and positive feedback was balanced, and this symmetry was similar for the two main block types used in our study (see Table 1). This was an important pre-requisite at the methodological level to perform subsequently adequate statistical comparisons between the main experimental conditions.

All in all, our new EEG results add to the existing literature on defensive motivation and cognitive control by suggesting strong ties between them (Inzlicht et al., 2015; Dreisbach & Fischer, 2012b), especially when the N2 and MFT are considered (Cavanagh & Frank, 2014). In this framework, negative affect resulting from defensive motivation is not epiphenomenal, but instead, instrumental and maybe even causal for adjusting levels of cognitive control, on a trial-by-trial basis. These dynamic changes likely arise because defensive motivation usually corresponds to a potent, internal and meaningful affective state that informs individuals about a nearing threat or challenge, and hence, the need to take remedial actions to overcome it (Shackman et al., 2016). Therefore, defensive motivation, when controllable and of mild intensity, usually fuels enhanced cognitive control (Inzlicht et al., 2015). Our new findings lend support to this general theoretical framework, but also help to pinpoint the actual stage of information processing following stimulus onset where defensive motivation dynamically enhances cognitive control, namely the N2 component. Interestingly, this boost in cognitive control by defensive motivation seen at the N2 level could be dissociated from a more global change visible across all conditions, as shown by the complementing MFT results.

To conclude, the current results show that defensive motivation can increase cognitive control, and more specifically conflict adaptation. This gain is mostly reflected by an early change following stimulus onset at the N2 level, selectively, whose amplitude varied systematically with conflict adaptation found at the behavioral level. This neurophysiological effect was dissociable from other, likely more general cognitive and/or affective control effects visible at the MFT level, whose power augmented for all stimuli when embedded in a context where defensive motivation was elicited. Accordingly, we suggest that the influence of defensive motivation on cognitive control can be traced at different levels during conflict processing, translating either specific or general modulatory effects. Locally, defensive motivation facilitates conflict-driven adaptive control processes with effects visible at the CAE and the N2 levels. Globally, it increases the need for enhanced cognitive control by augmenting stimulus-locked MFT. Additional EEG studies are needed to assess whether these two effects necessarily co-occur during conflict processing under defensive motivation, or one might emerge in the absence of the other one.

Declarations of interest

None

Acknowledgements

Qian Yang is supported by a grant (201606990022) from China Scholarship Council (CSC) and co-funding (BOF) grant (BOFCHN2017000101) from Ghent University. Katharina Paul is supported by a PhD mandate (11U9216N) granted from the Research Foundation Flanders (FWO).

References

- Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 479–492.
- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin, 142*, 693.
- Alpay, G., Goerke, M., & Stürmer, B. (2009). Precueing imminent conflict does not override sequence-dependent interference adaptation. *Psychological Research PRPF, 73*, 803.
- Beste, C., Ness, V., Lukas, C., Hoffmann, R., Stüwe, S., Falkenstein, M., & Saft, C. (2012). Mechanisms mediating parallel action monitoring in fronto-striatal circuits. *Neuroimage, 62*, 137–146.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624.
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences, 18*, 414–421.
- Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience, 29*, 98–105.
- Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49*, 220–238.
- Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology-Paris, 109*, 3–15.
- Clayson, P. E., & Larson, M. J. (2011a). Conflict adaptation and sequential trial effects: Support for the conflict monitoring theory. *Neuropsychologia, 49*, 1953–1961.
- Clayson, P. E., & Larson, M. J. (2011b). Effects of repetition priming on electrophysiological and behavioral indices of conflict adaptation and cognitive control. *Psychophysiology, 48*, 1621–1630.
- Cohen, M. X. (2011). Hippocampal-prefrontal connectivity predicts midfrontal oscillations and long-term memory performance. *Current Biology, 21*, 1900–1905.

- Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in Psychology*, 2, 30.
- Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology*, 110, 2752–2763.
- Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences*, 37, 480–490.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21.
- Desender, K., van Opstal, F., & van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science*, 25, 675–683.
- Dreisbach, G., & Fischer, R. (2012a). Conflicts as aversive signals. *Brain and Cognition*, 78, 94–98.
- Dreisbach, G., & Fischer, R. (2012b). The role of affect and reward in the conflict-triggered adjustment of cognitive control. *Frontiers in Human Neuroscience*, 6, 342.
- Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, 24, 255–260.
- Dreisbach, G., & Fischer, R. (2016). Conflicts as aversive signals: Motivation for control adaptation in the service of affect regulation.
- Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 7, 380–390.
- Egner, T., Ely, S., & Grinband, J. (2010). Going, going, gone: characterizing the time-course of congruency sequence effects. *Frontiers in Psychology*, 1, 154.
- Engelen, U., De Peuter, S., Victoir, A., van Diest, I., & van den Bergh, O. (2006). Verdere validering van de Positive and Negative Affect Schedule (PANAS) en vergelijking van twee Nederlandstalige versies. [Further validation of the Positive and Negative Affect Schedule (PANAS) and comparison of two Dutch versions]. *Gedrag en Gezondheid*, 34, 61–70.

- Forster, S. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2011). Parametric manipulation of the conflict signal and control-state adaptation. *Journal of Cognitive Neuroscience*, *23*, 923–935.
- Folstein, J. R., & van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*, 152–170.
- Freitas, A. L., Banai, R., & Clark, S. L. (2009). When cognitive control is calibrated: Event-related potential correlates of adapting to information-processing conflict despite erroneous response preparation. *Psychophysiology*, *46*, 1226–1233.
- Fröber, K., Stürmer, B., Frömer, R., & Dreisbach, G. (2017). The role of affective evaluation in conflict adaptation: an LRP study. *Brain and Cognition*, *116*, 9–16.
- Gheza, D., De Raedt, R., Baeken, C., & Pourtois, G. (2018). Integration of reward with cost anticipation during performance monitoring revealed by ERPs and EEG spectral perturbations. *NeuroImage*, *173*, 153–164.
- Gray, J. R. (2004). Integration of emotion and cognitive control. *Current directions in Psychological Science*, *13*, 46–48.
- Hayden, B. Y., & Platt, M. L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. *Journal of Neuroscience*, *30*, 3339–3346.
- Hobson, N. M., Saunders, B., Al-Khindi, T., & Inzlicht, M. (2014). Emotion down-regulation diminishes cognitive control: A neurophysiological investigation. *Emotion*, *14*, 1014.
- Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*, 126–132.
- Jiang, J., Zhang, Q., & van Gaal, S. (2015). Conflict awareness dissociates theta-band neural dynamics of the medial frontal and lateral frontal cortex during trial-by-trial cognitive control. *Neuroimage*, *116*, 102–111.
- Kanske, P., & Kotz, S. A. (2010). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, *48*, 3661–3664.
- Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neuroscience & Biobehavioral Reviews*, *46*, 71–84.

- Kuhbandner, C., & Zehetleitner, M. (2011). Dissociable effects of valence and arousal in adaptive executive control. *PloS One*, *6*, e29287.
- Lang, P. J., & Bradley, M. M. (2013). Appetitive and defensive motivation: Goal-directed or goal-determined? *Emotion Review*, *5*, 230–234.
- Larson, M. J., Kaufman, D. A., & Perlstein, W. M. (2009). Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia*, *47*, 663–670.
- Larson, M. J., Clawson, A., Clayson, P. E., & South, M. (2012a). Cognitive control and conflict adaptation similarities in children and adults. *Developmental Neuropsychology*, *37*, 343–357.
- Larson, M. J., Clayson, P. E., & Baldwin, S. A. (2012b). Performance monitoring following conflict: internal adjustments in cognitive control? *Neuropsychologia*, *50*, 426–433.
- Larson, M. J., Clawson, A., Clayson, P. E., & Baldwin, S. A. (2013). Cognitive conflict adaptation in generalized anxiety disorder. *Biological Psychology*, *94*, 408–418.
- Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: a theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology*, *93*(3), 283–297.
- Larson, M. J., Clayson, P. E., Primosch, M., Leyton, M., & Steffensen, S. C. (2015). The effects of acute dopamine precursor depletion on the cognitive control functions of performance monitoring and conflict processing: an event-related potential (ERP) study. *PLoS One*, *10*, e0140770.
- Larson, M. J., Clayson, P. E., Kirwan, C. B., & Weissman, D. H. (2016). Event-related potential indices of congruency sequence effects without feature integration or contingency learning confounds. *Psychophysiology*, *53*, 814–822.
- Lindström, B. R., Mattsson-Mårn, I. B., Golkar, A., & Olsson, A. (2013). In your face: risk of punishment enhances cognitive control and error-related activity in the corrugator supercilii muscle. *PLoS One*, *8*, e65692.
- Luft, C. D. B. (2014). Learning from feedback: the neural mechanisms of feedback processing facilitating better performance. *Behavioural Brain Research*, *261*, 356–368.

- Marco-Pallarés, J., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *Journal of Cognitive Neuroscience*, *20*, 1595–1610.
- Mansouri, F. A., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience*, *10*, 141.
- McKay, C. C., van den Berg, B., & Woldorff, M. G. (2017). Neural cascade of conflict processing: Not just time-on-task. *Neuropsychologia*, *96*, 184–191.
- Mordkoff, J. T. (2012). Observation: Three reasons to avoid having half of the trials be congruent in a four-alternative forced-choice experiment on sequential modulation. *Psychonomic Bulletin & Review*, *19*, 750–757.
- núñez Castellar, E., Kühn, S., Fias, W., & Notebaert, W. (2010). Outcome expectancy and not accuracy determines posterror slowing: ERP support. *Cognitive, Affective, & Behavioral Neuroscience*, *10*, 270–278.
- Nigbur, R., Ivanova, G., & Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, *122*, 2185–2194.
- Pastötter, B., Dreisbach, G., & Bäuml, K. H. T. (2013). Dynamic adjustments of cognitive control: oscillatory correlates of the conflict adaptation effect. *Journal of Cognitive Neuroscience*, *25*, 2167–2178.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*, 148.
- Questienne, L., van Opstal, F., van Dijck, J. P., & Gevers, W. (2016). Metacognition and cognitive control: behavioural adaptation requires conflict experience. *The Quarterly Journal of Experimental Psychology*, 1–15.
- Sallet, J., Camille, N., & Procyk, E. (2013). Modulation of feedback-related negativity during trial-and-error exploration and encoding of behavioral shifts. *Frontiers in Neuroscience*, *7*, 209.
- Santesso, D. L., Bogdan, R., Birk, J. L., Goetz, E. L., Holmes, A. J., & Pizzagalli, D. A. (2011). Neural responses to negative feedback are related to negative emotionality in healthy adults. *Social Cognitive and Affective Neuroscience*, *7*, 794–803.

- Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews*, *34*, 1015–1022.
- Steinhauser, M., & Yeung, N. (2010). Decision processes in human performance monitoring. *Journal of Neuroscience*, *30*, 15643–15653.
- Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS One*, *9*, e102337.
- Schuch, S., & Koch, I. (2015). Mood states influence cognitive control: The case of conflict adaptation. *Psychological Research*, *79*, 759–772.
- Schuch, S., Zweerings, J., Hirsch, P., & Koch, I. (2017). Conflict adaptation in positive and negative mood: Applying a success-failure manipulation. *Acta Psychologica*, *176*, 11–22.
- Schouppe, N., Braem, S., De Houwer, J., Silvetti, M., Verguts, T., Ridderinkhof, K. R., & Notebaert, W. (2015). No pain, no gain: the affective valence of congruency conditions changes following a successful response. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 251–261.
- Severo, M. C., Walentowska, W., Moors, A., & Pourtois, G. (2017). Goal impact influences the evaluative component of performance monitoring: Evidence from ERPs. *Biological Psychology*, *129*, 90–102.
- Severo, M. C., Walentowska, W., Moors, A., & Pourtois, G. (2018). Goals matter: Amplification of the motivational significance of the feedback when goal impact is increased. *Brain and Cognition*, *128*, 56–72.
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*, 154.
- Shackman, A. J., Stockbridge, M. D., Tillman, R. M., Kaplan, C. M., Tromp, D. P., Fox, A. S., & Gamer, M. (2016). The neurobiology of dispositional negativity and attentional biases to threat: implications for understanding anxiety disorders in adults and youth. *Journal of Experimental Psychopathology*, *7*, 311–342.

- Tillman, C. M., & Wiens, S. (2011). Behavioral and ERP indices of response conflict in Stroop and flanker tasks. *Psychophysiology*, *48*, 1405–1411.
- Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*, 35–79.
- van den Berg, I., Shaul, L., van der Veen, F. M., & Franken, I. H. (2012). The role of monetary incentives in feedback processing: why we should pay our participants. *Neuroreport*, *23*, 347–353.
- van Steenbergen, H., Band, G. P., & Hommel, B. (2009). Reward counteracts conflict adaptation: Evidence for a role of affect in executive control. *Psychological Science*, *20*, 1473–1477.
- van Steenbergen, H., Band, G. P., & Hommel, B. (2010). In the mood for adaptation: How affect regulates conflict-driven control. *Psychological Science*, *21*, 1629–1634.
- van Steenbergen, H., Band, G. P., & Hommel, B. (2012). Reward valence modulates conflict-driven attentional adaptation: Electrophysiological evidence. *Biological Psychology*, *90*, 234–241.
- Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Sciences*, *13*, 252–257.
- Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, *46*, 2545–2555.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063.
- Weissman, D. H., Jiang, J., & Egnér, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 2022.
- Whitehead, P. S., Brewer, G. A., & Blais, C. (2017). ERP evidence for conflict in contingency learning. *Psychophysiology*, *54*, 1031–1039.
- Yang, Q., & Pourtois, G. (2018). Conflict-driven adaptive control is enhanced by integral negative emotion on a short time scale. *Cognition and Emotion*, *32*, 1–17.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931.

Yeung, N., & Cohen, J. D. (2006). The impact of cognitive deficits on conflict monitoring: Predictable dissociations between the error-related negativity and N2. *Psychological Science*, *17*, 164–171.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*, 6258–6264.

Figures Legends

Fig. 1 Experimental procedure. Each trial started with a fixation cross, followed by the Stroop stimulus. A blank screen ensued, before the performance-contingent feedback was presented (being either negative or positive). In punishment blocks, negative feedback led to monetary loss whereas in neutral blocks, no loss followed it. The figure shows an example of a cI trial.

Fig. 2 Subjective ratings and behavioral results. **A.** Positive affect was lower in punishment than neutral blocks, while negative affect did not differ between these two conditions. **B.** In punishment blocks, RT speed (for correct response) was influenced significantly by the Previous and Current congruency concurrently, in agreement with the CAE. **C.** In comparison, in neutral blocks only the effect of Current congruency was significant. The error bar represents the standard error of the mean (SE). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig. 3 Stimulus-locked ERP results. **A.** Grand average ERP waveforms from Fz and FCz (pooled together) for cC, cI, iC, and iI trials in punishment blocks. **B.** Mean amplitudes of N2 (extracted in the 280-320 ms interval following stimulus onset) for cC, cI, iC, and iI trials in punishment blocks. For each trial type, the corresponding horizontal topographical map is shown. **C.** Mean amplitudes of N450 (extracted in the 400-500 ms interval following stimulus onset) for cC, cI, iC, and iI trials in punishment blocks, with the corresponding horizontal topographical map each time. **D-F.** Same as **A-C** for neutral blocks. The error bar represents the standard error of the mean (SE). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fig. 4 Feedback-locked ERP results. **A.** Grand average ERP waveforms from FCz for Punishment block-Negative feedback (Pun-NegFB), Punishment block-Positive feedback (Pun-PosFB), Neutral block-Negative feedback (Neu-NegFB), and Neutral block-Positive feedback (Neu-PosFB). The FRN was extracted as the difference between the most negative peak (in the 250-350 ms interval) and the average voltage of the preceding (P2) and the following positive peaks (P3) following feedback onset. For each condition separately, the corresponding topographical map is shown. **B.** Grand average ERP waveforms from Pz for the same four conditions separately. The P3 amplitude was extracted in the 300-400 ms interval following feedback onset. **C.** Mean amplitudes of the P3 for the four main conditions and showing that it increased selectively for negative feedback in punishment blocks. **D.** In punishment blocks, a significant positive correlation was found between P3 amplitudes (averaged for negative and following positive feedback together) and the CAE at the behavioral level. **E.** In neutral blocks, no significant correlation was found between P3 amplitudes and the CAE at the

behavioral level. The error bar represents the standard error of the mean (SE). $***p < 0.001$, $**p < 0.01$, $*p < 0.05$.

Fig. 5 A. MFT (4-8Hz) power changes (FCz) at the stimulus level following negative feedback, for the four main conditions separately: Punishment-Congruent (negFB-punC, i.e. Congruent stimulus following negative feedback during punishment blocks), Punishment-Incongruent (negFB-punI, i.e. Incongruent stimulus following negative feedback during punishment blocks), Neutral-Congruent (negFB-neuC, i.e. Congruent stimulus following negative feedback during neutral blocks), and Neutral-Incongruent (negFB-neuI, i.e. Incongruent stimulus following negative feedback during neutral blocks). **B.** The corresponding horizontal topographical maps (extracted during the 300-600 ms interval following stimulus onset) are shown. **C.** Mean MFT power for the four main conditions separately, showing that it increased selectively for incongruent trials in punishment compared with neutral blocks. The error bar represents the standard error of the mean (SE). $***p < 0.001$, $**p < 0.01$, $*p < 0.05$.

Fig. 6 A. MFT (4-8Hz) power changes (FCz) at the feedback level for the four main conditions (see Figure 4 for descriptions and labels). **B.** The corresponding topographical maps (extracted in the 200-400 ms interval following feedback onset) are shown. **C.** Mean MFT power for the four main conditions, showing that MFT increased selectively for negative feedback in punishment blocks. The error bar represents the standard error of the mean (SE). $***p < 0.001$, $**p < 0.01$, $*p < 0.05$.

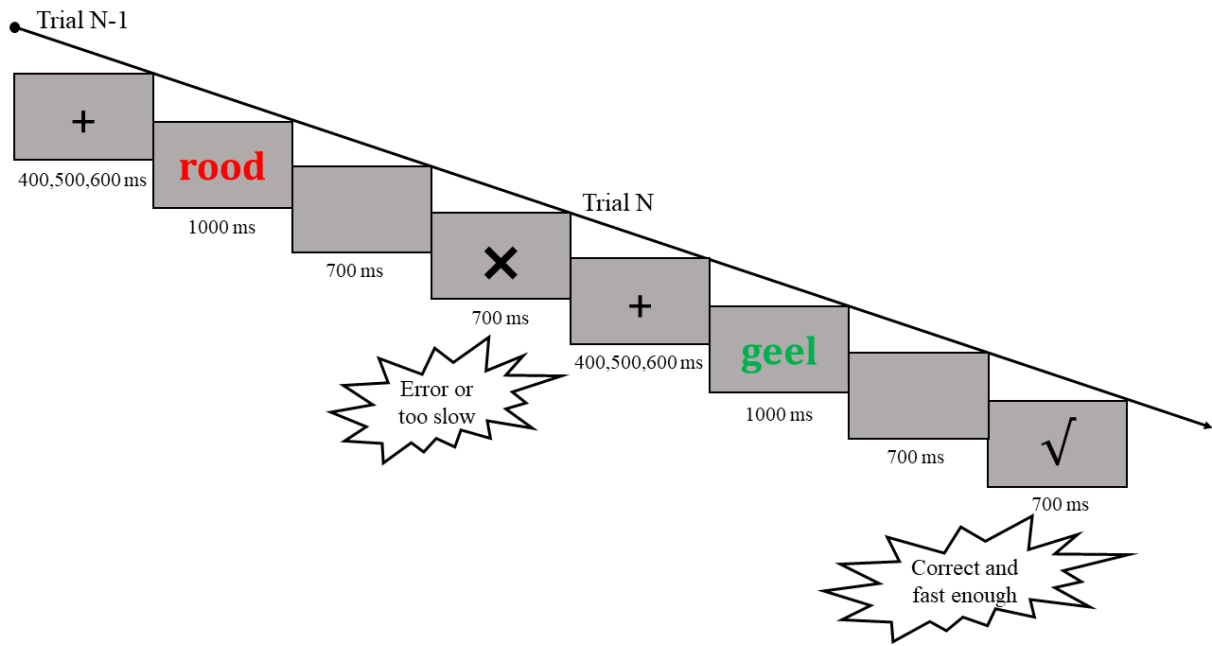


Figure 1

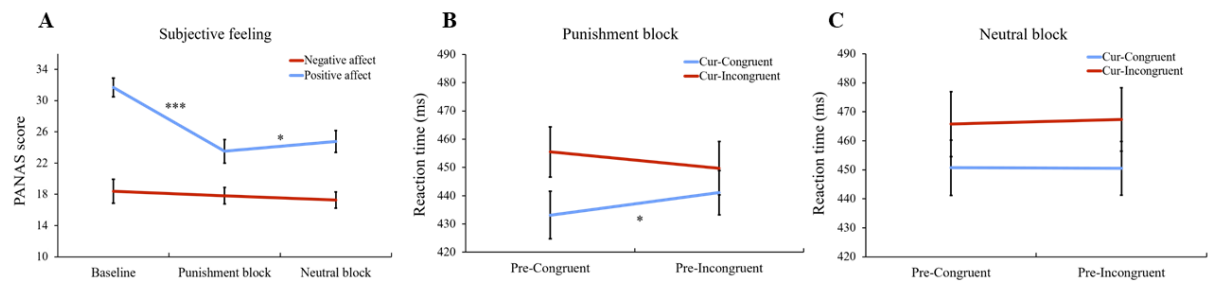
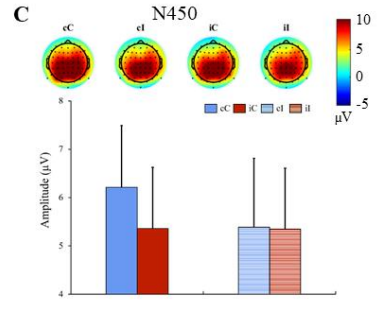
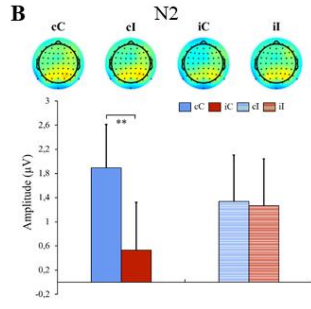
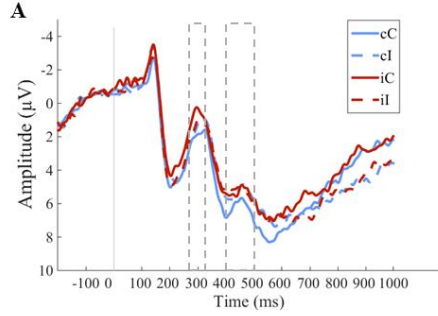


Figure 2

Punishment block



Neutral block

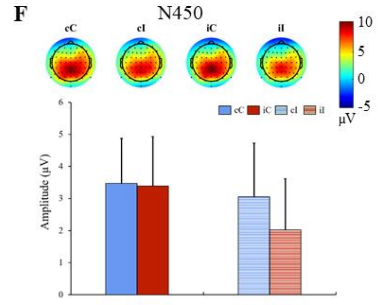
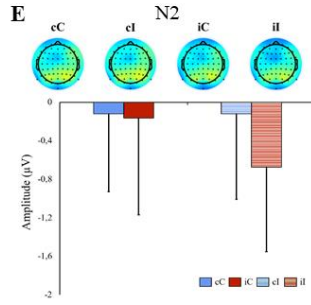
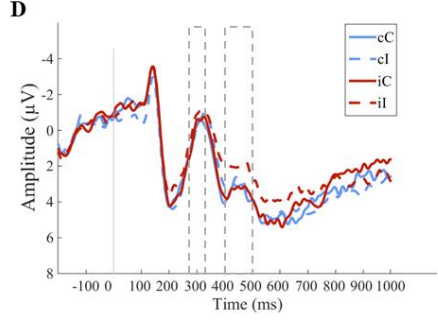


Figure 3

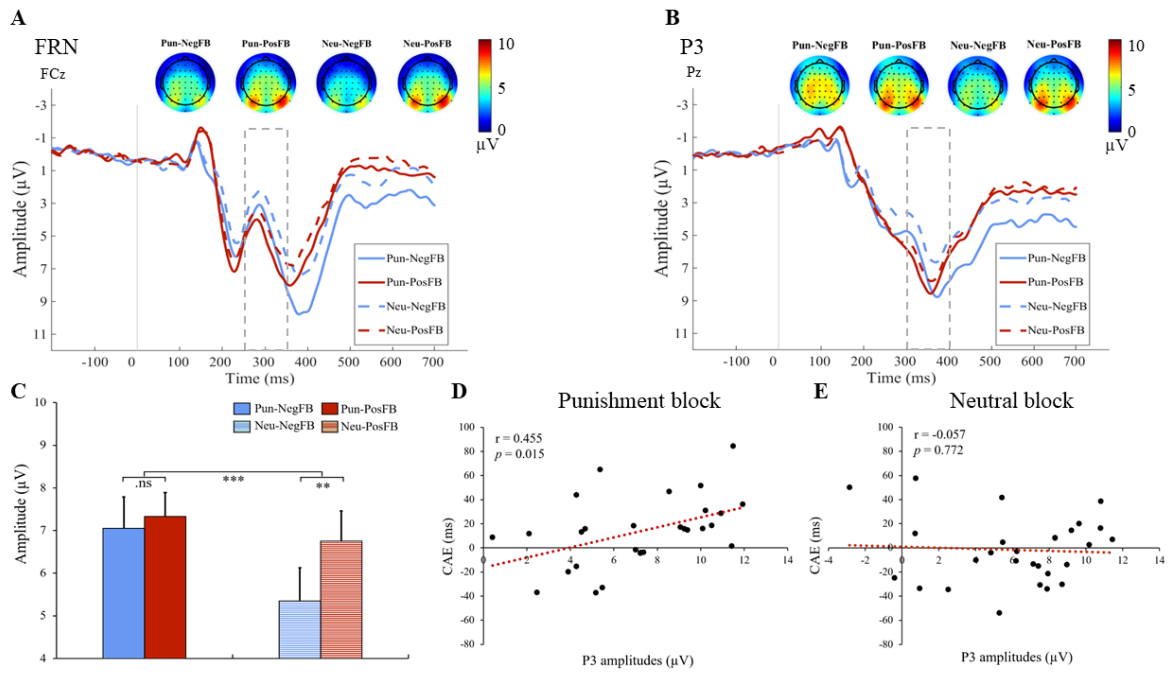


Figure 4

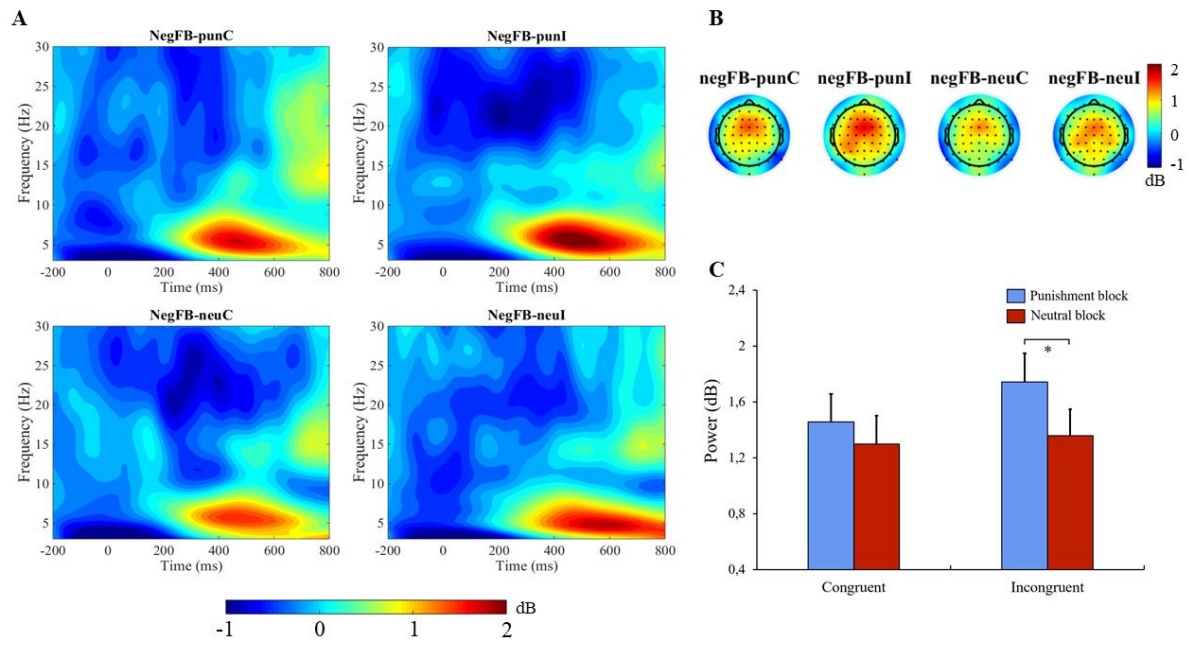


Figure 5

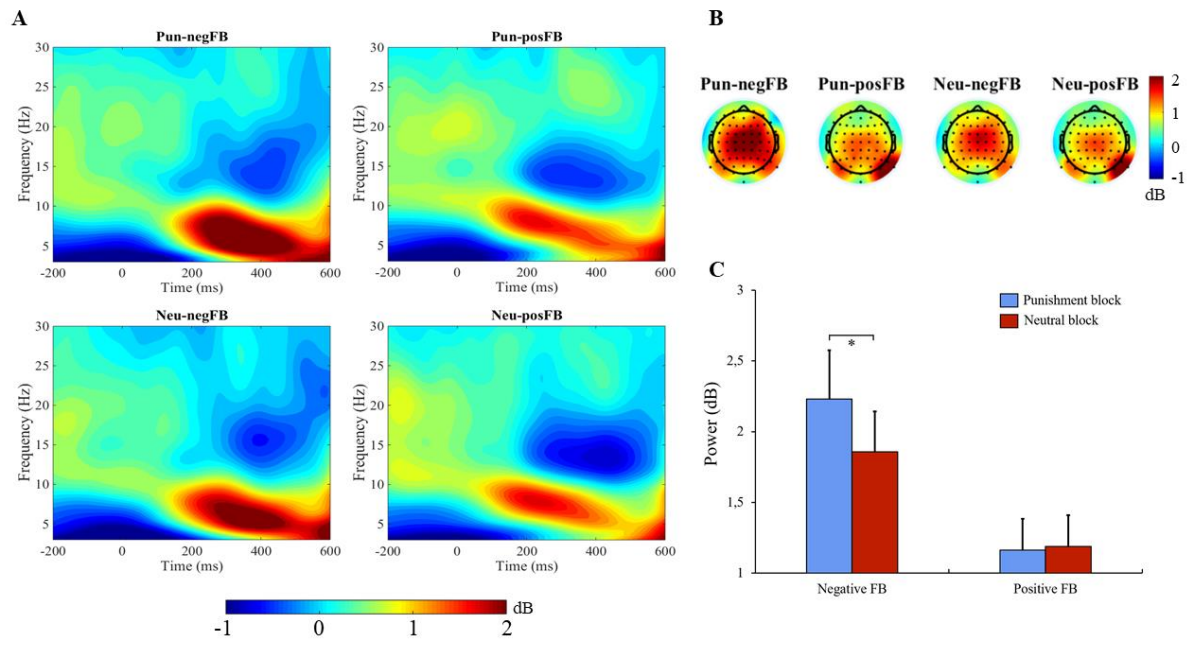


Figure 6

Table 1

Proportion (expressed in percentage) of negative feedback for each experimental condition separately

	In general	Following Congruent	Following Incongruent
Punishment block	60	56	63
Neutral block	61	58	64

Table 2

Ratings (Mean and standard deviation in parenthesis) of congruent and incongruent stimuli for punishment and neutral blocks, separately. The larger the value, the more negative the stimulus was judged by participants.

Baseline period		Punishment		Neutral	
Congruent trials	Incongruent trials	Congruent trials	Incongruent trials	Congruent trials	Incongruent trials
31(26)	50(16)	37(18)	61(17)	35(18)	60(17)

Table 3

Mean amplitude (and standard deviation) expressed in μV of the FRN and P3 elicited for negative and positive feedback in punishment and neutral blocks, separately.

	Punishment block		Neutral block	
	Negative feedback	Positive feedback	Negative feedback	Positive feedback
FRN	-0.73(1.84)	0.32(1.16)	-0.61(1.43)	0.24(1.55)
P3	7.05(3.89)	7.33(2.94)	5.34(4.12)	6.75(3.72)

Table 4

Mean (and standard deviation) of MFT power expressed in dB for cC, cI, iC, iI trials in punishment and neutral blocks, separately.

Punishment block				Neutral block			
cC	cI	iC	iI	cC	cI	iC	iI
1.59(1.09)	1.94(1.20)	1.60(0.98)	1.81(1.22)	1.42(0.94)	1.42(1.20)	1.20(1.40)	1.43(1.11)

Highlights

- Defensive motivation increases conflict adaptation at the behavioural level.
- The N2 captures this modulation of defensive motivation on conflict adaptation.
- Mid-frontal theta is increased by defensive motivation and conflict processing.
- Mid-frontal theta does not covary locally with conflict adaptation.