

What happens when right means wrong? The impact of conflict arising from competing feedback responses

Mathew R. Hammerstrom^{*}, Thomas D. Ferguson, Chad C. Williams, Olave E. Krigolson

Theoretical and Applied Neuroscience Laboratory, University of Victoria, Canada

ARTICLE INFO

Keywords:

Feedback evaluation
Reward positivity
P300
Cognitive load
Stroop effect
Cognitive control

ABSTRACT

Humans often rely on feedback to learn. Indeed, in learning the difference between feedback and an expected outcome is computed to inform future actions. Further, recent work has found that reward and feedback have a unique role in modulating conflict processing and cognitive control. However, it is still not clear how conflict, especially concerning the processing and evaluation of feedback, impacts learning. To address this, we examined the effects of feedback competition on feedback evaluation in a reinforcement learning task. Specifically, we had participants play a simple two-choice gambling game while electroencephalographic (EEG) data were recorded. On half of the experiment blocks, we reversed the meaning of performance feedback for each trial from its prepotent meaning to induce response conflict akin to the Stroop effect (e.g., '✓' meant incorrect). Behaviourally, we found that participants' accuracy was reduced as a result of incongruent feedback. Paralleling this, an analysis of our EEG revealed that incongruent feedback resulted in a reduction in amplitude of the reward positivity and the P300, components of the human event-related brain potential implicated in reward processing. Our results demonstrate the negative impact of conflict on feedback evaluation and the impact of this on subsequent performance.

1. Introduction

Learning is dependent upon the comparison of outcomes with expectations. Consider the following scenario of a school mathematics exam. Normally, most students have learned that when their teacher returns their marked exam a '✓' indicates a correct response whereas an '×' indicates an incorrect response. Now imagine what might happen if the teacher started grading exams the opposite way – suddenly a '✓' indicates a mistake and an '×' indicates a correct answer. With this new feedback mapping, students now have to overcome a well-learned feedback-to-outcome mapping to understand their score (Schiffer et al., 2017). The aforementioned example parallels the well-known Stroop phenomenon. In the Stroop task, participants are typically asked to report or read-aloud the colour of a word presented on a screen (Stroop, 1935). A consistent finding from this body of literature is that participants are faster and more accurate at responding when the word-colour pairing is congruent (i.e., the word blue in blue colour) than when the word-colour pairing is incongruent (i.e., the word red in blue colour; see Macleod, 1991 for a detailed review). The most prominent

explanation of the Stroop Phenomenon posits that slowed or erroneous responses arise from competition between the two potential responses (Doehman et al., 1978).

The impact of response competition is well-studied within the Stroop literature. Response conflict has been shown to have a role in the Stroop effect with the suggestion that in an incongruent Stroop case there is competition between the word reading and colour naming ability of participants (Dyer, 1973; Shichel and Tzelgov, 2018; Tillman and Wiens, 2011; Kiyonaga and Egner, 2014; Szucs and Soltész, 2010; Levin and Tzelgov, 2014) which in turn leads to slower reaction times and an increase in errors. In other words, participants have to inhibit the faster word reading response in the face of the slower colour naming response, this, in turn, produces slower responses when the two responses conflict and thus the observed incongruency effect of slower reaction time and an increase in errors. Neuroimaging of the Stroop task has shown differences in the brain regions activated by congruent and incongruent conditions. Specifically, Egner and Hirsch (2005) demonstrated with functional magnetic resonance imaging that incongruent trials elicited greater frontal activity in the anterior cingulate cortex (ACC) and

Abbreviations: EEG, electroencephalography; ERP, event-related potentials; sTPCA, spatiotemporal principal components analysis; ACC, anterior cingulate cortex; DLPFC, dorsolateral prefrontal cortex.

^{*} Corresponding author at: McKinnon Building, 3800 Finnerty Road, Rm 120, University of Victoria, Victoria, British Columbia V8P 5C2, Canada.

E-mail address: mathewhammerstrom@uvic.ca (M.R. Hammerstrom).

<https://doi.org/10.1016/j.brainres.2021.147393>

Received 9 October 2020; Received in revised form 19 February 2021; Accepted 22 February 2021

Available online 25 February 2021

0006-8993/Crown Copyright © 2021 Published by Elsevier B.V. All rights reserved.

dorsolateral prefrontal cortex (DLPFC) than in congruent Stroop trials, a result the authors interpreted as evidence for increased cognitive control in the incongruent condition to resolve response conflict. Indeed, others have identified the ACC-DLPFC cortical-subcortical circuit as a locus for cognitive control computations (Vanderhasselt et al., 2009; Floden et al., 2011; Mitchell, 2010).

Interestingly, the ACC has also been posited to play a role in feedback evaluation (Holroyd and Coles, 2002). As such, it stands to reason that competition between feedback processes may lead to conflict. For instance, research by Krebs et al. (2010) has shown a link between reward and conflict processing within the context of the Stroop task. In their work, the traditional Stroop task was modified so that a subset of task colours was associated with the possibility of reward and participants were tasked with pressing a button that corresponded with the word color. For example, colour names presented in green or blue ink colours were associated with reward while colour names presented in red or yellow ink were not. While the authors did find that the potential for reward decreased reaction time, a curious finding was that subjects were slower on incongruent Stroop trials where they were shown a word that was previously associated with a rewarding colour (e.g., "GREEN" or "BLUE" in the example) compared to trials where the word was not previously associated with a rewarding colour (e.g., "YELLOW" or "RED" in the example). That is, on trials where the written word was an ink colour that was previously rewarded, subjects were slower to respond in comparison to trials where the written word was not previously an ink colour that was rewarded – suggesting that the reward information of the relevant stimulus dimension (ink colour) had generalized to an irrelevant stimulus dimension (written word) in turn causing the observed increase in reaction time on those trials relative to trials where the written word had was not an ink colour that had been rewarded previously. Krebs and colleagues' results suggest that while reward does offer a performance benefit in the Stroop task, previously learned reward associations might also inflict a cost on performance – which the authors hypothesized was due to the greater interference between the response mapping (Krebs et al., 2010; see also Lu et al., 2013). In other words, Krebs et al.'s results suggest that the impact of a previously learned reward mapping can induce greater response competition in the Stroop task in some instances.

Event-related brain potentials (ERPs) provide a means to examine the neural processes that underlie reward/feedback processing. The reward positivity (or feedback related-negativity – see Krigolson, 2018; Proudfit, 2015 for more detail) is an ERP component that is maximal around 250 ms after performance feedback (Miltner et al., 1997; Proudfit, 2015). This component is most prominent in frontocentral sites on the scalp, and source-localization indicates it reflects a signal originating from within the medial-frontal cortex (specifically, the ACC; Holroyd et al., 2004). The reward positivity is thought to reflect a positive prediction error - when feedback indicates an outcome is better than expected - in reinforcement learning tasks, including time-estimation (Holroyd and Krigolson, 2007; Holroyd et al., 2008; Becker et al., 2014; Williams et al., 2017), gambling (Yeung et al., 2005; Zhou et al., 2010; Williams et al., 2016; Hassall et al., 2013), and multi-armed bandit tasks (Hassall et al., 2019a, 2019b). Further, the amplitude of the reward positivity diminishes with learning, reflecting a decreased dependence on feedback for performance evaluation (Krigolson et al., 2009, 2014; Walsh and Anderson, 2012; Williams et al., 2020b). Also evoked by feedback, the P300 is an ERP component prominent around 200–600 ms after stimulus onset over posterior electrode sites (Duncan-Johnson and Donchin, 1977; Polich, 2007). Importantly, the P300 is thought to reflect the functional significance of feedback stimuli (Hajcak et al., 2005, 2007; Leng and Zhou, 2010; Zhou et al., 2010; Yeung et al., 2005; Sato et al., 2005; Wu and Zhou, 2009). Together, the reward positivity and P300 provide an effective means for studying feedback processing and the factors that impact it such as feedback competition or incongruency.

Previous work (i.e., Krebs et al., 2010) suggests that reward/

feedback itself can create or add to existing response conflict to impact performance. In the present experiment, we examined the impact of conflict induced by performance feedback by providing both intuitive and counter-intuitive feedback in a reinforcement learning task. Specifically, we hypothesized that *feedback competition* - a discrepancy between the prepotent meaning of a feedback stimulus and a newly learned associated meaning for a feedback stimulus - would essentially create a Stroop-like effect that would impact both behaviour and neural processing. To accomplish this, participants completed a two-armed bandit task while electroencephalographic data were recorded. During the performance of the bandit-task participants received feedback in one of two experimental conditions. In the first experimental condition, the stimulus-to-feedback mappings were congruent, a checkmark ('✓') indicated a 'Win' trial and a cross ('×') indicated a 'Loss' trial. To induce feedback competition, in the second experimental condition the feedback stimulus was incongruent with the outcome of the trial; the feedback stimulus normally associated with positive outcomes ('✓') was presented on loss trials and the feedback stimulus normally associated with negative outcomes ('×') was presented on win trials. We predicted that incongruent trials would be associated with decreased behavioural performance and a reduction in the amplitude of the reward positivity and P300 relative to congruent trials. As an underlying mechanism, we hypothesized that the misleading feedback in the incongruent condition would result in a smaller positive prediction error (less feedback value), and thus reduce reward positivity amplitude. Similarly, we predicted that because the incongruent feedback symbol would not be explicitly indicative of performance, thus the functional relevance of that feedback will be diminished and P300 amplitude would be reduced.

2. Results

2.1. Behavioural results

Participants selected the square with a higher probability of providing a 'Win' an average of 64.8% [57.4%, 72.1%] of all experimental trials. Further, we found that participants were selecting the higher probability square more often than chance, $t = 5.5108$, $p < 0.001$, $d = 1.006$, indicating that they did learn the task.

A two (condition: congruent, incongruent) by two (outcome: correct, incorrect) repeated measures ANOVA demonstrated that response time did not differ between outcomes for congruent and incongruent trials, $F(1, 29) = 1.23$, $p = 0.277$. However, participants did respond faster in incongruent blocks ($M = 214.9$ ms, [200.1 ms, 229.5 ms]) than in congruent blocks ($M = 228.0$ ms, [210.1 ms, 245.9 ms]), $t(29) = 2.19$, $p = 0.037$, $d = 0.287$.

The repeated measures ANOVA examining accuracy rate revealed an interaction between condition and outcome, $F(1, 29) = 4.54$, $p = 0.042$. Decomposing this interaction revealed that accuracy rate was higher for congruent feedback blocks ($M = 46.3\%$, [43.7%, 48.9%]) than incongruent feedback blocks ($M = 41.3\%$, [36.6%, 45.9%]), $t(29) = -1.76$, $p = 0.044$, $d = -0.093$.

Our first linear regression analysis on condition and block produced a significant reaction ($F(2, 585) = 11.8$, $p < 0.001$, $R = 0.197$), but revealed that block order was not a significant predictor of accuracy ($B = 0.0005$, $p = 0.210$). The second linear regression analysis on condition and trial also produced a significant reaction ($F(2, 1177) = 25.180$, $p < 0.001$, $R = 0.203$), but instead revealed that trial was a significant predictor of accuracy ($B = 0.003$, $p < 0.001$). In both analyses, condition was a significant predictor of accuracy (By block: $B = -0.112$, $p = 0.000$; By trial: $B = -0.080$, $p = 0.000$).

2.2. ERP results

Statistical analysis revealed that the amplitude of the reward positivity at channel FCz was greater for congruent outcomes ($M = 2.8$ μ V, [1.4 μ V, 4.3 μ V]) than for incongruent outcomes ($M = -0.12$ μ V, [-1.2

μV , $0.92 \mu\text{V}$], $t(29) = 3.486$, $p = 0.002$, $d = 0.881$, see Fig. 1. Examination of the P300 at channel Pz also revealed that the amplitude of this component was greater for congruent outcomes ($M = 3.5 \mu\text{V}$, [$2.5 \mu\text{V}$, $4.7 \mu\text{V}$]) than for incongruent outcomes ($M = -1.0 \mu\text{V}$, [$-1.9 \mu\text{V}$, $-0.1 \mu\text{V}$]), $t(29) = 6.579$, $p = 0.000$, $d = 1.686$, see Fig. 2.

2.3. stPCA results

stPCA analysis revealed four spatial factors, accounting for 57.1%, 10.3%, 7.6%, and 7.2% of the variance, respectively. An examination of the spatial factor loadings for each of these factors revealed that the second factor had maximal loadings in a frontal-central location consistent with the topography of the reward positivity (see Fig. 3). Further investigation of this spatial factor via temporal PCA revealed a

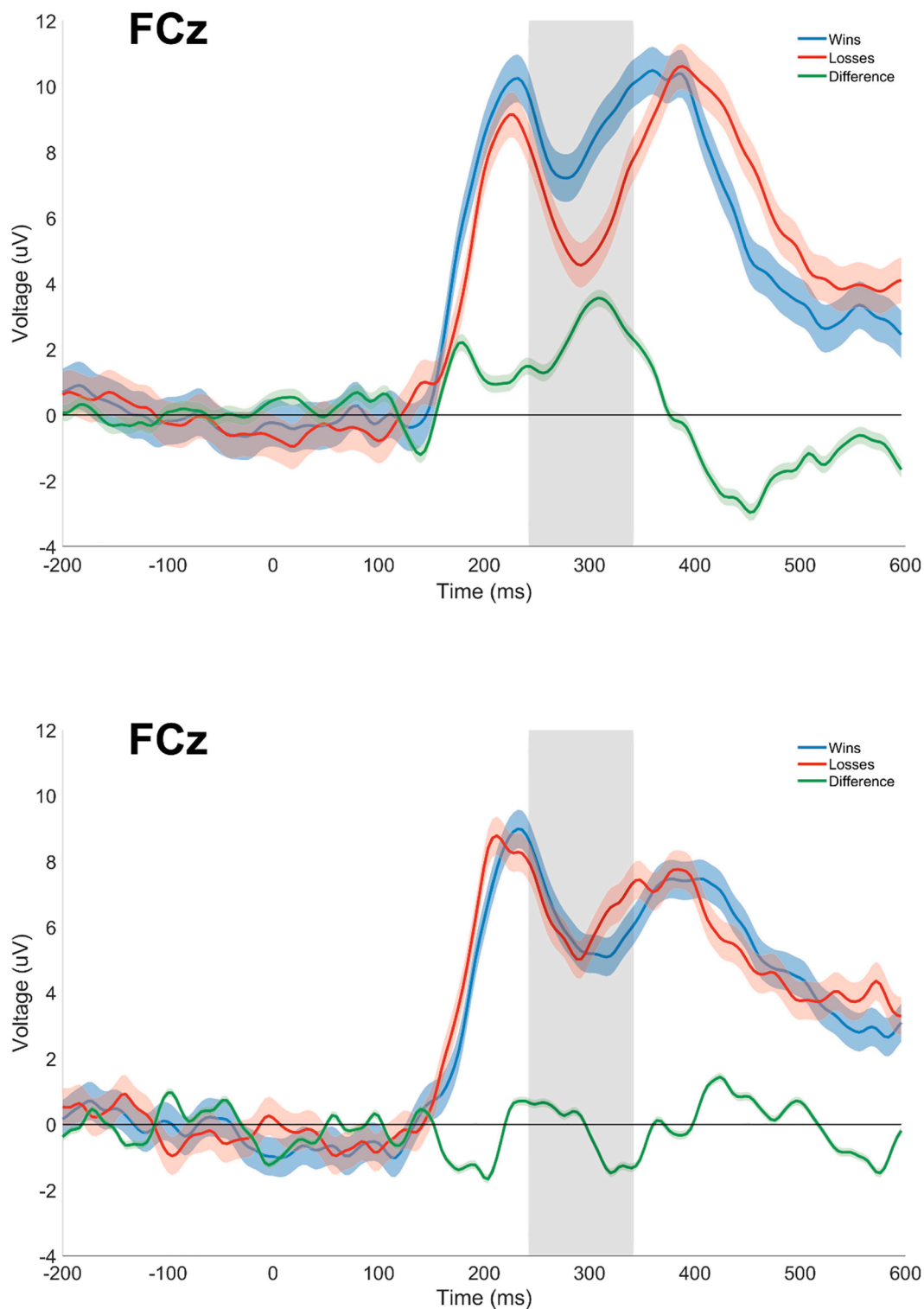


Fig. 1. Grand average conditional ERP waveforms for Channel FCz. Top Panel: Conditional waveforms for congruent trials. Bottom Panel: Conditional waveforms for incongruent trials. Shading reflects 99% Within-subjects confidence intervals.

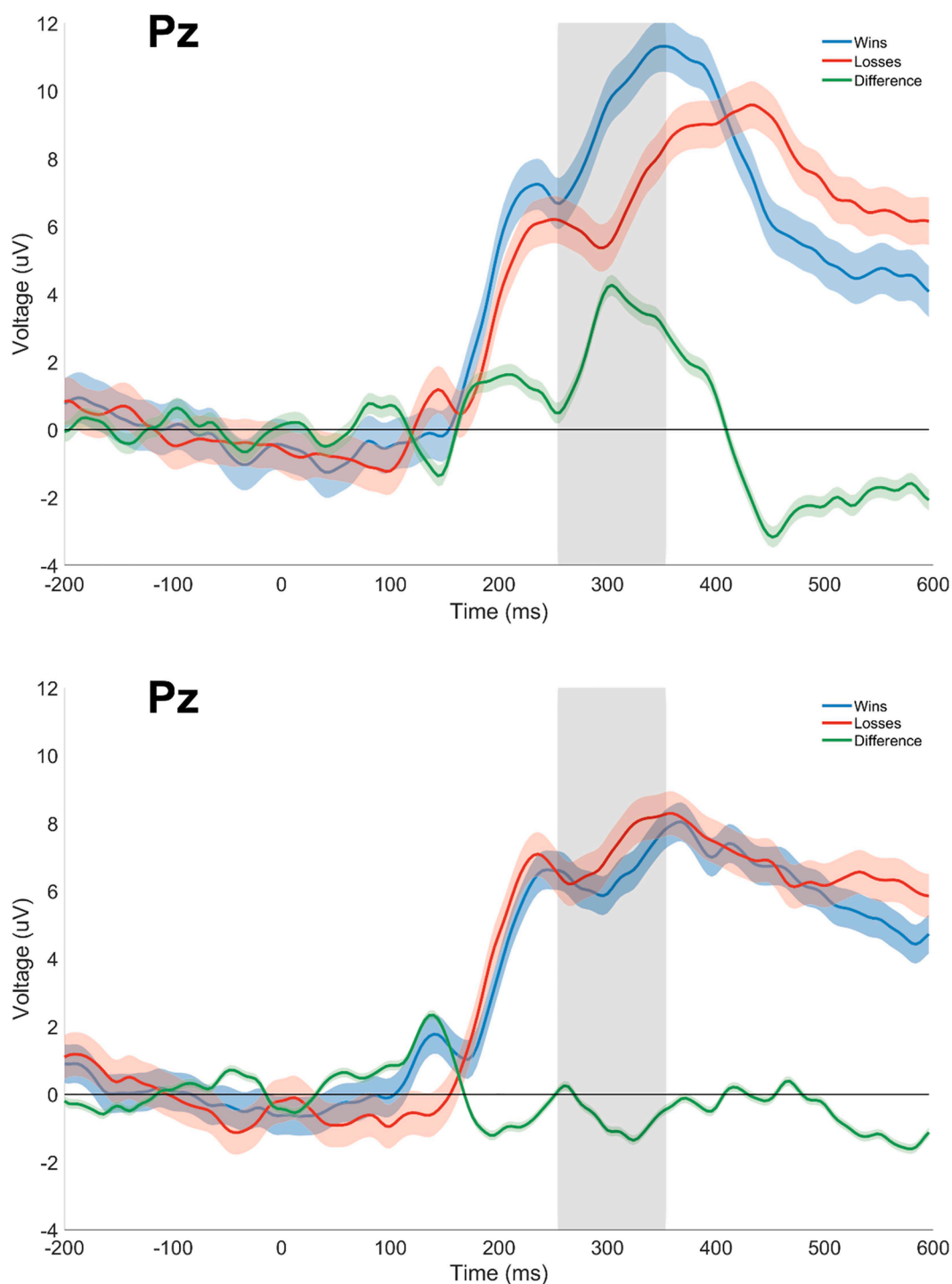


Fig. 2. Grand average conditional ERP waveforms for Channel Pz. Top Panel: Conditional waveforms for congruent trials. Bottom Panel: Conditional waveforms for incongruent trials. Shading reflects 99% Within-subjects confidence intervals.

temporal factor that accounted for 14.6% of the variance that had maximal temporal loadings between 275 and 325 ms, consistent with the timing of the reward positivity. When examining the spatial factor loadings, we also found that the first factor had maximal loadings in a parietal location, consistent with the topography of the P300 (see Fig. 3). Further investigation of this spatial factor via temporal PCA revealed a temporal factor that accounted for 32.2% of the variance that had maximal temporal loadings between 325 and 375 ms, consistent with the timing of the P300.

3. Discussion

In the present experiment, we examined the impact of conflict arising from competing feedback responses on performance and the amplitude of event-related potentials associated with feedback processing. Our findings demonstrate that feedback incongruent with prepotent meaning affected performance and two neural signals associated with feedback evaluation. Specifically, we observed a decrease in response accuracy and a decrease in the amplitude of the reward positivity when feedback was incongruent with meaning. Analysis of the P300 revealed

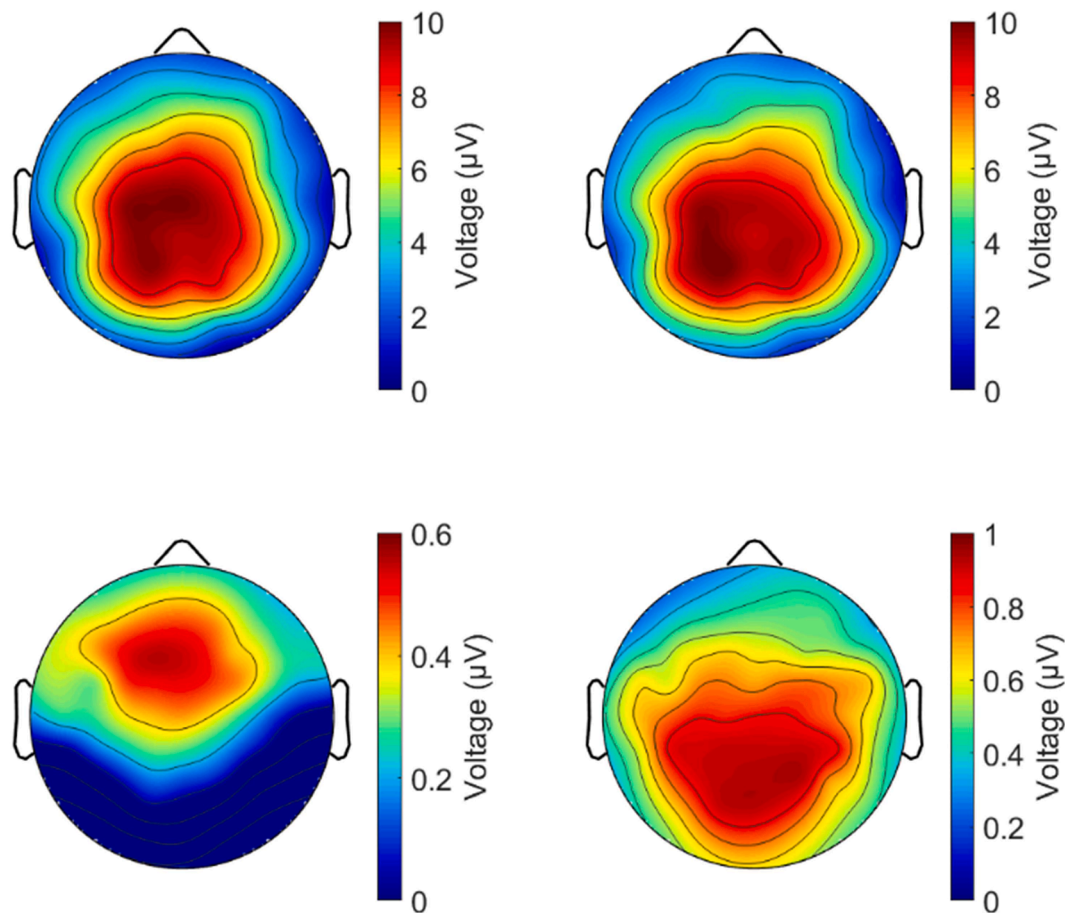


Fig. 3. Topographic maps. Top Panel: Grand average topographic maps for congruent wins, centered on maximal points for the reward positivity (left) and the P300 (right). Bottom Panel: Topographic maps for stPCA analysis. Left: Spatial factor two, accounting for 10.3% of the variance. Right: Spatial factor one, accounting for 57.1% of the variance.

the same pattern of results, a reduction in component amplitude in the incongruent feedback condition. Further, a spatiotemporal principal component analysis affirmed the ERP components we observed were indeed the reward positivity and the P300. Our results can be explained with two different, but convergent, explanations.

Our results can be interpreted within the framework of previous investigations from the Stroop paradigm – extending findings from the Stroop paradigm to conflict/competition during feedback processing. We suggest that our accuracy and ERP effects, lower accuracy and a reduction in the amplitude of the reward positivity and the P300, for incongruent as compared to congruent trials are similar to the traditional Stroop effect brought about by response conflict (Stroop, 1935; MacLeod, 1991; see also Entel et al., 2015). Specifically, the reduction in component amplitude that we observed could be driven by an increase in N2 amplitude related to an increase in cognitive control and conflict monitoring (Folstein and Van Petten, 2007; Yeung and Cohen, 2006). In our paradigm, on incongruent trials, the presentation of stimuli with well-learned associations with correct and incorrect outcomes (the ‘✓’ and the ‘×’) were superimposed with actual meanings that were opposite to their prepotent interpretation. We posit this created conflict similar to the response conflict that is observed in a traditional Stroop paradigm. However, in our case, the conflict is related to feedback-outcome mappings, or competition between what the feedback symbol traditionally indicates and what it indicated within the incongruent trials. In line with previous work (i.e., MacLeod, 1991; Krebs et al., 2010), the conflict resulted in reduced performance and, in a novel finding, also a reduction in ERP components associated with reward processing (i.e., the reward positivity and the P300). Supporting

this, prior work has shown that congruency effects with reward predictive stimuli increase response conflict in the Stroop paradigm (Krebs et al., 2010). In our task, the prepotent feedback-reward association (‘✓’ means correct) may have overcome the newly learned stimulus association (‘✓’ means wrong) in incongruent blocks, disrupting performance in incongruent trials. Indeed, our findings and those reported by Krebs and colleagues suggest that rewarding stimuli can bias congruent/incongruent associations and thus may have done so in the present experiment via prepotent feedback-reward associations.

An alternative account for our results is that incongruent feedback increased cognitive load. In other words, by reversing the meaning of feedback on incongruent trials, we created an increased demand for cognitive resources. As such, the ability of the medial-frontal system to evaluate feedback was reduced, which we saw as a reduction in accuracy and a decrease in the amplitude of the reward positivity and the P300. Indeed, our results of the reduction in the amplitudes of the reward positivity and the P300 are similar those in previous experiments from our laboratory (Krigolson et al., 2012, 2015) where performance and ERP component amplitudes were reduced by increased cognitive load. Introducing the need for two cognitive processes, feedback evaluation and conflict resolution, also introduces competition for cognitive resources. The sum of required resources for both processes was greater than those available, and as a result feedback evaluation did not have sufficient resources. Put simply, cognitive resources typically used for feedback evaluation may be instead diverted to conflict processing, resulting in a reduced or absent reward signal (Krigolson et al., 2015).

Perhaps the most satisfying explanation for our findings is a synthesis of the two accounts presented above. A unifying explanation may be that

when feedback is evaluated in the incongruent condition, additional cognitive resources are allocated to aid in conflict resolution which in effect an increase in cognitive load. As a result, cognitive resources that may have been used for the evaluation of feedback within the anterior cingulate cortex (see Holroyd et al., 2004) may have instead been used by processes within this same neural region to resolve conflict (Kerns et al., 2004; Botvinick et al., 2004; Mansouri et al., 2017). Another way of stating this would be to say that as there is an increased demand within the anterior cingulate cortex to resolve feedback conflict then is a concomitant decreased ability to evaluate feedback. We note here that we did not use source localization in the present experiment and are relying on converging evidence from a multitude of studies that have localized the reward positivity and the evaluation of response conflict to the anterior cingulate cortex (Holroyd et al., 2004; Brown and Braver, 2005; Critchley et al., 2005; Holroyd and Yeung, 2012; Rushworth et al., 2007; Alexander and Brown, 2019; Orr and Hester, 2012; Kerns et al., 2004; Botvinick et al., 2004; Mansouri et al., 2017). The next step in this research line is to dissociate the effects of cognitive load and response conflict on reinforcement learning systems.

The present research is not without limitations. First, the reward positivity is typically elicited in tasks that emphasize rewarding outcomes (ex. Williams et al., 2016; Hassall et al., 2013). In our study, we participants did not explicitly receive a reward, but rather a binary indication of performance ('Win' or 'Loss'), which could have influenced feedback processing in itself. Further, our behavioural results do not explicitly follow the expected patterns observed in Stroop tasks (increased reaction time and decreased accuracy in incongruent trials), suggesting that feedback conflict may not be similar to response conflict in terms of cognitive processing. One possible explanation is that the inter-trial waiting period was not random enough, leading to an effect of response predictability.

In conclusion, here we demonstrate that feedback conflict (and/or cognitive load) impacts a neural system associated with human reinforcement learning. Specifically, we have shown that altering the valence of feedback in a gambling task has negative effects on performance and neural learning signals. We have also presented potential explanations for these effects, namely a feedback-conflict-induced Stroop effect and cognitive load. Overall, our results suggest that competition between well-established and counterintuitive feedback mappings can lead to diminished feedback evaluation, and this perceptual conflict can be indexed by prevalent ERP components. Understanding how different sources of conflict processing can be elicited and measured in experimental paradigms is important for expanding our understanding of the Stroop phenomenon and cognitive control in general. Further, our results emphasize the prevalence of pre-existing feedback mappings in learning, which has implications for methods of providing feedback in education. Further work should aim to investigate how either of these phenomena may uniquely affect the medial learning system and other aspects of cognition.

4. Experimental procedure

4.1. Participants

Thirty undergraduate students (6 male, 24 female, mean age 21 years old [CI: ± 1 year]) from the University of Victoria were recruited to participate in this study. All participants volunteered through the University of Victoria's online research participation system and were compensated with course credit in a psychology course. Previous work in our laboratory (Williams et al., 2020a) that conducted an ERP experiment with a sample size of 500 found that detecting a reward positivity elicits a large effect size of 0.8, following recommendations from Cohen (1988). As such, we conducted a power analysis for a repeated measures *t*-test using this standardized effect size, an alpha of 0.05, and the desired power of 0.95, which yielded a prospective sample size of 22 participants. Moreover, our laboratory follows a protocol

wherein ERP studies include a minimum of 30 participants, which would correspond to a power of 0.99. To avoid conducting underpowered research (see Ioannidis, 2005) we tested participants until we had thirty data sets of sufficient behavioural and EEG data quality (actual $n = 37$, data with an EEG artifact rejection rate greater than 40% resulted in the testing of another participant). Prior to commencing the experiment, all participants provided informed consent in agreement with the guidelines established by the University of Victoria Human Research Ethics Board (Ethics Protocol Number: 16-428) and followed the ethical standards specified in the 1964 Declaration of Helsinki.

4.2. Apparatus and procedure

Participants completed a modified computer version of a two-armed bandit task (Sutton and Barto, 1998; see Fig. 4) in a sound dampened room in which EEG data was recorded. During the performance of the two-armed bandit task participants completed a series of gambles by selecting between one of two coloured squares presented on a standard 19" LCD monitor.

Each trial began with the presentation of a black fixation cross for 500 ms after which two-coloured squares also appeared onscreen for an additional 500 ms. Next, a change in the colour of the fixation cross to grey signalled participants to select one of the two squares using a computer gamepad. Following the selection of a square, both of the squares disappeared and the fixation cross changed back to black fixation for 300 to 500 ms. After this, a feedback symbol indicating the outcome of the trial appeared in place of the fixation for 1000 ms. The next experimental trial began immediately after the offset of the feedback stimulus. Based on previous work in our laboratory (Colino et al., 2020; Krigolson et al., 2017; Howse et al., 2018), one square colour had a higher probability of winning than the other square colour (60% versus 10%) to ensure an approximate number of win and loss trials per participant. As stated, these specific percentages were used to make the task learnable (i.e., one square won more often than the other so there was an optimal choice on each trial) and to ensure a win/loss rate of roughly 50% to avoid frequency contamination of the N200 (see Holroyd, 2004). As the task was learnable, the participants were instructed that their goal was to determine the "winning" coloured square each block and pick that colour to win as many times as possible. They were also informed that one of the squares was more likely to provide a win than the other, but also that both squares had the potential to be a winning square throughout the block.

In the present experiment the key manipulation related to the variation of feedback congruency. Specifically, for half of the experimental blocks feedback was congruent and participants saw a '✓' each time they won a gamble and an '×' each time they lost a gamble. For the other half of the experimental blocks, the feedback was incongruent; participants saw an '×' each time they won a gamble and a '✓' each time they lost a gamble. Thus, our definition of congruency was yoked to the notion that participants would instinctively process a '✓' as a win and an '×' as a loss during incongruent blocks and would have to override this processing with the incongruent ruleset to determine the actual gamble outcome (see Schiffer et al., 2017). Each block began with an instruction screen emphasizing the meaning of the feedback symbol (congruent: '✓' = win and '×' = loss; incongruent '×' = win and '✓' = loss) and a research assistant verified verbally that each participant knew the feedback structure for the block. Each block ended with a screen informing the participant of the number of times they had selected the optimal response option. The colours for the squares were randomly chosen each block and not repeated and the location of each square (i.e. left versus right) was randomly determined on each trial. The experiment consisted of 20 blocks of 20 trials, with 10 blocks of each condition randomly sequenced. The task was programmed in MATLAB (Version 9.6, Mathworks, Natick, USA) using the Psychophysics Toolbox extension (Brainard, 1997) and participants used a ResponsePixx, VPixx Technologies, button box to make their selection.

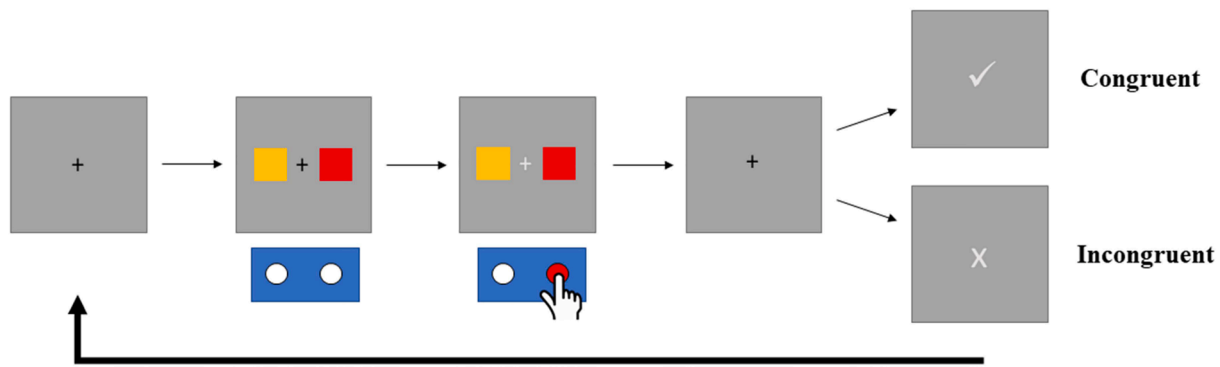


Fig. 4. A single trial of the modified two-armed bandit task. After a fixation symbol is presented, two different-coloured squares appear on the screen and the participant must select one of them using the button box. In this example, the red square gives a ‘Win’ outcome, so in congruent blocks the participants see a ‘✓’, and in incongruent blocks participants see a ‘X’.

4.3. Data acquisition

Response time (ms) and accuracy (Win or Loss) were recorded by the MATLAB (Version 9.6, Mathworks, Natick, USA) experiment script. EEG data were collected from 64 active electrodes, mounted in a 10–20 layout fitted cap (ActiCAP, Brain Products GmbH, Munich, Germany), using Brain Vision Recorder software (Version 1.21, Brain Products GmbH, Munich, Germany). All electrodes were referenced to electrode AFz during recording, and impedances were maintained below 20 k Ω at all times. EEG data were recorded at a sampling rate of 500 Hz, amplified (ActiChamp, Revision 2, Brain Products GmbH, Munich, Germany), and filtered through an antialiasing low-pass filter of 245 Hz.

4.4. Data analysis

4.4.1. Behavioural Data Analysis

For both the congruent and incongruent conditions, we computed each participant’s mean response time and accuracy rate. The response time was calculated as the mean time it took for participants to select one of the squares using the button box after the fixation cross turned grey, signalling when the participant was permitted to respond in each block. The accuracy rate was calculated as the percentage of all valid trials (excluding early responses or responses made with an invalid button) within a condition (congruent or incongruent) that resulted in ‘Win’ feedback for the participant. Additionally, to confirm participants were learning how to complete the task regardless of ‘Win’ feedback, we assessed if the participant was choosing the higher probability square more often. Next, we binned participant’s accuracy data by trial (Participant \times Trial \times Condition) and by block (Participant \times Block \times Condition).

We submitted mean response time and accuracy to a two (condition: congruent, incongruent) by two (outcome: win, loss) repeated measures ANOVA to determine if there was an interaction between condition and trial outcome. To examine significant interactions, we used paired samples t-tests ($\alpha = 0.05$), 95% confidence intervals, and effect sizes (Cohen’s d). We submitted binned trial and block data to two linear regression analysis to determine if linear combinations of Condition and trial or block, respectively, could predict accuracy. We completed this analysis to determine if any potential accuracy effects revealed in the repeated measures ANOVA were actually due to accuracy differences at the beginning and end of a given block. All statistical operations were also conducted in SPSS (Version 26.0, IBM, Armonk, USA), except for the statistical power calculation mentioned in section 4.1, which was done using the “pwr” R package (Champely et al., 2018; R Core Team, 2020).

4.4.2. EEG data processing

All EEG data were processed using MATLAB (Version 9.6,

Mathworks, Natick, USA), using the EEGLAB open-source toolbox (Delorme and Makeig, 2004) and custom software developed in the Krigolson Laboratory, available here: <https://github.com/neuro-tools>. First, channels were visually inspected for noisy data and removed accordingly. Continuous EEG data were then re-referenced to mastoid channels (TP9, TP10). A dual-pass phase free Butterworth filter with a band-pass of 0.1 Hz to 30 Hz and a 60 Hz notch filter were applied to the re-referenced EEG data. To identify and remove ocular artifacts, an independent component analysis (ICA) was conducted on the filtered data to identify components associated with ocular artifacts (Delorme and Makeig, 2004). Visual examination of the ICA factor loadings, as well as cross-correlations between EEG data and ICA component activations, guided selection of components that contained eye blinks so that they could be removed. Following the removal of components that were associated with ocular artifacts, the EEG data were reconstructed from the remaining ICA components. At this point, channels that were removed at the beginning of the analysis were re-interpolated using the method of spherical splines. All data were segmented by condition and feedback outcome into shorter epochs spanning from -200 ms to 600 ms after the stimulus presentation. The segments were then baseline-corrected using the 200 ms window prior to feedback onset. Segments were then examined for artifacts and segments of data containing a gradient larger than 10 $\mu\text{V}/\text{ms}$ or segments with an absolute difference of more than 150 μV were removed which resulted in an average of 19.9% [12.5%, 27.4%] of data being lost across participants.

4.4.3. EEG component analysis

Following artifact rejection, ERPs were created by averaging the segments for each of the two experimental conditions (congruent, incongruent) and each of the two outcomes within each condition (win, loss) resulting in separate ERP waveforms for win congruent, loss congruent, win incongruent, and loss incongruent trials. First, we created conditional difference waveforms by subtracting the average loss waveform for a condition from the average win waveform of the same condition (ex. Congruent: [congruent win – congruent loss]). Next, overall difference waveforms for each participant were created by subtracting the average loss waveform from the average win waveform collapsed across both conditions ([congruent win + incongruent win] – [congruent loss + incongruent loss]). Finally, grand average condition and difference ERPs were generated by averaging the respective individual ERP waveforms.

To ensure we did not bias our timings based on conditional effects, we used the overall difference waveforms to find our component timings and calculated ERP amplitudes based on their respective channels. We found that Channel FCz was maximal at 292 ms, so the reward positivity was then quantified for each participant and condition as the mean amplitude ± 25 ms of the grand average peak (292 ms) on the conditional difference waveforms. The P300 was then quantified for each

participant and condition as the mean amplitude \pm 25 ms of the grand average peak (304 ms) channel Pz. Using these values, we compared the congruent and incongruent difference waves using a dependent-samples *t*-test. Topographies for the reward positivity and the P300 were computed at the individual point of maximal difference.

4.4.4. Spatiotemporal principal components analysis

To provide further evidence of the topography and timing of the reward positivity and the P300 we submitted our ERP data to a spatiotemporal Principle Components Analysis (stPCA; see Dien, 2010a) using custom MATLAB code and the EP Toolkit (Dien, 2010b). The spatial data (channels) were submitted to a PCA with an Infomax rotation. Visual inspection of the resulting spatial factors guided the selection of a factor with maximal frontal-central loadings for the reward positivity and with maximal central-parietal loadings for the P300. The data for these spatial factors were then reshaped with time as the independent variable and were submitted to a temporal PCA with Promax rotation (see Dien, 2010a).

CRediT authorship contribution statement

Mathew R. Hammerstrom: Corresponding author of this manuscript. Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Thomas D. Ferguson:** Data curation, Methodology, Writing - original draft, Writing - review & editing. **Chad C. Williams:** Data curation, Visualization, Writing - original draft, Writing - review & editing. **Olave E. Krigolson:** Principal investigator for aforementioned authors. Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Software, Supervision, Writing - review & editing.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN 2016-0943) awarded to the last author.

References

- Alexander, W.H., Brown, J.H., 2019. The role of the anterior cingulate cortex in prediction error and signaling surprise. *Top. Cognit. Sci.* 11 (1), 119–135. <https://doi.org/10.1111/tops.12307>.
- Becker, M.P.I., Nitsch, A.M., Miltner, W.H.R., Straube, T., 2014. A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *J. Neurosci.* 34, 3005–3012. <https://doi.org/10.1523/JNEUROSCI.3684-13.2014>.
- Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2004.10.003>.
- Brainard, D.H., 1997. The Psychophysics toolbox. *Spatial Vision* 10 (4), 433–436. <https://doi.org/10.1163/156856897X00357>.
- Brown, J.W., Braver, T.S., 2005. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 80-.). 307, 1118–1121. <https://doi.org/10.1126/science.1105783>.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., De De Rosario, H., Rosario, M.H., 2018. Package 'pwr'. R package version, 1, 2.
- Colino, F.L., Heath, M., Hassall, C.D., Krigolson, O.E., 2020. Electroencephalographic evidence for a reinforcement learning advantage during motor skill acquisition. *Biol. Psychol.* 151, 107849. <https://doi.org/10.1016/j.biopsycho.2020.107849>.
- Cohen, J., 1988. *Statistical power analysis for the behavioral sciences*, (2nd ed.). Routledge.
- Critchley, H.D., Tang, J., Glaser, D., Butterworth, B., Dolan, R.J., 2005. Anterior cingulate activity during error and autonomic response. *Neuroimage* 27, 885–895. <https://doi.org/10.1016/j.neuroimage.2005.05.047>.
- Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Dien, J., 2010a. Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. *Psychophysiology* 47, 170–183. <https://doi.org/10.1111/j.1469-8986.2009.00885.x>.
- Dien, J., 2010b. The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *J. Neurosci. Methods* 187, 138–145. <https://doi.org/10.1016/j.jneumeth.2009.12.009>.
- Doehman, S., Landau, R., O'Connell, D., 1978. The Stroop phenomenon: perceptual conflict or response competition? *Percept. Mot. Skills* 47, 1127–1131. <https://doi.org/10.2466/pms.1978.47.3f.1127>.
- Duncan-Johnson, C.C., Donchin, E., 1977. On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology* 14, 456–467.
- Dyer, F.N., 1973. The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. *Mem. Cognit.* 1, 106–120. <https://doi.org/10.3758/BF03198078>.
- Egner, T., Hirsch, J., 2005. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* 8, 1784–1790. <https://doi.org/10.1038/nn1594>.
- Entel, O., Tzelgov, J., Bereby-Meyer, Y., Shahar, N., 2015. Exploring relations between task conflict and informational conflict in the Stroop task. *Psychol. Res.* 79, 913–927. <https://doi.org/10.1007/s00426-014-0630-0>.
- Floden, D., Vallesi, A., Stuss, D.T., 2011. Task context and frontal lobe activation in the stroop task. *J. Cogn. Neurosci.* 23, 867–879. <https://doi.org/10.1162/jocn.2010.21492>.
- Folstein, J.R., Van Petten, C., 2007. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* 45 (1), 152–170. <https://doi.org/10.1111/j.1469-8986.2007.00602.x>.
- Hajcak, G., Holroyd, C.B., Moser, J.S., Simons, R.F., 2005. Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology* 42, 161–170. <https://doi.org/10.1111/j.1469-8986.2005.00278.x>.
- Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F., 2007. It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology* 44, 905–912. <https://doi.org/10.1111/j.1469-8986.2007.00567.x>.
- Hassall, C.D., Hajcak, G., Krigolson, O.E., 2019a. The importance of agency in human reward processing. *Cogn. Affect. Behav. Neurosci.* 19, 1458–1466. <https://doi.org/10.3758/s13415-019-00730-2>.
- Hassall, C.D., McDonald, C.G., Krigolson, O.E., 2019b. Ready, set, explore! Event-related potentials reveal the time-course of exploratory decisions 1719, 183–193. <https://doi.org/10.1016/j.brainres.2019.05.039>.
- Hassall, C.D., Holland, K., Krigolson, O.E., 2013. What do I do now? An electroencephalographic investigation of the explore/exploit dilemma. *Neuroscience* 228, 361–370. <https://doi.org/10.1016/j.neuroscience.2012.10.040>.
- Holroyd, C.B., 2004. A Note on the Oddball N200 and the Feedback ERN. *Errors, Conflicts and the Brain*, 211–218. <https://doi.org/10.1037/0033-295X.109.4.679>.
- Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109 (4), 679–709. <https://doi.org/10.1037/0033-295X.109.4.679>.
- Holroyd, C.B., Krigolson, O.E., 2007. Reward prediction error signals associated with a modified time estimation task. *Psychophysiology* 44, 913–917. <https://doi.org/10.1111/j.1469-8986.2007.00561.x>.
- Holroyd, C.B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R.B., Coles, M.G.H., Cohen, J.D., 2004. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.* 7, 497–498. <https://doi.org/10.1038/nn1238>.
- Holroyd, C.B., Pakzad-Vaezi, K.L., Krigolson, O.E., 2008. The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* 45, 688–697. <https://doi.org/10.1111/j.1469-8986.2008.00668.x>.
- Holroyd, C.B., Yeung, N., 2012. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2011.12.008>.
- Howse, A.D., Hassall, C.D., Williams, C.C., Hajcak, G., Krigolson, O.E., 2018. Alcohol hangover impacts learning and reward processing within the medial-frontal cortex. *Psychophysiology* 55, e13081. <https://doi.org/10.1111/psyp.13081>.
- Ioannidis, J.P.A., 2005. Why most published research findings are false. *PLoS Med.* 2, e124. <https://doi.org/10.1371/journal.pmed.0020124>.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. Anterior cingulate conflict monitoring and adjustments in control. *Science* (80-.). 303, 1023–1026. <https://doi.org/10.1126/science.1089910>.
- Kiyonaga, A., Egner, T., 2014. The working memory stroop effect: when internal representations clash with external stimuli. *Psychol. Sci.* 25, 1619–1629. <https://doi.org/10.1177/0956797614536739>.
- Krebs, R.M., Boehler, C.N., Woldorff, M.G., 2010. The influence of reward associations on conflict processing in the Stroop task. *Cognition* 117, 341–347. <https://doi.org/10.1016/j.cognition.2010.08.018>.
- Krigolson, O.E., Hassall, C.D., Handy, T.C., 2014. How we learn to make decisions: Rapid propagation of reinforcement learning prediction errors in humans. *J. Cogn. Neurosci.* 26, 635–644. <https://doi.org/10.1162/jocn.2014.00509>.
- Krigolson, O.E., Pierce, L.J., Holroyd, C.B., Tanaka, J.W., 2009. Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. *J. Cogn. Neurosci.* 21, 1833–1840. <https://doi.org/10.1162/jocn.2009.21128>.
- Krigolson, O.E., 2018. Event-related brain potentials and the study of reward processing: Methodological considerations. *Int. J. Psychophysiol.* <https://doi.org/10.1016/j.ijpsycho.2017.11.007>.
- Krigolson, O.E., Hassall, C.D., Satel, J., Klein, R.M., 2015. The impact of cognitive load on reward evaluation. *Brain Res.* 1627, 225–232. <https://doi.org/10.1016/j.brainres.2015.09.028>.
- Krigolson, O.E., Heinekey, H., Kent, C.M., Handy, T.C., 2012. Cognitive load impacts error evaluation within medial-frontal cortex. *Brain Res.* 1430, 62–67. <https://doi.org/10.1016/j.brainres.2011.10.028>.
- Krigolson, O.E., Williams, C.C., Norton, A., Hassall, C.D., Colino, F.L., 2017. Choosing MUSE: validation of a low-cost, portable EEG system for ERP research. *Front. Neurosci.* 11, 109. <https://doi.org/10.3389/fnins.2017.00109>.

- Leng, Y., Zhou, X., 2010. Modulation of the brain activity in outcome evaluation by interpersonal relationship: An ERP study. *Neuropsychologia* 48, 448–455. <https://doi.org/10.1016/j.neuropsychologia.2009.10.002>.
- Levin, Y., Tzelgov, J., 2014. Conflict components of the Stroop effect and their “control”. *Front. Psychol.* 5, 463. <https://doi.org/10.3389/fpsyg.2014.00463>.
- Lu, A., Yang, L., Hu, S., Shao, Y., Zhang, H., Liu, L., 2013. Effects of reward and punishment on conflict processing: same or different? *Int. J. Psychol. Stud.* 5 <https://doi.org/10.5539/ijps.v5n1p22>.
- MacLeod, C.M., 1991. Half a century of research on the stroop effect: An integrative review. *Psychol. Bull.* 109, 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>.
- Mansouri, F.A., Egner, T., Buckley, M.J., 2017. Monitoring demands for executive control: shared functions between human and nonhuman primates. *Trends Neurosci* 40, 15–27. <https://doi.org/10.1016/j.tins.2016.11.001>.
- Miltner, W.H.R., Braun, C.H., Coles, M.G.H., 1997. Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a “Generic” neural system for error detection. *J. Cognit. Neurosci.* 9 (6), 788–798. <https://doi.org/10.1162/jocn.1997.9.6.788>.
- Mitchell, R.L.C., 2010. Linear increases in BOLD response associated with increasing proportion of incongruent trials across time in a colour Stroop task. *Springer* 203, 193–204. <https://doi.org/10.1007/s00221-010-2225-3>.
- Orr, C., Hester, R., 2012. Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Front. Hum. Neurosci.* 6, 177. <https://doi.org/10.3389/fnhum.2012.00177>.
- Polich, J., 2007. Updating P300: An integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118 (10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>.
- Proudfit, G.H., 2015. The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology* 52, 449–459. <https://doi.org/10.1111/psyp.12370>.
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E., Bannerman, D.M., 2007. Functional organization of the medial frontal cortex. *Curr. Opin. Neurobiol.* <https://doi.org/10.1016/j.conb.2007.03.001>.
- Sato, A., Yasuda, A., Ohira, H., ... K.M.-, 2005, undefined, n.d. Effects of value and reward magnitude on feedback negativity and P300. [journals.lww.com](https://www.journals.lww.com).
- Schiffer, A.M., Siletti, K., Waszak, F., Yeung, N., 2017. Adaptive behaviour and feedback processing integrate experience and instruction in reinforcement learning. *Neuroimage* 146, 626–641. <https://doi.org/10.1016/j.neuroimage.2016.08.057>.
- Shichel, I., Tzelgov, J., 2018. Modulation of conflicts in the Stroop effect. *Acta Psychol. (Amst)* 189, 93–102. <https://doi.org/10.1016/j.actpsy.2017.10.007>.
- Stroop, J.R., 1935. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. <https://doi.org/10.1037/h0054651>.
- Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. Cambridge: MIT Press. <https://doi.org/10.1109/TNN.1998.712192>.
- Szucs, D., Soltész, F., 2010. Stimulus and response conflict in the color-word Stroop task: A combined electro-myography and event-related potential study. *Brain Res.* 1325, 63–76. <https://doi.org/10.1016/j.brainres.2010.02.011>.
- Tillman, C.M., Wiens, S., 2011. Behavioral and ERP indices of response conflict in Stroop and flanker tasks. *Psychophysiology* 48 (10), 1405–1411. <https://doi.org/10.1111/j.1469-8986.2011.01203.x>.
- Vanderhasselt, M.-A., De Raedt, R., Baeken, C., 2009. Dorsolateral prefrontal cortex and Stroop performance: Tackling the lateralization. *Psychon. Bull. Rev.* 16, 609–612. <https://doi.org/10.3758/PBR.16.3.609>.
- Walsh, M.M., Anderson, J.R., 2012. Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2012.05.008>.
- Williams, C.C., Ferguson, T.D., Hassall, C.D., Abimbola, W., Krigolson, O.E., 2020a. The ERP, frequency, and time–frequency correlates of feedback processing: Insights from a large sample study. *Psychophysiology* 00, e13722. <https://doi.org/10.1111/psyp.13722>.
- Williams, C.C., Hassall, C.D., Lindenbach, T., Krigolson, O.E., 2020b. Reward prediction errors reflect an underlying learning process that parallels behavioural adaptations: a trial-to-trial analysis. *Comput. Brain Behav.* 3, 189–199. <https://doi.org/10.1007/s42113-019-00069-4>.
- Williams, C.C., Hassall, C.D., Trska, R., Holroyd, C.B., Krigolson, O.E., 2017. When theory and biology differ: The relationship between reward prediction errors and expectancy. *Biol. Psychol.* 129, 265–272. <https://doi.org/10.1016/j.biopsycho.2017.09.007>.
- Williams, C.C., Saffer, Y., Mcculloch, R.B., Krigolson, O.E., 2016. The scarcity heuristic impacts reward processing within the medial-frontal cortex. *Neuroreport* 27, 522–526. <https://doi.org/10.1097/WNR.0000000000000575>.
- Wu, Y., Zhou, X., 2009. The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Res.* 1286, 114–122. <https://doi.org/10.1016/j.brainres.2009.06.032>.
- Yeung, N., Cohen, J.D., 2006. The impact of cognitive deficits on conflict monitoring: Predictable dissociations between the error-related negativity and N2. *Psychol. Sci.* 17 (2), 164–171. <https://doi.org/10.1111/j.1467-9280.2006.01680.x>.
- Yeung, N., Holroyd, C.B., Cohen, J.D., 2005. ERP Correlates of Feedback and Reward Processing in the Presence and Absence of Response Choice. *Cereb. Cortex* 15, 535–544.
- Zhou, Z., Yu, R., Zhou, X., 2010. To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation. *Neuropsychologia* 48, 3606–3613. <https://doi.org/10.1016/j.neuropsychologia.2010.08.010>.