The selective use of punishments on congruent versus incongruent trials in the Stroop task

Qian Yang^a, Jintao Xing^a, Senne Braem^b, & Gilles Pourtois^c

^aInstitute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China

^bDepartment of Experimental Psychology, Ghent University, Ghent, Belgium

^cCognitive & Affective Psychophysiology Laboratory, Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

Corresponding author:

Dr. Qian Yang

Institute of Brain and Psychological Sciences

Sichuan Normal University

Chengdu, China

Email: qianyang_psy@163.com

Abstract

Conflict adaptation refers to the dynamic modulation of conflict processing across successive trials and reflects improved cognitive control. Interestingly, aversive motivation can increase conflict adaptation, although it remains unclear through which process this modulation occurs because previous studies presented punishment feedback following suboptimal performance on both congruent and incongruent trials. According to integrative accounts of conflict monitoring and aversive motivation in the dorsal anterior cingulate cortex, punishment feedback following slow or erroneous performance on incongruent trials in particular should lead to improved conflict adaptation. Second, selectively increasing motivation on incongruent trials should reduce the overall congruency effect. The current study sought to test both hypotheses. Specifically, we administered the confound-minimized Stroop task to a large group of participants and manipulated the position of feedback (following either congruent or incongruent trials) and aversive motivation (tied to a monetary loss or not) across different blocks. As expected, the congruency effect was found to be smaller when punishment was coupled with incongruent versus congruent trials. However, results showed that conflict adaptation was increased when punishment feedback was selectively coupled with congruent rather than incongruent trials. Together, these results suggest that aversive motivation does not uniformly improve cognitive control but this gain appears to be context dependent.

Keywords: Conflict adaptation; adaptation by binding; aversive motivation; reinforcement; cognitive control

Introduction

We are often faced with conflicting information or response tendencies, and are thought to be flexible in adjusting our attentional settings to deal with subsequent conflicts more efficiently. This process is sometimes referred to as "conflict adaptation", and often studied using the congruency sequence effect (Egner, 2007), which refers to a decreased congruency effect following incongruent compared to congruent trials (Botvinick et al., 2001). According to the adaptation by binding model (Abrahamse, Braem, Notebaert, & Verguts, 2016; Verguts & Notebaert, 2009), there is a specific binding process that underlies this behavioral effect: the detection of conflict reinforces all active (usually task-relevant) associations and thereby helps the processing of subsequent incongruent versus congruent trials, resulting in a reduced congruency effect. At the neuro-anatomical level, a loop operating from the dorsal anterior cingulate cortex (dACC) to the dorsolateral prefrontal cortex (DLPFC) is thought to underpin this conflict monitor-controller system (Carter & Van Veen, 2007).

This carry-over effect from the preceding to the current trial, being driven either by increased cognitive control (Egner, 2007) or binding (Abrahamse et al., 2016; Egner, 2014), is mostly reactive and short-lived. Accordingly, conflict adaptation is usually attenuated at long compared to short inter-trial intervals (see Yang & Pourtois, 2018; Aben, Verguts, & van den Bussche, 2017; Duthoo, Abrahamse, Braem, & Notebaert, 2014; Egner, Ely, & Grinband, 2010). According to the adaptation by binding model, the binding process responsible for conflict adaptation is transient and its strength decays over time (see also Fritz & Dreisbach, 2015). In theory, one could hypothesize that this decay could be counteracted if a stimulus or feature relevant to the surmised conflict-related representation would be presented during the inter-trial interval, especially when it would be long. In fact, such an external event could serve as an extra reinforcement of the putative binding process. Evaluative feedback on task-performance provides one class of external stimuli that meets these conditions. In line with this idea,

integrative accounts of dACC function have suggested that its role as a conflict monitor is best understood as one of continuously registering performance feedback and performance prediction errors more generally (Alexander & Brown, 2011; Holroyd et al., 2004; Silvetti, Vassena, Abrahamse, & Verguts, 2018), which can be thought to exert similar reinforcing influences on ongoing binding processes.

Previous studies have shown that the inclusion of punishment-related feedback in between successive trials increased conflict adaptation (see Braem, Duthoo, & Notebaert, 2013; Van Steenbergen, Band, & Hommel, 2009; Yang & Pourtois, 2018; Yang, Paul, & Pourtois, 2019). Punishment feedback is closely related to defensive or aversive motivation¹ (Yee, Leng, Shenhav, & Braver, 2021; Lang & Bradley, 2013), and conflict processing indirectly (Botvinick, 2007). It manifests a discrepancy between the goal and the outcome or action, which in turn influences motivation (Carver & Scheier, 1990; Botvinick & Braver, 2015). Moreover, similarly to aversive motivation, conflict processing is closely related to negative affect, and perceived as aversive (Dignath, Eder, Steinhauser, & Kiesel, 2020a; Inzlicht, Bartholow, & Hirsh, 2015; Dreisbach & Fischer, 2015). During conflict processing, conflict and negative affect both contribute to performance monitoring and in turn, they improve cognitive control, as reflected by a gain in conflict adaptation (Botvinick et al., 2001; Dreisbach & Fischer, 2015; Dignath et al., 2020a). Consistent with this idea, negative affect and conflict have been found to evoke similar neural (pattern) responses in the dACC (Braem et al., 2017; Shackman et al., 2011; Vermeylen et al., 2020). By virtue of these properties, punishment feedback could easily be integrated into the active conflict-related representation (Kanske & Kotz, 2010), thereby reinforcing the cognitive control signal driving conflict processing across successive trials

¹ In our view and given the manipulation we have used here (i.e., loss-related feedback contingent on task performance), "defensive motivation" could very well be used as synonym or proxy of "aversive motivation". For the sake of consistency, we use aversive motivation throughout.

(Braem et al., 2013; King et al., 2021; Lindström, Mattsson-Mårn, Golkar & Olsson, 2013). In this framework, punishment feedback should have a beneficial effect on conflict adaptation mostly when conflict processing as such, and thus incongruent trials, are reinforced (Dignath, Johannsen, Hommel, & Kiesel, 2019, Inzlicht et al., 2015; Lindström et al., 2013). However and more generally, it should be noted that this gain in cognitive control driven by aversive motivation is probably not absolute or invariant. As convincingly discussed in a recent review article by Yee et al. (2021), the benefit of aversive motivation for cognitive control is mostly visible when conflict resolution is successful or efficient, in the sense that it allows avoiding or escaping from the (feared) negative outcome (i.e., punishment or loss). If this negative outcome cannot be avoided, aversive motivation impedes rather than facilitates cognitive processing in general (Lovibond, Saunders, Weidemann, & Mitchell, 2008), and conflict processing more specifically (Choi & Cho, 2020). Hence, beneficial effects of aversive motivation on cognitive control are probably context-dependent and if the reinforcer (i.e. punishment) cannot be avoided (for example, because it is frequent and/or it is not instrumental to performance or learning), they can be mitigated.

In the existing literature, there have been few studies assessing the modulation of conflict processing by motivation, especially when a specific contingency is created between the incentive (either reward or punishment) and conflict (i.e., incongruent trials). To the best of our knowledge, only indirect evidence for a preferential link or association between evaluative feedback (i.e., motivation) and conflict processing is available in the existing cognitive psychology literature. Moreover, this evidence mostly concerns appetitive motivation (i.e., reward) and the congruency effect focusing on conflict processing in the current trial, as opposed to aversive motivation (i.e., punishment) and conflict adaptation focusing on conflict processing across successive trials. In two recent studies (Chen, Tan, Liu, & Wang, 2020; Prével, Krebs, Kukkonen, & Braem, 2021; Experiment 3), the authors found that a smaller congruency

effect was observed in blocks where incongruent trials (correctly processed) were paired with reward selectively, compared to blocks where congruent trials were paired with this incentive instead. These results suggest that conflict processing can be improved by selective reinforcement through reward, and this effect could stem from a stronger binding created by this incentive between conflict signals and control states (Braem & Egner, 2018). However, these studies did not look at conflict adaptation and moreover, they did not use a confound-minimized design (see below).

In comparison, existing studies on conflict adaptation and aversive motivation have not sought to assess whether conflict processing selectively (i.e., incongruent trials) needed to be reinforced to influence the size of this trial-by-trial adjustment. More specifically, earlier studies mostly used punishment feedback for incorrect or slow responses, yet irrespective of whether they concerned congruent or incongruent trials (Braem et al., 2013; Stürmer, 2011; Van Steenbergen et al., 2009; Yang & Pourtois, 2018; Yang et al., 2019). Hence, in those studies, the selective association between conflict processing and aversive motivation was not reinforced. Instead, a rather unspecific binding between levels of cognitive control (active throughout the block) and aversive motivation was probably established. Despite this limitation, it must be noted that these studies found an increase of conflict adaptation under aversive motivation (see Yang & Pourtois, 2018; Yang et al., 2019). Thus, it is conceivable that this gain would be larger if the binding between punishment feedback and conflict processing selectively, would be reinforced (Dignath et al., 2019; Abrahamse et al., 2016).

The main goal of our study was to test this specific hypothesis. To this end, we used the confound-minimized Stroop task combined with evaluative feedback contingent on task performance, as in our previous studies (see Yang & Pourtois, 2018; Yang et al., 2019). In some blocks, negative feedback led to monetary loss while in others, it did not. Hence, aversive motivation was elicited in the former condition. In those two earlier studies however, when

aversive motivation was induced, punishment feedback was not associated to slow responses or errors on incongruent trials only, but all trials uniformly (i.e., congruent and incongruent). Moreover, by design, congruent and incongruent trials led to comparable amounts of slow responses and errors in these earlier studies. In comparison, in the current study, in some blocks, punishment feedback was paired with incongruent trials selectively, while in others with congruent trials, enabling us to assess whether cognitive control could benefit from the selective association created in the former condition. We hypothesized that conflict adaptation should be larger in blocks where aversive motivation was elicited compared to blocks without punishment feedback (see Yang & Pourtois, 2018; Yang et al., 2019). Moreover, we reckoned that this gain should be larger when punishment feedback was paired selectively with incongruent compared to congruent trials. Second, we also assessed if the congruency effect could change depending on the specific contingency created. We surmised that it could be smaller when incongruent trials were selectively paired with punishment compared to congruent trials (see Prével et al., 2021).

Methods

Participant

Fifty-three Dutch-speaking participants took part in this experiment, and all of them were included in the final data analyses (8 males, mean age = 21.09 years, σ^2 = 2.02). Based on our previous work (Yang & Pourtois, 2018), we ran a power analysis (using Gpower) that indicated that 50 participants had to be included in the sample, when an effect size of d = 0.45 with a power of 80% were set, and a 2 × 2 × 2 × 2 within-subject design used. They had normal or corrected-to normal vision, and no history of psychiatric or neurological disorders. On average, participants lost 2 Euro during blocks with punishment feedback and were compensated 12 Euro for their participation.

Stimuli and task

We used a confound-minimized Stroop task to control for effects of feature repetition and contingency learning (Schmidt & Weissman, 2014; Weissman, Jiang, & Egner, 2014). The Stroop stimuli consisted of four words (in Dutch) ("rood"/red, "blauw"/blue, "groen"/green, or "geel"/yellow; font size, 30 points) presented in one out of four possible colors (red, RGB: 255, 0, 0; blue, RGB: 0, 176, 240; green, RGB: 0, 255, 0; yellow, RGB: 255, 255, 0). To rule out contingency learning, two pairs of stimulus-response rules 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. According to this rule, 8 stimuli 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 a congruent and incongruent color in each block within each mapping (Mordkoff, 2012). For a given participant, each word was presented in only two of the possible four colors however (see below). To rule out feature repetitions across successive trials, different stimuli were presented across successive trials to ensure that there was no stimulus (or response) repetition for both goal-relevant (color) and goal-irrelevant (word meaning) dimensions. On each trial, participants were instructed to identify the color of the word (i.e., color naming task) as fast and accurate as possible by using four predefined keys of a response box. These four keys corresponded to four colors (i.e. red, blue, green, yellow). More specifically, they used their left middle finger to respond to red color, left index finger to blue color, right index finger to green color, and right middle finger to yellow color.

Each trial started with a fixation cross that was used as inter-trial interval (ITI), with a mean duration of 500 ms (range: 400–600 ms). After this, the Stroop stimulus was presented in the middle of the screen for 1000 ms or until a response was given, followed by a blank screen

shown for 700 ms. Finally, either an evaluative or a neutral feedback was provided. For the evaluative feedback, either a negative feedback (black cross) was provided if the response was incorrect or too slow (i.e., slower than an arbitrary time limit; see below), or a positive feedback (black tick mark) was provided if the response was correct and fast enough (i.e., within this time limit; see Figure 1). The neutral feedback (black square) indicated a response had been made, without specific information about the task performance. For the time limit, we used an algorithm that enforces fast responding (as also used in Aarts & Pourtois, 2010; Vocat, Pourtois, & Vuilleumier, 2008), and ensures a balanced proportion of positive and negative feedback on average without yielding excessive frustration (see Table 1 and Yang & Pourtois, 2018; Yang et al., 2019). 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 Figure 1 about here
Insert Table 1 about here

2.3. Procedure

We manipulated both the feedback type (Punishment vs. No-punishment) and evaluative feedback contingency (following Congruent vs. Incongruent trials) concurrently. Consequently, the experiment consisted of four sessions corresponding to four different conditions, whose order was alternated across participants: Punishment-Feedback following Congruent trials (P-FB-C, see Figure 1A), Punishment-Feedback following Incongruent trials (P-FB-I, see Figure 1B), No-punishment-Feedback following Congruent trials (N-FB-C, see Figure 1C), and Nopunishment-Feedback following Incongruent trials (N-FB-I, see Figure 1D). More specifically, for the P-FB-C and N-FB-C sessions, evaluative feedback (either negative or positive) was only presented following congruent trials, while neutral feedback was presented following incongruent trials. In comparison, for the P-FB-I and N-FB-I sessions, the evaluative feedback was only presented following incongruent trials, while neutral feedback was presented following congruent trials. In addition, negative feedback received in the two punishment conditions (P-FB-C and P-FB-I) was converted into monetary loss (2 cents), whereas in the two No-punishment conditions (N-FB-C and N-FB-I), it did not lead to any consequence. Each session included three blocks, each including 81 trials. The three blocks belonging to one condition were presented successively, and the order of conditions was random. Prior to each session, participants were instructed on the screen about whether negative feedback would be converted to monetary loss (Punishment) or not (No-punishment). However, they were not informed about whether the evaluative feedback was paired with congruent or incongruent trials. They were encouraged to make accurate and fast response throughout the experiment.

After having signed the informed consent, participants started with a short practice session that consisted of two blocks of 16 trials each, without any punishment involved. Afterwards, the experimental sessions ensued (consisting of 12 blocks divided into four sessions, see above), where self-spaced breaks were allowed in between blocks. For each block separately, stimuli were shown in a pseudo-random order within each block to lead to the same number of congruent-Congruent (cC), congruent-Incongruent (cI), incongruent-Congruent (iC) and incongruent-Incongruent (iI) trials that were used offline to compute conflict adaptation (see below). Stimuli presentation and data recording were controlled using E-Prime (Version 2.0; Psychology Software Tools Inc., Sharpsburg, PA).

Questionnaires

Positive and negative affect schedule

A Dutch version of the Positive and Negative Affect Schedule (PANAS; Engelen et al., 2006; Watson, Clark, & Tellegen, 1988) was used to measure changes in affect between the four main sessions. 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). The PANAS was administered 5 times in total: after the practice for the first time, and after each session subsequently. Each time, participants rated the 20 items on a 5-point scale ranging from 1-*Very slightly* or not at all to 5-*Extremely*. The order of these 20 items was changed across the five measurement points to avoid the use of a specific response strategy.

Dislike feeling (negative feedback)

Participants were also asked to rate their dislike feelings towards negative feedback by means of a Visual Analog Scale (VAS) ranging from 0 (not at all) to 100 (a lot) along a putative like-dislike continuum. These ratings were presented for the first time after the practice session and subsequently after each session (hence, 5 times in total). These ratings assessed whether larger dislike feelings toward negative feedback could be found when it was embedded in the Punishment compared to the No-punishment sessions. Hence, they served as second check for the aversive motivation manipulation.

Feedback position

Although we did not inform participants beforehand about the manipulation of the evaluative feedback contingency (following either congruent or incongruent trials, selectively), we asked them at the end of the experiment to rate their awareness of this manipulation using another VAS ranging from 0 (not at all) to 100 (a lot). The mean score was 49.94 (SD: 30.13;

range: 0-100). As this variable might influence the results, we included it in the statistical analyses as predictor and refer to it as "contingency score" (see below; see also Yang & Pourtois, 2022).

Data analysis

Manipulation checks

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 Feedback type (Punishment vs. No-punishment), Feedback contingency (following Congruent vs. Incongruent trials) and Affect (Negative vs. Positive) as within-subject factors. Mean ratings of dislike feelings for negative feedback were first computed. The resulting VAS values were then submitted to an ANOVA with Feedback type (Punishment vs. No-punishment) and Feedback contingency (following Congruent vs. Incongruent trials) as within-subject factors.

Behavioral data analysis

Data preprocessing, visualization and analysis were carried out in R (R Core Team, 2018), using the tidyverse (Wickham, 2017), ggplot2 (Wickham, 2016), and Ime4 (Bates, Mächler, Bolker, & Walker, 2015) packages. For each subject separately, the first trial of each block, and outlier trials (over ± 3 SD from the mean) were excluded, leading to 50256 trials kept out of 50880 trials available in total (98.8%). Similarly for the RT data, the first trial of each block, error trials, post-error trials, and outliers were excluded from further analyses, leading to 37193 trials kept out of 50880 trials available in total (73.1%). RT data were analyzed using a linear mixed model (LMM) with which RT were log-transformed, and accuracy data, which are corresponding to a categorical dependent variable, were analyzed using a generalized linear mixed model (GLMM) with binomial distribution and a logit link function. We used the mean-centered deviation coding for the four factors (Feedback contingency, Feedback type, Previous

congruency, Current congruency) and the single predictor (Contingency score).

In order to explore whether inter-individual differences in negative affect at baseline could account for modulations in conflict adaptation during the Stroop task depending on aversive motivation and Feedback contingency (see Yang & Pourtois, 2022), we used the PANAS scores (including negative and positive affect) and Dislike feelings from the Practice session as predictors (separately) in the (G)LMMs (see results in the Supplementary Materials). Moreover, in order to investigate whether the awareness of the manipulation (cf. contingency of the evaluative feedback) influenced conflict adaptation, we added the contingency score as a predictor in the LMM. However, results showed that it did not influence conflict adaptation as such, $\chi^2(1) = 0.787$, p = .375, nor the interaction between feedback type and feedback contingency on conflict adaptation, $\chi^2(1) = 2.625$, p = .105. Given this null finding for the contingency score, the full LMMs were then created without it. The full model (1) was created based on four factors (i.e., Feedback contingency, Feedback type, Previous congruency, Current congruency, see Table 3) in order to investigate whether the selective binding of punishment feedback and conflict increased conflict adaptation (as well as the congruency effect).

For the two (G)LMMs, the random intercept for each subject was added as the random effect, and four main effects, six two-way interactions, four three-way interactions, and a four-way interaction were added as the fixed effects for the models. In order to assess effects of each factor of interest (i.e., the main and interaction effects) on accuracy and RT measures for the two models, we compared models with and without that fixed effect of interest using likelihood ratio test. For each comparison, the model included all other fixed effects that would conceivably influence the data, as well as identical random effects structures, and if the interaction term turned out to be significant based on the model comparison, we then further assessed its simple effects. All data are made publicly available via the Open Science Framework (https://osf.io/2pnzq/).

Results

Manipulation checks

PANAS

The ANOVA revealed a significant main effect of Feedback type, with higher subjective ratings for the Punishment (M = 23.8, SEM = 0.70) relative to the No-punishment sessions² (M = 22.2, SEM = 0.72), F(1, 52) = 17.357, p = .001, $\eta^2 = 0.014$. The main effect of Affect was also significant, F(1, 52) = 6.043, p = .017, $\eta^2 = 0.083$, with higher ratings for positive than negative affect. The main effect of Feedback contingency was not significant, p = .629. In addition, none of the interaction effects reached significance, $Fs \leq 3.097$, $ps \ge .084$. Planned paired *t*-tests showed that levels of positive affect during the practice session (M = 32.2, SEM = 0.79) were significantly higher than the experimental session (M = 25.0, SEM = 1.09), t = -8.575, p = .001; whereas levels of negative affect during the practice session (M = 17.2, SEM = 0.77) were significantly lower than the experimental session (M = 21.0, SEM = 1.04), t = 5.135, p = .001 (Figure 2A).

Dislike feelings (negative feedback)

The ANOVA showed a significant main effect of Feedback type, with higher dislike feelings for the Punishment (M = 75.1, SEM = 2.94) relative to the No-punishment sessions (M = 64.5, SEM = 3.10), F(1, 52) = 23.150, p = .001, $\eta^2 = 0.220$ (Figure 2B). No other effects reached significance, $Fs \le 0.157$, $ps \ge 0.694$. Planned paired *t*-tests showed that dislike feelings during the practice session (M = 45.1, SEM = 3.97) were significantly lower than the experimental sessions (M = 69.8, SEM = 2.81), t = -6.603, p = .001.

 $^{^{2}}$ Of note, "No-punishment sessions" referred to Neutral blocks (after the practice) where the negative feedback had no consequence.

Insert Figure 2 about here

Conflict processing

Accuracy

The model comparison based on the fixed effects (see Table 2) revealed a significant main effect of Feedback type, $\chi^2(1) = 4.458$, p = .035, indicating higher accuracy for the Punishment than the No-punishment sessions. A significant main effect of Feedback contingency was also found, $\chi^2(1) = 4.647$, p = .031, indicating higher accuracy when the evaluative feedback followed congruent compared to incongruent trials. Moreover, the two-way interaction effect between Previous congruency and Current congruency was also significant, $\chi^2(1) = 11.578$, p= .0007. This interaction showed that accuracy for iI trials was higher compared to cI trials, z= 2.514, SE = 0.037, p = .012; whereas accuracy for iC trials was significantly lower compared to cC trials, z = -2.301, SE = 0.037, p = .021, thereby confirming the presence of conflict adaptation (Figure 3A).

Insert Figure 3 about here
Insert Table 2 about here

RTs

The model comparison based on the fixed effect (see Table 3) showed significant main effects of Previous congruency, $\chi^2(1) = 47.097$, p < .001, and Current congruency, $\chi^2(1) = 382.2$, p < .001, indicating each time faster RTs for congruent than incongruent trials. The main effect

of Feedback type was also significant, $\chi^2(1) = 72.037$, p < .001, with faster RTs in the Punishment than No-punishment sessions. In addition, the two-way interaction between Feedback type and Feedback contingency was significant, $\chi^2(1) = 8.559$, p = .003. In the Nopunishment sessions, RTs were slower when the feedback followed incongruent than congruent trials, z = 3.317, SE = 0.001, p = .001; whereas in the Punishment sessions, this difference was not significant. The two-way interaction between Previous congruency and Current congruency was also significant, $\chi^2(1) = 13.551$, p = .0002. Participants made faster responses for cC trials compared to iC trials, z = 7.468, SE = 0.001, p < .001, whereas responses were significantly slower for iI trials compared to cI trials, z = 2.282, SE = 0.001, p = .023. In addition, the threeway interaction between Current congruency, Feedback type and Feedback contingency was marginally significant, $\chi^2(1) = 3.385$, $p = .066^3$. Importantly, the four-way interaction between Previous congruency, Current congruency, Feedback type and Feedback contingency was significant, $\chi^2(1) = 5.088$, p = .024, suggesting that conflict adaptation was influenced by the two latter variables.

To corroborate this assumption, two LMMs including three factors (Previous congruency, Current congruency and Feedback contingency) were computed, for the Punishment and Nopunishment sessions separately. During the No-punishment sessions (Figure 3B, left panel), the two-way interaction between Previous congruency and Current congruency was not significant, $\chi^2(1) = 1.433$, p = .231. The three-way interaction between Previous congruency, Current congruency and Feedback contingency was not significant either, $\chi^2(1) = 0.233$, p = .630. In comparison, during the Punishment sessions (Figure 3B, right panel), the three-way interaction between Previous congruency, Current congruency and Feedback contingency was significant, $\chi^2(1) = 7.296$, p = .007, indicating that when aversive motivation was elicited, the size of

³ We also assessed if the congruency effect could change depending on aversive motivation and the specific contingency created at the block level (see the Supplementary Materials).

conflict adaptation changed depending on the actual contingency of the evaluative feedback (see Figure 2C). To explore in which direction this modulation actually occurred, two LMMs including two factors (Previous congruency and Current congruency) were computed for two types of block separately. When the evaluative feedback followed congruent trials, the interaction between Previous congruency and Current congruency was highly significant, $\chi^2(1) = 21.138$, p < .001. It was explained by faster responses for iI trials compared to cI trials, z = -2.507, SE = 0.003, p = .012, whereas RTs for iC trials were significantly slower compared to cC trials, z = 3.998, SE = 0.003, p = 0.0001. In comparison, when the feedback followed incongruent trials, the interaction effect between Previous congruency and Current congruency was not significant, $\chi^2(1) = 0.747$, p = .387. In addition, the model comparison also indicated that in the Punishment sessions, the two-way interaction between Current congruency and Feedback contingency was significant, $\chi^2(1) = 4.579$, p = .032, indicating that the congruency effect was smaller in the FB-I relative to the FB-C sessions. In contrast, in the No-punishment sessions, the congruency effect (i.e. Current congruency) was not modulated by Feedback contingency, $\chi^2(1) = 0.274$, p = .601 (see Figure 2D).

Insert Table 3 about here

Discussion

The current study sought to investigate whether the transient increase in cognitive control under aversive motivation could be strengthened using a selective reinforcement of conflict processing through punishment feedback. To this end, we extracted a standard behavioral correlate of reactive cognitive control using the confound minimized Stroop task (Braem et al., 2019), namely conflict adaptation (Egner, 2007). In some blocks, aversive motivation was elicited because negative feedback led to monetary loss, while in others, it did not. Moreover and orthogonally to aversive motivation, in some blocks the evaluative feedback followed congruent trials selectively, while in others, it followed incongruent trials instead, allowing us to assess the influence of feedback position on conflict adaptation. As a result, in some blocks, a specific contingency was created between aversive motivation and either congruent or incongruent trials. Aversive motivation was elicited by the prospect and encounter of monetary loss upon slow or incorrect responses (see also Yang & Pourtois, 2018; Yang et al., 2019 for a similar manipulation). We reasoned that the selective pairing of conflict processing (i.e., incongruent trials) with evaluative feedback should increase conflict adaptation, when aversive motivation was elicited. We also examined whether this specific pairing could decrease the congruency effect compared to the pairing with congruent trials, in line with earlier studies using reward as reinforcer.

At the subjective level, the results showed that the main manipulation was successful. Negative affect substantially increased from the practice to the experimental session, and importantly, it was higher when aversive motivation was induced (Punishment conditions vs. No-punishment conditions). However, the actual position of the evaluative feedback (i.e., Feedback contingency) did not add up to this effect. In a similar vein, participants reported higher levels of dislike feelings towards the negative feedback when it led to monetary loss (Punishment sessions) compared to no loss (No-punishment sessions), with this effect remaining unaffected by the actual position of the evaluative feedback.

Under aversive motivation, the congruency effect was smaller when incongruent trials were selectively reinforced compared to congruent trials. This result accords well with two recent studies showing a gain in cognitive control at the level of the congruency effect when conflict was selectively reinforced by reward (Chen et al., 2020; Prével et al., 2021).

When focusing on conflict adaptation, we found that higher levels of negative affect (at

baseline) led to enhanced adaptive reactive control (see the Supplementary Materials; see also Yang & Pourtois, 2022). Moreover, it was absent in the No-punishment sessions where punishment feedback had no consequence, closely replicating our previous results (see Yang & Pourtois, 2018; Yang et al., 2019). The lack of conflict adaptation in the No-punishment sessions could be explained by the fact that reactive cognitive control underlying conflict adaptation is transient and short-lived (see Egner et al., 2010), and thus highly susceptible to intervening events. Conflict adaptation was likely interrupted by the distracting trial-by-trial evaluative feedback that was devoid of motivational significance in the No-punishment sessions.

Contrary to our hypothesis, we found that conflict adaptation was larger when punishment was selectively paired with congruent (P-FB-C) rather than incongruent trials (P-FB-I). In these two conditions, a comparable level of negative affect was reported by the participants. In other words, the motivation to increase temporarily cognitive control (and/or binding) was, as expected, enhanced by aversive motivation, but surprisingly, at first sight, if and only if the processing of congruent trials was selectively reinforced by evaluative feedback. In comparison, if incongruent trials were selectively reinforced by evaluative feedback, no such gain in cognitive control was evidenced. This dissociation therefore suggests that aversive motivation is a driving force of conflict adaptation (Dignath et al., 2020a, b), but other factors also mediate the strength of the association between them. Among them, the utility or function of punishment to guide and improve performance or learning during conflict processing appears as an important factor to take into account (Yee et al., 2021).

More specifically, we suggest that aversive motivation did not improve cognitive control when incongruent trials were selectively paired with punishment feedback because by design, in about 55% of the cases, this pairing led to monetary loss. For each and every participant, a stringent and individually-calibrated response deadline was used, and on average, they all received punishment on half the trials. Hence, the frequent and unavoidable negative feedback

(informing about monetary loss) likely counteracted the rewarding nature of conflict resolution (Braem et al., 2015; Ivanchei, Braem, Vermeylen, & Notebaert, 2021; Kohli et al., 2018; Schouppe et al., 2015). In comparison, in the P-FB-C sessions, the successful resolution of conflict (on incongruent trials) was not challenged by the evaluative (and often negative, lossrelated) feedback since it followed congruent trials selectively, in turn enabling a boost in adaptive reactive control, as shown by enhanced conflict adaptation in this condition. This interpretation accords well with the recent model proposed by Yee et al (2021) suggesting that aversive motivation can promote cognitive control, but this gain is context specific: it is mostly observed if the negative outcome (i.e., punishment) can be avoided and hence motivation is instrumental to performance. This corresponds to the P-FB-C sessions in our study. In comparison, in the P-FB-I sessions, successful conflict resolution was not reinforced by the evaluative feedback because on roughly half of the trials, punishment was presented. This interpretation is also reinforced by an auxiliary data analysis where we assessed the impact of feedback valence on subsequent task performance and showing that the negative feedback shown in the P-FB-I condition led to slower RTs on the subsequent trials than in the P-FB-C condition (see Supplementary Materials). Hence, aversive motivation impeded rather than facilitated cognitive control in this condition (P-FB-I). These results are in line with the notion that punishment performs a double-duty, in the sense that it can be either beneficial or deleterious to cognitive control depending on the specific reinforcement value or contingency of the former for the latter processes (Choi & Cho, 2020; Cubillo, Makwana, & Hare, 2019; Lindström et al., 3013). Importantly, our new results therefore suggest that aversive motivation can promote adaptive reactive control (as reflected by conflict adaptation), as long as the rewarding value of successful conflict resolution is not jeopardized by an external evaluative feedback that carries punishment and/or is not associated with clear benefits at the behavioral level. In our study, the P-FB-C condition likely met these requirements.

To conclude, the current study adds to a growing literature in cognitive psychology and neuroscience (see Botvinick & Braver, 2015) showing that cognitive control, as indexed here by conflict adaptation (as well as the congruency effect), is influenced by aversive motivation. However, aversive motivation does not uniformly improve cognitive control, but this gain appears to be context dependent. If successful conflict resolution is challenged by a frequent and unavoidable loss-related feedback, then cognitive control does not benefit from it. However, if conflict processing and resolution are not jeopardized by it, then a gain in cognitive control at the level of conflict adaptation can be found.

Figure legends

Figure 1. Experimental procedure. Each trial started with a fixation cross (that lasted on average 500 ms), followed by a Stroop stimulus. A blank screen was then presented, before either an evaluative (negative or positive) or neutral feedback selectively followed either congruent or incongruent stimulus. **A.** Punishment-Feedback following congruent trials condition (P-FB-C). The evaluative feedback selectively followed congruent trials, and each negative feedback was converted to monetary loss. **B.** Punishment-Feedback following incongruent trials, and each negative feedback selectively followed incongruent trials, and each negative feedback was converted to monetary loss. **C.** No-punishment-Feedback following congruent trials, and each negative feedback was not converted to monetary loss. **D.** No-punishment-Feedback following incongruent trials, and each negative feedback was not converted to monetary loss. **D.** No-punishment-Feedback following incongruent trials, and each negative feedback was not converted to monetary loss. **D.** No-punishment-Feedback following incongruent trials, and each negative feedback was not converted to monetary loss. **D.** No-punishment-Feedback following incongruent trials, and each negative feedback was not converted to monetary loss.

Figure 2. Ratings and behavioral results. **A.** PANAS results. Levels of positive affect (PA) and negative affect (NA) were significantly higher in Punishment sessions compared to Nopunishment sessions. PANAS scores during the practice session were significantly higher (for PA) or lower (for NA) than the experimental sessions. **B.** Dislike ratings of negative feedback were significantly higher in Punishment sessions than No-punishment sessions. Dislike ratings during the practice session were significantly lower than the experimental sessions. **C, D.** Conflict adaptation effect (**C**, using the formula ([CI – CC] – [II – IC])) and the congruency effect (**D**, using the formula [Incongruent – Congruent]) were significantly higher in P-FB-C than P-FB-I sessions. However, they did not differ between N-FB-C and N-FB-I sessions. In red: Mean response time. In dark purple: individual data points. Vertical bars correspond to standard errors of the mean. ** *p* < .01, * *p* < .05.

Figure 3. Behavioral results for conflict adaptation. **A.** Accuracy in the FB-C and FB-I conditions for No-punishment sessions (left panel) and Punishment sessions (right panel) separately. Conflict adaptation was present and similar in the four conditions. **B.** RTs in the FB-C and FB-I conditions for No-punishment sessions (left panel) and Punishment sessions (right panel) separately. Conflict adaptation was present in the P-FB-C session, but not in the three other conditions. Vertical bars correspond to standard errors of the mean.

Acknowledgements

This work was supported by a Grant (201606990022) from China Scholarship Council (CSC), a co-funding (BOF) Grant (BOFCHN2017000101) from Ghent University, and a Starting Research Fund (XJ20210029) from Sichuan Normal University awarded to Qian Yang. Senne Braem is supported by an ERC Starting grant (European Union Horizon 2020 programme, Grant agreement 852570).

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Ghent University and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Reference

- Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective, & Behavioral Neuroscience, 10*(4), 479–492.
- Aben, B., Verguts, T., & Van den Bussche, E. (2017). Beyond trial-by-trial adaptation: A quantification of the time scale of cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 509.
- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, *142*(7), 693.
- Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.
- Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624.
- Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 356–366.
- Braem, S., & Egner, T. (2018). Getting a grip on cognitive flexibility. *Current directions in psychological Science*, 27(6), 470–476.
- Braem, S., King, J. A., Korb, F. M., Krebs, R. M., Notebaert, W., & Egner, T. (2017). The role of anterior cingulate cortex in the affective evaluation of conflict. *Journal of Cognitive Neuroscience*, 29(1), 137– 149.
- Braem, S., Duthoo, W., & Notebaert, W. (2013). Punishment sensitivity predicts the impact of punishment on cognitive control. *PloS one*, 8(9), e74106.
- Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, 23(9), 769–783.
- Braem, S., Coenen, E., Bombeke, K., Van Bochove, M. E., & Notebaert, W. (2015). Open your eyes for prediction errors. *Cognitive, Affective, & Behavioral Neuroscience*, 15(2), 374–380.
- Carver, C. S., & Scheier, M. (1990). Principles of self-regulation: Action and emotion. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior*, Vol. 2, pp. 3–52). The Guilford Press.

- Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 367–379.
- Chen, J., Tan, L., Liu, L., & Wang, L. (2020). Reinforcement learning of irrelevant stimulus-response associations modulates cognitive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*
- Choi, J. M., & Cho, Y. S. (2020). Impaired cognitive control during reward pursuit and punishment avoidance. *Motivation and Emotion*, 44(6), 832–845.
- Cubillo, A., Makwana, A. B., & Hare, T. A. (2019). Differential modulation of cognitive control networks by monetary reward and punishment. *Social Cognitive and Affective Neuroscience*, *14*(3), 305–317.
- Duthoo, W., Abrahamse, E. L., Braem, S., & Notebaert, W. (2014). Going, going, gone? Proactive control prevents the congruency sequence effect from rapid decay. *Psychological Research*, 78(4), 483–493.
- Dignath, D., Eder, A. B., Steinhauser, M., & Kiesel, A. (2020a). Conflict monitoring and the affective-signaling hypothesis—An integrative review. *Psychonomic Bulletin & Review*, 27(2), 193–216.
- Dignath, D., Wirth, R., Kühnhausen, J., Gawrilow, C., Kunde, W., & Kiesel, A. (2020b). Motivation drives conflict adaptation. *Motivation Science*, *6*(1), 84.
- Dignath, D., Johannsen, L., Hommel, B., & Kiesel, A. (2019). Reconciling cognitive-control and episodicretrieval accounts of sequential conflict modulation: Binding of control-states into event-files. *Journal of Experimental Psychology: Human Perception and Performance*, 45(9), 1265.
- Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, *24*(4), 255–260.
- Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, *5*, 1247.
- Egner, T., Ely, S., & Grinband, J. (2010). Going, going, gone: characterizing the time-course of congruency sequence effects. *Frontiers in Psychology*, *1*, 154.
- Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 380–390.
- 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. *Gedrag en Gezondheid*, *34*(2), 61–70.
- Fritz, J., & Dreisbach, G. (2015). The time course of the aversive conflict signal. *Experimental psychology*.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7(5), 497–498.

- Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*(3), 126–132.
- Ivanchei, I. I., Braem, S., Vermeylen, L., & Notebaert, W. (2021). Correct responses alleviate the negative evaluation of conflict. *Quarterly Journal of Experimental Psychology*, 74(6), 1083–1095.
- Kanske, P., & Kotz, S. A. (2010). Modulation of early conflict processing: N200 responses to emotional words in a flanker task. *Neuropsychologia*, 48(12), 3661–3664.
- King, J. A., Braem, S., Korb, F. M., Diekmann, L-S., Roessner, V., Steding, J., & Ehrlich, S. (2021). The impact of punishment on cognitive control in a clinical population characterized by heightened punishment sensitivity. *Journal of Abnormal Psychology*.
- Kohli, A., Blitzer, D. N., Lefco, R. W., Barter, J. W., Haynes, M. R., Colalillo, S. A., ... & Zink, C. F. (2018).
 Using Expectancy Theory to Quantitatively Dissociate the Neural Representation of Motivation From Its Influential Factors in the Human Brain: An fMRI Study. *NeuroImage*, *178*, 552–561.
- Lang, P. J., & Bradley, M. M. (2013). Appetitive and defensive motivation: goal-directed or goaldetermined? *Emotion Review*, 5(3), 230–234.
- 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(6), e65692.
- Lovibond, P. F., Saunders, J. C., Weidemann, G., & Mitchell, C. J. (2008). Evidence for expectancy as a mediator of avoidance and anxiety in a laboratory model of human avoidance learning. *Quarterly Journal of Experimental Psychology*, 61(8), 1199–1216.
- Mordkoff, J. T. (2012). Observation: Three reasons to avoid having half of the trials be congruent in a fouralternative forced-choice experiment on sequential modulation. *Psychonomic Bulletin & Review*, 19(4), 750–757.
- Prével, A., Krebs, R., Kukkonen, N., & Braem, S. (2021). Selective reinforcement of conflict processing in the Stroop task.
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
- Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS One*, 9(7), e102337.
- 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*(1), 251–261.

- 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(3), 154–167.
- Silvetti, M., Vassena, E., Abrahamse, E., & Verguts, T. (2018). Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLoS Computational Biology*, *14*(8), e1006370.
- Stürmer, B. (2011). Reward and punishment effects on error processing and conflict control. *Frontiers in Psychology*, *2*, 335.
- 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(12), 1473–1477.
- Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Sciences*, 13(6), 252–257.
- Vermeylen, L., Wisniewski, D., González-García, C., Hoofs, V., Notebaert, W., & Braem, S. (2020). Shared neural representations of cognitive conflict and negative affect in the medial frontal cortex. *Journal of Neuroscience*, 40(45), 8715–8725.
- 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(10), 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(6), 1063.
- Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, 40(5), 2022.
- Wickham, H. (2017). The tidyverse. R package ver, 1(1), 836.
- Wickham, H. (2016). Ggplot2: Create elegant data visualisations using the grammar of graphics. Retrieved from https://ggplot2.tidyverse.org
- Yang, Q., & Pourtois, G. (2018). Conflict-driven adaptive control is enhanced by integral negative emotion on a short time scale. *Cognition and Emotion*, 32(8), 1637–1653.
- Yang, Q., Paul, K., & Pourtois, G. (2019). Defensive motivation increases conflict adaptation through local changes in cognitive control: evidence from ERPs and mid-frontal theta. *Biological Psychology*, 148, 107738.
- Yang, Q., & Pourtois, G. (2022). Modulation of conflict processing by reappraisal: An experimental investigation. *Brain Sciences*, 12(5), 564.

Yee, D. M., Leng, X., Shenhav, A., & Braver, T. S. (2021). Aversive Motivation and Cognitive Control. *Neuroscience & Biobehavioral Reviews*.

Table 1. Proportion (expressed in percentage) of negative feedback for each session separately.

Session types	Punishment-FBc	Punishment-FBi	No-punishment-FBc	No-punishment-FBi
Negative FB (%)	56.04	56.26	56.96	56.45

Predictor	Estimate	SE	z-value	$Pr\left(>/z/\right)$
(Intercept)	1.936	0.079	24.23	< 2e-16 ***
PreCon	-0.003	0.026	-0.121	0.904
CurCon	0.031	0.026	1.204	0.228
FB type	0.055	0.026	2.114	0.035 *
FB contingency	0.056	0.026	2.158	0.031 *
PreCon: CurCon	0.178	0.052	3.405	0.001 ***
PreCon: FB type	0.086	0.052	1.647	0.099 .
CurCon: FB type	0.046	0.052	0.894	0.371
PreCon: FB contingency	0.065	0.052	1.242	0.214
CurCon: FB contingency	-0.047	0.052	-0.902	0.367
FB type: FB contingency	0.071	0.052	1.364	0.172
PreCon: CurCon: FB type	0.017	0.104	0.169	0.866
PreCon: CurCon: FB contingency	-0.030	0.104	-0.293	0.769
PreCon: FB type: FB contingency	0.062	0.104	0.593	0.553
CurCon: FB type: FB contingency	0.052	0.104	0.499	0.618
PreCon: CurCon: FB type: FB contingency	-0.011	0.209	-0.055	0.956

Table 2. Summary of fixed effects for accuracy analysis.

'***' < .001; '**' < .01; '*' < .05; '.' < .1
PreCon: Previous congruency; CurCon: Current congruency; FB type: Feedback type; FB contingency:</pre> Feedback contingency

Predictor	Estimate	SE	t value	$Pr\left(> z \right)$
(Intercept)	2.671e+00	7.385e-03	361.699	< 2e-16 ***
PreCon	-6.223e-03	9.067e-04	-6.864	6.82e-12 ***
CurCon	-1.777e-02	9.066e-04	-19.596	< 2e-16 ***
FB type	-7.699e-03	9.068e-04	-8.490	< 2e-16 ***
FB contingency	-1.575e-03	9.072e-04	-1.737	0.083.
PreCon: CurCon	-6.672e-03	1.813e-03	-3.681	0.000 ***
PreCon: FB type	2.783e-04	1.813e-03	0.154	0.878
CurCon: FB type	-7.131e-04	1.813e-03	-0.393	0.694
PreCon: FB contingency	8.319e-03	1.813e-03	4.589	4.48e-06 ***
CurCon: FB contingency	-2.251e-03	1.813e-03	-1.241	0.214
FB type: FB contingency	5.306e-03	1.814e-03	2.925	0.003 **
PreCon: CurCon: FB type	-6.076e-03	3.625e-03	-1.676	0.094 .
PreCon: CurCon: FB contingency	-5.816e-03	3.625e-03	-1.604	0.109
PreCon: FB type: FB contingency	-1.052e-03	3.626e-03	-0.290	0.772
CurCon: FB type: FB contingency	-6.670e-03	3.626e-03	-1.840	0.066 .
PreCon: CurCon: FB type: FB contingency	-1.635e-02	7.251e-03	-2.255	0.024 *

Table 3. Summary of fixed effects for RTs analysis.

'***' < .001; '**' < .01; '*' < .05; '.' < .1
PreCon: Previous congruency; CurCon: Current congruency; FB type: Feedback type; FB contingency:</pre> Feedback contingency



Figure 1



Figure 2



Figure 3

Supplementary materials

We performed three additional analyses. (i) First, we assessed if the congruency effect could change depending on aversive motivation and the specific contingency created at the block level. In this analysis⁴, we added the valence of the preceding feedback as factor to assess whether negative feedback could have a stronger effect on the congruency effect than positive feedback (Dignath et al., 2020a; Frank, Woroch, & Curran, 2005). (ii) Second, we calculated bivariate (Pearson) correlations between the congruency effect and conflict adaptation, separately for the four main conditions: Punishment-Feedback following Congruent trials (P-FB-C), Punishment-Feedback following Incongruent trials (P-FB-C), Punishment-Feedback following Incongruent trials (N-FB-C), iii) Third, we also explored whether inter-individual differences in negative affect at baseline could account for modulations in conflict adaptation during the Stroop task depending on aversive motivation and Feedback contingency (see also Yang & Pourtois, 2022). To this end, we used the ratings from the PANAS as well as dislike feelings (towards the negative feedback).

The influence of previous-feedback on current conflict processing

Data analysis

A model based on four factors (Feedback contingency, Feedback type, Previous feedback valence, Current congruency) was created. Log-transformed RT data were analyzed using a linear mixed model (LMM) (see Table 2), and accuracy data were analyzed using a generalized linear mixed model (GLMM) with binomial distribution and a logit link function (see Table 1).

⁴Because few trials per cell remained when we added feedback valence as fifth fixed factor in the LMM, we removed the factor "Previous congruency" to explore effect of feedback valence on subsequent conflict processing.

Results

Accuracy

The model comparison based on the fixed effect (see Table 1) showed that the two-way interaction between Feedback type and Feedback contingency was significant, $\chi^2(1) = 4.318$, p = .037. In addition, the three-way interaction between Current congruency, Feedback type, and Previous feedback valence was also significant, $\chi^2(1) = 4.794$, p = .028. Moreover, the fourway interaction between Feedback contingency, Feedback type, Previous feedback valence, and Current congruency was significant as well, $\gamma^2(1) = 4.737$, p = .030. To further explore it, two GLMMs including three factors (Previous feedback valence, Current congruency, and Feedback contingency) were computed, for the Punishment and No-punishment sessions separately. In the No-punishment sessions, the three-way interaction between Previous feedback valence, Current congruency, and Feedback contingency was not significant, $\chi^2(1) = 0.957$, p = .328. In contrast, in the Punishment sessions, the three-way interaction between Previous feedback valence, Current congruency, and Feedback contingency was significant, $\chi^2(1) = 4.150$, p = .042. This effect was explained by the fact that when the previous feedback was positive, the twoway interaction between Feedback contingency and Current congruency was significant, $\chi^2(1)$ = 7.223, p = .007. In the P-FB-I sessions, accuracy was significantly higher for incongruent than congruent trials (z = 2.539, SE = 0.117, p = 0.011), whereas accuracy did not differ between congruent and incongruent trials in the P-FB-C sessions (z = -1.250, SE = 0.111, p = 0.211). In contrast, when the previous feedback was negative, the two-way interaction between Feedback contingency and Current congruent was not significant, $\chi^2(1) = 0.001$, p = .975.

Table 1. Results for accuracy

Predictor	Estimate	SE	z-value	Pr(> z)
(Intercept)	1.96281	0.08483	23.139	< 2e-16 ***
FB contingency	0.05429	0.0374	1.452	0.147
CurCon	0.02064	0.03739	0.552	0.581
FB type	0.07217	0.03739	1.93	0.054 .
PreFBValence	-0.03716	0.03788	-0.981	0.327
FB contingency: CurCon	0.13207	0.07478	1.766	0.077.
FB contingency: FB type	0.15573	0.07477	2.083	0.037 *
CurCon: FB type	0.03285	0.07476	0.439	0.660
FB contingency: PreFBValence	0.03781	0.07558	0.5	0.617
CurCon: PreFBValence	-0.03686	0.07549	-0.488	0.625
FB type: PreFBValence	0.03482	0.07549	0.461	0.645
FB contingency: CurCon: FB type	0.10255	0.14945	0.686	0.493
FB contingency: CurCon: PreFBValence	0.11331	0.1509	0.751	0.453
FB contingency: FB type: PreFBValence	0.02651	0.1509	0.176	0.861
CurCon: FB type: PreFBValence	-0.33128	0.15075	-2.197	0.028 *
FB contingency: CurCon: FB type: PreFBValence	0.65856	0.30059	2.191	0.029 *
				-

·***'<.001; ·**'<.01; ·*'<.05; ·.'<.1

CurCon: Current congruency; FB type: Feedback type; FB contingency: Feedback contingency; PreFBValence: Previous-feedback valence

RTs 8 1

The model comparison based on the fixed effect (see Table 2) showed a significant main effect of Previous feedback valence, $\chi^2(1) = 108.79$, p = .001, indicating faster RTs when the previous feedback was positive than negative. The main effect of Feedback contingency was significant, $\chi^2(1) = 28.165$, p = .001, indicating faster RTs when the evaluative feedback followed congruent compared to incongruent trials. In addition, the two-way interaction between Previous-feedback valence and Feedback contingency was significant, $\chi^2(1) = 21.683$, p = .001.

Importantly, the three-way interaction between Previous-feedback valence, Feedback type, and Feedback contingency was significant, $\chi^2(1) = 3.958$, p = .047. To further explore it, two LMMs including two factors (Previous feedback valence and Feedback contingency) were computed, for the punishment and no punishment sessions separately. In the Punishment sessions, the two-way interaction between Previous feedback valence and Feedback contingency was significant, $\chi^2(1) = 20.197$, p = .001. It was explained by the fact that when the previous feedback was negative, longer RTs were observed in the P-FB-I sessions compared to the P-FB-C sessions (z = 4.897, SE = 0.00260, p = .001). When the previous feedback was positive, RTs did not differ significantly between the P-FB-C and the P-FB-I sessions (z = -1.421, SE = 0.00253, p = .155; see sm-Figure 1A). In the No-punishment sessions, the two-way interaction between Previous-feedback valence and Feedback contingency was not significant, $\chi^2(1) = 2.554$, p = .110. Moreover, the three-way interaction between Current congruency, Feedback type, and Feedback contingency was significant, $\chi^2(1) = 4.808$, p = .028. Accordingly, two LMMs including two factors (Current congruency and Feedback contingency) were computed, for the Punishment and No-punishment sessions separately. In the Punishment sessions, the two-way interaction between Feedback contingency and Current congruency was significant, $\chi^2(1) = 16.785$, p < .001, indicating that the congruency effect (RTs[incongruent] – RTs[congruent]) was smaller in the P-FB-I sessions than the P-FB-C sessions. In the Nopunishment sessions, the two-way interaction between Feedback contingency and Current congruency was not significant, $\chi^2(1) = 0.500$, p = .479 (see sm-Figure 1A).

Table 2. Results for RTs

Predictor	Estimate	SE	t-value	Pr(>/t/)
(Intercept)	2.673	0.007223	370.069	< 2e-16 ***
FB contingency	-0.00678	0.001277	-5.307	1.1e-07 ***
CurCon	-0.01948	0.001277	-15.26	< 2e-16 ***
FB type	-0.008171	0.001277	-6.4	1.6e-10 ***
PreFBValence	-0.01338	0.001281	-10.441	<2e-16 ***
FB contingency: CurCon	-0.008609	0.002554	-3.372	0.001 ***
FB contingency: FB type	0.005626	0.002554	2.202	0.028 *
CurCon: FB type	-0.004956	0.002553	-1.941	0.052.
FB contingency: PreFBValence	0.01191	0.002559	4.656	3.3e-06 ***
CurCon: PreFBValence	0.0008671	0.002556	0.339	0.734
FB type: PreFBValence	-0.002462	0.002556	-0.964	0.335
FB contingency: CurCon: FB type	-0.01119	0.005106	-2.192	0.029 *
FB contingency: CurCon: PreFBValence	-0.00076	0.005113	-0.149	0.882
FB contingency: FB type: PreFBValence	0.01017	0.005112	1.989	0.047 *
CurCon: FB type: PreFBValence	-0.003627	0.00511	-0.71	0.478
FB contingency: CurCon: FB type: PreFBValence	-0.01558	0.01022	-1.524	0.127

'***'<.001; '**'<.01; '*'<.05; '.'<.1</pre>

CurCon: Current congruency; FB type: Feedback type; FB contingency: Feedback contingency;

PreFBValence: Previous-feedback valence



Figure 1. A. Results for the congruency effect, when the valence of the previous feedback was considered. **B.** Correlation between the congruency effect (CE) and conflict adaptation, for each condition separately. In the Punishment sessions (upper panel), a significant positive correlation was found between conflict adaptation and the CE when the evaluative feedback followed congruent trials (i.e. FB-Congruent condition). No significant correlation was found between conflict adaptation and the CE when the evaluative feedback followed incongruent trials (i.e. FB-Incongruent condition). In the No-punishment sessions (lower panel), no

significant correlations were found between conflict adaptation and the CE, neither for the FB-Congruent nor for the FB-Incongruent condition.

Correlations

Results

The results showed a significant positive correlation between conflict adaptation and the CE (r = 0.375, p = .005) in the P-FB-C condition selectively. However, this relationship was not found in the three other conditions, r ≤ 0.243 , $p \ge .08$ (see sm-Figure 1B).

Effects of inter-individual differences in negative affect at baseline on subsequent conflict adaptation

Data analysis

To model these inter-individual differences, we used the PANAS scores (including negative and positive affect) and Dislike feelings from the Practice session, because they provided a neutral and neat baseline regarding mood and feelings, and they were recorded prior to the introduction of punishment contingent on task performance (i.e., aversive motivation) for which we found that negative affect substantially increased. We used them as predictors (separately) in the (G)LMMs.

Results

Of note, when adding them, the joint influence of Punishment and Feedback contingency on conflict processing was still present and significant. The results showed that the three-way interaction between Negative affect, Previous congruency, and Current congruency was significant: For RTs ($\chi^2(1) = 6.015$, p = .014), conflict adaptation increased with higher scores of negative affect (see sm-Figure 2A). For accuracy ($\chi^2(1) = 3.838$, p = .05), it decreased with higher scores of negative affect (see sm-Figure 2B). These results suggest that higher levels of negative affect (at baseline) led to enhanced cognitive control.



Figure 2. Conflict adaptation is modulated by negative affect, as shown by the RTs (A) and accuracy data **(B)**.