International Journal of Psychophysiology 150 (2020) 11–19

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

A tradeoff relationship between internal monitoring and external feedback during the dynamic process of reinforcement learning

Chunlei Liu*,[1], Zhenzhen Huo[1]

Department of Psychology, Faculty of Education, Qufu Normal University, Qufu, Shandong, China

ARTICLE INFO

Keywords:
Reinforcement learning
ERN
FRN
Avoidance learner
Approach learner

ABSTRACT

Effective behavior monitoring, including internal monitoring/error detection and external monitoring/feedback, is very pivotal for reinforcement learning. However, less attention has been paid to internal monitoring and the dynamic learning performance in reinforcement learning, and there is still a heated debate on which kind of external feedback is relied on in the reinforcement learning. In order to address these questions, an adaption probabilistic selection task was used to examine the effect of the internal monitoring, external feedback and the relationship between them for approach learners and avoidance learners during dynamic learning process of reinforcement learning and behavior adaption. Error-related negativity (ERN), feedback-related negativity (FRN) and feedback-related P300 are three ERPs components, which can be used as the indexes of internal monitoring, external feedback and behavior adaption. For our results, the ERN effect of avoidance learners become large in block 3, which is earlier than approach learners (block 4). This phenomenon suggests that avoidance learners learned faster than approach learners. In addition, the FRN amplitude of avoidance learners in block 4 was significantly smaller than the other three blocks. The aforementioned results demonstrated a tradeoff relationship between the ERN and FRN effects.

1. Introduction

Learning is a fundamental and pivotal cognitive competence that every person needs to respond the ongoing change of environment with this ability (Decker et al., 2016; Muller-Gass et al., 2017; Uehara et al., 2017). As reinforcement learning is one kind of learning style and can be defined as an improvement in the ability relying on external feedback, which is an important and necessary information to adjust behavior (Collins and Frank, 2018; Luque et al., 2015; West et al., 2018). Meanwhile, external feedback is a form of reinforcement learning, in which learners are usually giving two kinds of feedback: positive feedback leads to approach repetition in the similar situations, whereas negative feedback leads to same behavior avoidance in the similar situations (Shephard et al., 2014; Barch et al., 2017).

During reinforcement learning, more attentions are mainly payed on external feedback by researchers. As one of the ERPs components, feedback-related negativity (FRN) is commonly used to record changes in learning (Luft, 2014). The FRN is a negative deflection on frontocentral midline that differs in amplitude for negative and positive feedback between 200 and 400 ms after onset of the feedback stimulus (Miltner et al., 1997; Ullsperger et al., 2014; Martin, 2012; West et al., 2018; Yin et al., 2018). Typically, the FRN is larger for negative compared to positive feedback (Sambrook and Goslin, 2015), and the FRN of unexpectedness is larger than expectedness (Walsh and Anderson, 2012; Zubarev and Parkkonen, 2018). In addition, the feedback-related P300 (P300) is another indicator, which has been observed during the external feedback phase (Hämmerer et al., 2010). P300 is a positive deflection that can be observed across large parts of the head from 300 to 600 ms after feedback (Wang et al., 2017; Gheza et al., 2017). The P300 has been commonly associated with the updating of working memory in the learning process (Fischer and Ullsperger, 2013) and expectedness (Hajcak et al., 2005a; Ernst and Steinhauser, 2012). Moreover, P300 is a necessary signal to adapt behavior for the formation of good performance (Hamamé et al., 2011). A dominant Reinforcement Learning Theory posits that behavioral adaption is based on negative feedback (Holroyd and Coles, 2002), and the theory is supported by many studies (Chase et al., 2010; Ichikawa et al., 2010; Bartra et al., 2013; Lou et al., 2015; Schiffler et al., 2016). By contrast, some studies' findings showed that positive feedback could also improve performance (Baker and Holroyd, 2011; Bartra et al., 2013). However, these researches pay less attention to the dynamic learning process. And the results consist of grand average feedback amplitudes

* Corresponding author at: Faculty of Education, Qufu Normal University, Qufu City, Shandong Province, China.
E-mail address: liuchunlei@qfnu.edu.cn (C. Liu).

[1] The authors are signed as co-first author.

https://doi.org/10.1016/j.ijpsycho.2020.01.004
Received 14 July 2019; Received in revised form 10 November 2019; Accepted 21 January 2020
Available online 23 January 2020
0167-8760/

of different learning stages. FRN may not play an important role in the later learning period. Therefore, the grand average across all the trials in the experiment task may cause this kind of contradictory results. Meanwhile, there is still a heated debate on which kind of feedback is relied on in reinforcement learning. And the best way to solve this problem is the dynamic learning process analysis. First of all, researchers should discriminate that external feedback plays a pivotal role in which period. Then, researchers can ensure that behavioral adaption is relied on which kind of feedback in the important feedback period.

Behavior monitoring includes internal monitoring and external monitoring, and both of them are important. While the FRN and P300 are the external monitoring indexes, Error-related negativity (ERN) can be applied as the index of internal monitoring (Falkenstein et al., 1991; Potts, 2011). The ERN is a negative deflection after making an error at the fronto-central midline and peaks within 100 ms (Gehring et al., 2011). In addition, ERN is typically larger when errors are less frequent or less expected (Jessup et al., 2010). Furthermore, both of ERN and FRN are illustrated by the reported Reinforcement Learning Theory (Holroyd and Coles, 2002). According to the Reinforcement Learning Theory, the mesencephalic dopamine system relays an error signal to anterior cingulate cortex (ACC), where it is used to improve performance. More specifically, the influence of dopamine signals in the ACC modulates to the amplitude of ERN (Crowley et al., 2009). Outcomes worse than expected leads to phasic decreases in dopamine activity and forms large ERNs, while outcomes better than expected induces phasic increasing and appears small ERNs. When learning something new, individuals make errors easily and error detection in knowledge-based tasks necessitates some kinds of external feedback simultaneously, which indicates the appropriateness of executed response (Stahl, 2010). Subsequently, when individuals respond throughout external feedback, they can gradually recognize the rules, reduce errors and learn to detect errors (Dambacher and Hübner, 2015). Therefore, the learning process can be considered as a transition from external feedback to internal monitoring (Bultena et al., 2017). Although a considerable number of electrophysiological studies on reinforcement learning have focused on ultimate changes after learning, less attention has been paid on how dynamically change between external feedback and internal monitoring, and the relationship between them. Hence, it is necessary to explore dynamic learning performance throughout the whole reinforcement learning process.

In addition, the individual difference is an important influencing factor for reinforcement learning and some studies have investigated on age (Van den Bos et al., 2009), cognitive control capacity (Decker et al., 2016) and so on. Herein, this study focuses on the individual differences between approach learners and avoidance learners, which are discriminated by a probabilistic selection task (Frank et al., 2005; Schmid et al., 2018). Recent Reinforcement Learning Computational models suggest that two systems contribute to approach and avoidance learning (Frank et al., 2007; Frank and Claus, 2006). Approach system is related to rapid updating of reward information in working memory. Avoidance system is related to the slow integration of reward information and habitual responding (Frank and Claus, 2006). Using this computational model, approach learners learn relatively slower following positive outcomes and faster following negative outcomes. As for avoidance learners, they display reverse behavioral predispositions (Aberg et al., 2016). The current study wants to find learning differences during the dynamic process of reinforcement learning.

The present study aims to investigate two kinds of learners' dynamic learning performance change between external feedback and internal monitoring, and the relationship between them. “Which kind of feedback” and “when does the feedback happen” are important for learning. As we all know, learning is a process. In the initial learning process, ERN is characterized by a relatively small amplitude resulting from lacking subjective knowledge. In this stage, external effective feedback presents great informative value, which is reflected by a relatively large FRN amplitude. As subjective knowledge increases, learners begin to rely on their own evaluation of outcomes. An increase ERN amplitude and a decrease FRN amplitude are gradually observed. EEG researches show that an inverse/tradeoff relationship between the amplitude of ERN and FRN can be observed (Gawlowska et al., 2018; Krigolson et al., 2009). Therefore we hypothesize that FRN and ERN will play a leading role during the early learning period and the later learning period, respectively. Namely, the ERN amplitude is relatively small and the FRN amplitude is relatively large in the early learning period. The ERN amplitude is relatively large and the FRN amplitude is relatively small in the later learning period. In addition, behavioral adaption is based on negative feedback in the early period.

2. Methods

2.1. Participants

A total of seventy-eight undergraduate students from Qufu Normal University, who received either course credit or cash payment, participated in the present study. Data of thirteen participants had to be discarded for the following reasons: Five participants were excluded from analysis due to excessive electroencephalogram (EEG) artifact; three participants were excluded from analysis due to accuracy rate $< 80\%$ in the last block; five participants, who could not be classified as approach learners or avoidance learners, were excluded from the present study. Hence, the effective sample consisted of 35 approach learners (mean age = 20.4 years, *SD* = 1.65 years, 26 female, 9 male) and 30 avoidance learners (mean age = 21.19 years, *SD* = 2.02 years, 24 female, 6 male). All participants are right-handed, have normal hearing and normal or corrected-to-normal vision. None of them has a history of psychiatric, neurological or medical illness. Informed consent was obtained from all participants. The present study was approved by the local ethical committee and conducted in line with the Declaration of Helsinki.

2.2. Stimuli and task

Participants performed a probabilistic selection task, which was adapted from the previous literature (Frank et al., 2004) and consisted of learning phase and testing phase. Participants were told that they would see some pairs of Japanese symbols. They needed to learn which symbols had higher correct probability than the others in some pairs of Japanese symbols. They were asked to choose a higher correct probability symbol in each pair as soon as possible. The more positive feedback they received, the higher course credit or the more money they would receive.

During the learning phase, three different types of stimulus pairs (AB, CD, EF) presented in a randomized order. Symbols were presented in the same pairs with their placement counterbalanced across trials in learning phase. This is to say, symbol A and symbol B, symbol C and symbol D, symbol E and symbol F were always paired together in learning phase. Each kind of stimulus type included 20 trials in learning phase. On each trial, a fixation cross was present for a jittered between 600 and 1000 ms, followed by two Japanese symbols presented side-by-side on the screen. The visual angle is 8.02°. In a forced-choice paradigm, participants were asked to choose one of these symbols. Participants responded by selecting either left-hand or right-hand response symbols using the “F” or “J” key of a standard keyboard. If participants did not clear which symbol should be to choose, they could choose by following their own gut feelings. After making the choice, the blank screen was presented for 500 ms, then positive or negative feedback on the response was presented for 1500 ms. Positive feedback consists of an image of a smiling green cartoon face with the word “correct” above and negative feedback consists of an image of an angry red cartoon face the word “wrong” above. If no response was given within 4 s, a “Too slow” information would appear. Inter-trial intervals

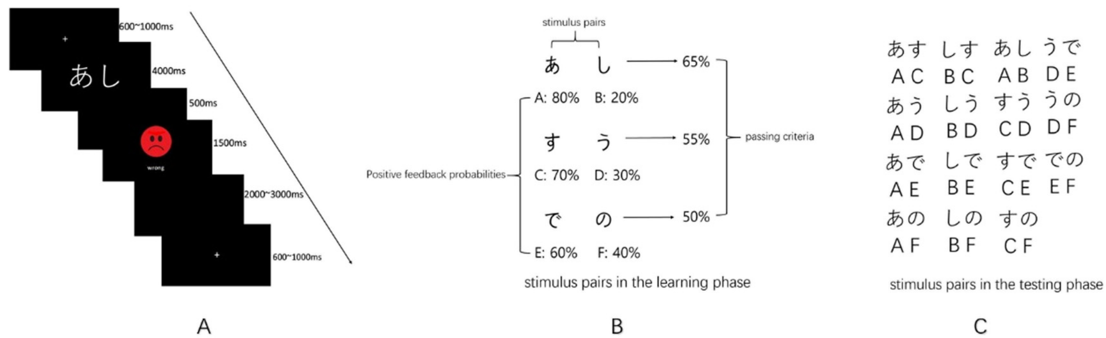

Fig. 1. (A) A trial sequence in probabilistic selection task. (B) Japanese symbols that participants learned in learning phase. (C) Japanese symbols that participants learned in testing phase.

ranged between 2000 and 3000 ms (see Fig. 1A).

Each symbol had unique feedback probabilities in learning phase. Every symbol had different positive/negative feedback frequencies in the learning phase. On AB trials, A could lead to 80% positive feedback and 20% negative feedback, whereas B could lead to 20% positive feedback and 80% negative feedback. On CD trials, C could lead to 70% positive feedback and 30% negative feedback, whereas D could lead to 30% positive feedback and 70% negative feedback. On EF trials, E could lead to 60% positive feedback and 40% negative feedback, whereas F could lead to 40% positive feedback and 60% negative feedback. Three different stimulus types had different passing criteria, and accuracy standard is 65%, 55% and 50% for AB trials, CD trials and EF trials, respectively (see Fig. 1B). Participants were aware that there was no absolute correct answer in this task, but some symbols (e.g., A) had a higher correct probability than the other. Through trial and error, participants would know that they could pass learning phase, continue to complete the testing phase and get more rewards by approaching higher probability symbols (e.g., A) and avoiding lower probability symbols (e.g., B). Participants could not proceed testing phase until a previously set passing criteria were met (Frank et al., 2005).

There were three blocks in the testing phase. Each block included 90 trials, presented in a randomized order. The stimulus types in testing phase were fifteen possible combinations composed by the six Japanese symbols from the learning phase (see Fig. 1C). The order of symbols correct probability is A > C > E > F > D > B. Participants were asked to choose a higher correct probability symbol in each pair. Thus, choosing A is correct and choosing C is wrong in AC pair; choosing C is correct and choosing B is wrong in BC pair. Stimulus arrangement in testing phase was almost the same as the learning phase, the stimulus types (AC, AD, AE, AF, BC, BD, BE, BF) were not given effective feedback. Participants were classified as approach learners or avoidance learners according to foregoing stimulus types without feedback at the testing phase (Schmid et al., 2018). The stimulus types at the testing phase would be divided into two parts to classify two types of learners. Stimulus pairs with symbol A (AC, AD, AE, AF) marked as A pairs and stimulus pairs with symbol B (BC, BD, BE, BF) marked as B pairs. A pairs and B pairs were only calculated accuracy. If accuracy in A pairs was higher than accuracy in B pairs across the three testing blocks, the participant was called an approach learner. If accuracy in B pairs was higher than accuracy in A pairs across the three testing blocks, the participant was called an avoidance learner. The remaining stimulus arrangement was the same as the learning phase.

2.3. Procedure

Participants were provided informed consent when arriving. Then experimental procedure was explained to participants. Each block included 60 trials in learning phase. The number of learning blocks was based on participant's accuracy. Only reaching or exceeding passing criteria, participants could continue to complete the testing phase. There were three blocks in testing phase, each block included 90 trials. The entire experiment task lasted for approximately 50 min.

2.4. Recording and preprocessing of electrophysiological data

EEG was recorded with a Brain Products Active Two system (Brain Products GmbH, Munich, Germany) at a 1000 Hz sampling rate from 64 active scalp electrodes placed in an electrode cap according to the standard international 10-20 system (Jasper, 1958). EEG recordings were referenced on-line to the left mastoid. The ground was positioned above the forehead. Impedances were kept below 5 kΩ. FP1 and FP2 were used to record electrooculogram (EOG).

EEG data were processed offline by using Brain Vision Analyzer 2 software. The recorded data were resampled at 250 Hz and re-referenced to an average of the left and right mastoid electrodes. EEG data were filtered offline with a low cutoff filter of 0.01 Hz and a high cutoff filter of 30 Hz. Eye blinks artifacts were semi-automatically decomposed of the clear data with the Ocular Independent Component Analysis (ICA) method on the continuous data. For feedback-locked analyses, epochs were subtracted that ranged from 200 ms before until 800 ms after feedback onset. For response-locked analyses, epochs were extracted from 200 ms before to 500 ms after the keypress response. A baseline correction was performed on the 200 ms time window before feedback-locked and response-locked data. A semi-automatic artifact correction procedure was applied to screen for artifacts and contaminate trials according to the following criteria: Any abrupt voltage change over 50 μV between adjacent sample points, any difference from peaks to peaks in a 200 ms interval that exceeded 200 μV, any amplitude that exceeded ± 100 μV, any activity that was consistently < 0.5 μV in a 100 ms interval were considered artifacts.

2.5. Statistical analyses

Accuracy and reaction time were analyzed for each condition: different types of learners (approach learners and avoidance learners) and different blocks (block 1, block 2, block 3 and block 4) for reaction time (RT), correct rates, using a mixed two-way repeated measure analysis of variance (ANOVA) of 2 (Learner type: approach learner, avoidance learner) × 4 (Period: block 1, block 2, block 3, block 4).

The ERPs, included the ERN, FRN and P300, were systematically analyzed. The ERN was evaluated as the mean amplitude within the time window, which is ranging from 50 ms before and 50 ms after the keypress response. The FRN was measured as mean amplitude within the time window 220–320 ms following the feedback outcome presentation. The P300 was measured as mean amplitude within the time window 320–420 ms after the feedback stimulus onset. The three ERPs components were calculated across five electrode locations (Fz, FCz, Cz, CPz, Pz). The results indicated that the effect of the ERN, FRN and P300 were largest at the FCz site, which is consistent with previous studies (Eppinger et al., 2008; Gheza et al., 2017; Ehlis et al., 2018). Hence, ERN, FRN and P300 analyses were conducted on electrode FCz. The ERN, FRN and P300 mean amplitudes were separately analyzed by using a mixed three-way repeated measure ANOVA. The Greenhouse-Geisser correction was conducted when sphericity violations. To control for Type-I error, Bonferroni correction was applied to post-hoc tests. In addition, Pearson correlation analyses were performed on the relationship between magnitude of ERN and FRN effects.

3. Results

3.1. Behavioral results

3.1.1. Response time

The ANOVA revealed a marginally significant main effect of learner type, $F(1,63) = 3.26$, $p = 0.08$, $\eta_p^2 = 0.05$, showing that approach learners (M = 1076.57 ms, SD = 41.53 ms) reacted a little faster than avoidance learners (M = 1186.96 ms, SD = 44.86 ms). The main effect of period was significant, $F(3,189) = 52.30$, $p < 0.001$, $\eta_p^2 = 0.45$. Polynomial orthogonal contrasts were used to examine the trends in RT of period. The results of the trend analysis indicated significant linear, $F(1,63) = 89.58$, $p < 0.001$, $\eta_p^2 = 0.58$. There was no significant quadratic ($F(1,63) = 0.64$, $p = 0.43$) and cubic ($F(1,63) = 2.60$, $p = 0.11$). That is, participants reacted faster from block1 to block4 gradually. The interaction of learner type and period did not approach significance, $F(3,189) = 0.97$, $p = 0.41$ (see Table 1).

3.1.2. Accuracy data

The main effect of period was significant, $F(3,189) = 376.77$, $p < 0.001$, $\eta_p^2 = 0.86$. Polynomial orthogonal contrasts were used to examine the trends in accuracy of period. The results of the trend analysis indicated a significant linear effect ($F(1,63) = 951.58$, $p < 0.001$, $\eta_p^2 = 0.94$), quadratic effect ($F(1,63) = 134.71$, $p < 0.001$, $\eta_p^2 = 0.68$) and cubic effect ($F(1,63) = 4.07$, $p = 0.05$, $\eta_p^2 = 0.06$). The linear effect, as expected, indicated that accuracy rates significantly improved after every block. The main effect of learner type was not significant, $F(1, 63) = 0.27$, $p = 0.61$. The period and learner type interaction did not approach significance, $F(3,189) = 1.98$, $p = 0.12$ (see Table 2).

3.2. ERP results

3.2.1. ERN

The ERN mean amplitudes were analyzed using a mixed three-way repeated measure ANOVA of 2 (Learner type: approach, avoidance) × 4 (Period: block 1, block 2, block 3, block 4) × 2 (response type: correct, error) in ± 50 ms in time window at the FCz site. The learner type was a between-subjects factor and the other two were within-subjects factors. The response type had a highly significant main effect, $F(1,63) = 21.02$, $p < 0.001$, $\eta_p^2 = 0.25$, reflecting greater negativity for error response (M = −2.53 μV, SD = 0.28 μV) than correct response (M = −1.39 μV, SD = 0.17 μV). The period had not a significant main effect, $F(3,189) = 0.12$, $p = 0.95$. The main effect of learner type was not significant, $F(1,63) = 0.02$, $p = 0.89$. Interestingly, a two-way interaction between learner type and period was significant, $F(3, 189) = 3.87$, $p = 0.01$, $\eta_p^2 = 0.06$. Trend analysis indicated a significant quadratic effect ($F(1,63) = 7.03$, $p = 0.01$, $\eta_p^2 = 0.10$) and cubic effect ($F(1,63) = 4.36$, $p = 0.04$, $\eta_p^2 = 0.07$). This result revealed that amplitudes of avoidance learners were significantly more negative than approach learners in block 3. However, amplitudes of approach learners were significantly more negative than avoidance learners in block 4. The two-way interaction of response type and period was significant, $F(3,189) = 4.32$, $p = 0.006$, $\eta_p^2 = 0.06$. Subsequently, trend analysis indicated a significant linear effect ($F(1,63) = 15.67$, $p < 0.001$, $\eta_p^2 = 0.20$). This result revealed that correct response mean amplitudes were smaller than error response mean amplitudes in block 2, block 3, and block 4. The two-way interaction of response type and learner type was not significant, $F(1,63) = 0.05$, $p = 0.82$. In addition, a reliable three-way interaction between learner type, response type and period was not obtained, $F(3,189) = 1.91$, $p = 0.13$ (see Fig. 2; see also Table S1 in the Supplemental material).

3.2.2. FRN

The FRN mean amplitudes were analyzed using a mixed three-way repeated measure ANOVA of 2 (Learner type: approach, avoidance) × 4 (Period: block 1, block 2, block 3, block 4) × 2 (feedback valence: positive, negative) in 220–320 ms in time window at the FCz site. A reliable main effect of period was obtained, $F(3,189) = 16.42$, $p < 0.001$, $\eta_p^2 = 0.21$. Polynomial orthogonal contrasts were used to examine the trends in FRN mean amplitudes of period. The results of the trend analysis indicated a significant linear effect ($F(1,63) = 27.89$, $p < 0.001$, $\eta_p^2 = 0.31$) and quadratic effect ($F(1,63) = 13.94$, $p < 0.001$, $\eta_p^2 = 0.18$). There was no significant cubic effect, $F(1,63) = 1.32$, $p = 0.26$. The linear effect, as expected, revealed that FRN mean amplitudes in block 4 were negative than other three blocks (see Fig. 3; see also Table S2 in the Supplemental material). There was no main effect of learner type, $F(1,63) = 0.61$, $p = 0.44$ and feedback valence, $F(1,63) = 0.01$, $p = 0.91$. Interestingly, there was a significant two-way interaction effect for feedback valence and period, $F(3,189) = 12.51$, $p < 0.001$, $\eta_p^2 = 0.17$. Trend analysis indicated a significant linear effect ($F(1,63) = 23.96$, $p < 0.001$, $\eta_p^2 = 0.28$) and quadratic effect ($F(1,63) = 12.66$, $p = 0.001$, $\eta_p^2 = 0.17$). This result revealed that negative feedback mean amplitudes were more negative than positive feedback in block1. The two-way interaction of learner type and period was marginally significant, $F(3,189) = 2.63$, $p = 0.052$, $\eta_p^2 = 0.04$. Trend analysis indicated a significant linear effect ($F(1,63) = 4.97$, $p = 0.03$, $\eta_p^2 = 0.07$). This result revealed that the amplitudes of approach learners were more negative than the amplitudes of avoidance learners in four blocks. The two-way interaction effect of feedback valence and learner type was not significant, $F(1,63) = 0.52$, $p = 0.48$. Moreover, a significant three-way interaction between learner type, feedback valence and period was not obtained, $F(3,189) = 0.96$, $p = 0.41$.

3.2.3. P300

The P300 mean amplitudes were analyzed using a mixed three-way repeated measure ANOVA of 2 (Learner type: approach, avoidance) × 4 (Period: block 1, block 2, block 3, block 4) × 2 (feedback valence: positive, negative) in 320–420 ms in time window at the FCz site. There was a significant main effect of feedback valence, $F(1,63) = 38.63$, $p < 0.001$, $\eta_p^2 = 0.38$, suggesting more positive for negative feedback (M = 15.64 μV, SD = 1.39 μV) than positive feedback (M = 11.15 μV, SD = 1.10 μV). A reliable main effect of period was obtained, $F(3,189) = 23.10$, $p < 0.001$, $\eta_p^2 = 0.27$ (see Fig. 3; see also Table S3 in the Supplemental material). Polynomial orthogonal contrasts were used to examine the trends in P300 mean amplitudes of

Table 1
Mean and standard deviation of reaction time in four blocks for approach learners and avoidance learners.

	Block1 (M ± SD)	Block2 (M ± SD)	Block3 (M ± SD)	Block4 (M ± SD)
Avoidance learners	1387.53 ± 287.31	1282.39 ± 341.00	1084.63 ± 284.98	993.29 ± 246.49
Approach learners	1257.67 ± 333.86	1118.94 ± 320.69	1007.47 ± 275.79	922.19 ± 216.10

Table 2
Mean and standard deviation of accuracy rate in four blocks for approach learners and avoidance learners.

	Block1 (M ± SD)	Block2 (M ± SD)	Block3 (M ± SD)	Block4 (M ± SD)
Avoidance learners	0.60 ± 0.05	0.80 ± 0.10	0.92 ± 0.06	0.93 ± 0.05
Approach learners	0.60 ± 0.06	0.81 ± 0.10	0.88 ± 0.07	0.92 ± 0.07

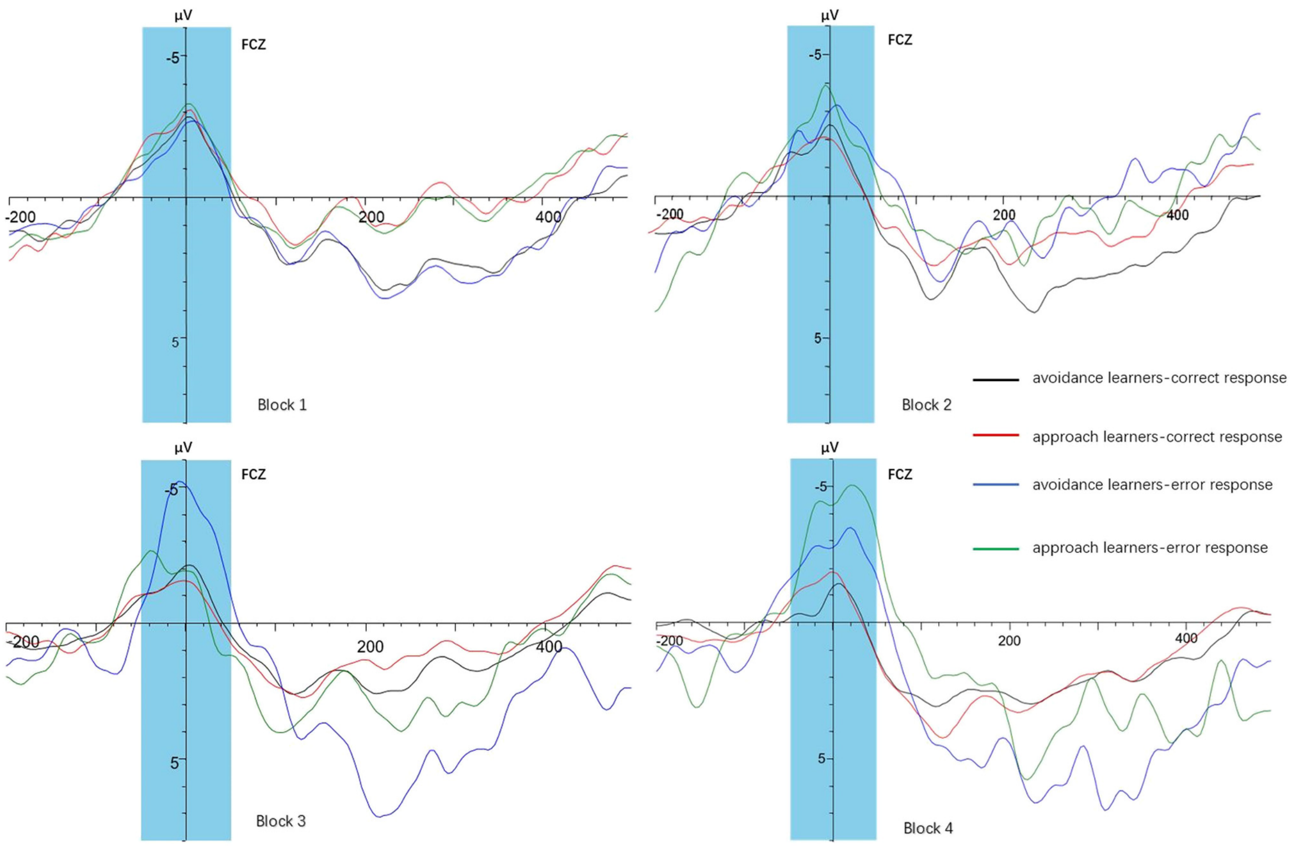

Fig. 2. Grand-averaged response-locked ERP waveforms at FCz in four blocks for approach learners and avoidance learners.

period. The results of the trend analysis indicated a significant linear effect ($F(1,63) = 34.80, p < 0.001, \eta_p^2 = 0.36$) and quadratic effect ($F(1,63) = 17.03, p < 0.001, \eta_p^2 = 0.21$). There was no significant cubic effect, $F(1,63) = 2.04$, $p = 0.16$. The linear effect, as expected, indicated that P300 mean amplitudes in block 4 is less positive in comparison to other three blocks (see Fig. 3). There was no main effect of learner type, $F(1,63) = 1.15$, $p = 0.29$. In addition, there was no significant two-way interaction effect for feedback valence and period, $F(3,189) = 1.69$, $p = 0.17$; feedback valence and learner type, $F(1,63) = 0.25$, $p = 0.62$; learner type and period, $F(3,189) = 0.10$, $p = 0.96$. A significant three-way interaction between learner type, feedback valence and period was not obtained, $F(3,189) = 0.11$, $p = 0.95$.

3.2.4. Relationship between ERN and FRN effect

The mean magnitude values of ERN effect (the difference between error and correct response) and FRN effect (the difference between negative and positive feedback) were separately computed for each of the four-block between approach learners and avoidance learners. The relationship between ERN and FRN effects averaged across the four blocks for approach learners and avoidance learners was investigated by Pearson correlation analyses. The correlation between the ERN and FRN effects in four blocks were negative for avoidance learners, $r(4) = -0.99$, $p = 0.01$, and approach learners, $r(4) = -0.35$, $p = 0.65$. Generally, the tendency of the ERN effect was gradually larger from block 1 to block 4 and the FRN effect was gradually smaller from block 1 to block 4. Namely, the tradeoff relationship was shown between the ERN and FRN effects during the learning process. The tradeoff relationship means a negative correlation relationship. The opposite-direction effects of learning periods on the ERN and the FRN were illustrated in Fig. 4.

4. Discussion

The dynamic learning performance of approach learners and avoidance learners was examined via an adapted probabilistic selection task. This experiment, including internal monitoring and external feedback, investigates that "which monitoring mechanism plays an important role in learning at different periods" and "which kind of feedback is the foundation of behavior modification".

The response time results indicated that approach learners reacted a little faster than avoidance learners in the whole task. Moreover, accuracy rates of two kinds of learners significantly improved after every block. The ERPs results demonstrated that external feedback only played a role in early learning period and negative feedback influenced individuals' behavior adaption in the early learning period. In the later period, internal monitoring took part in learning. In line with our hypotheses, the underlying ERN and FRN processes are functionally linked by a tradeoff relationship during reinforcement learning. For both negative feedback and positive feedback in four blocks, the P300 was

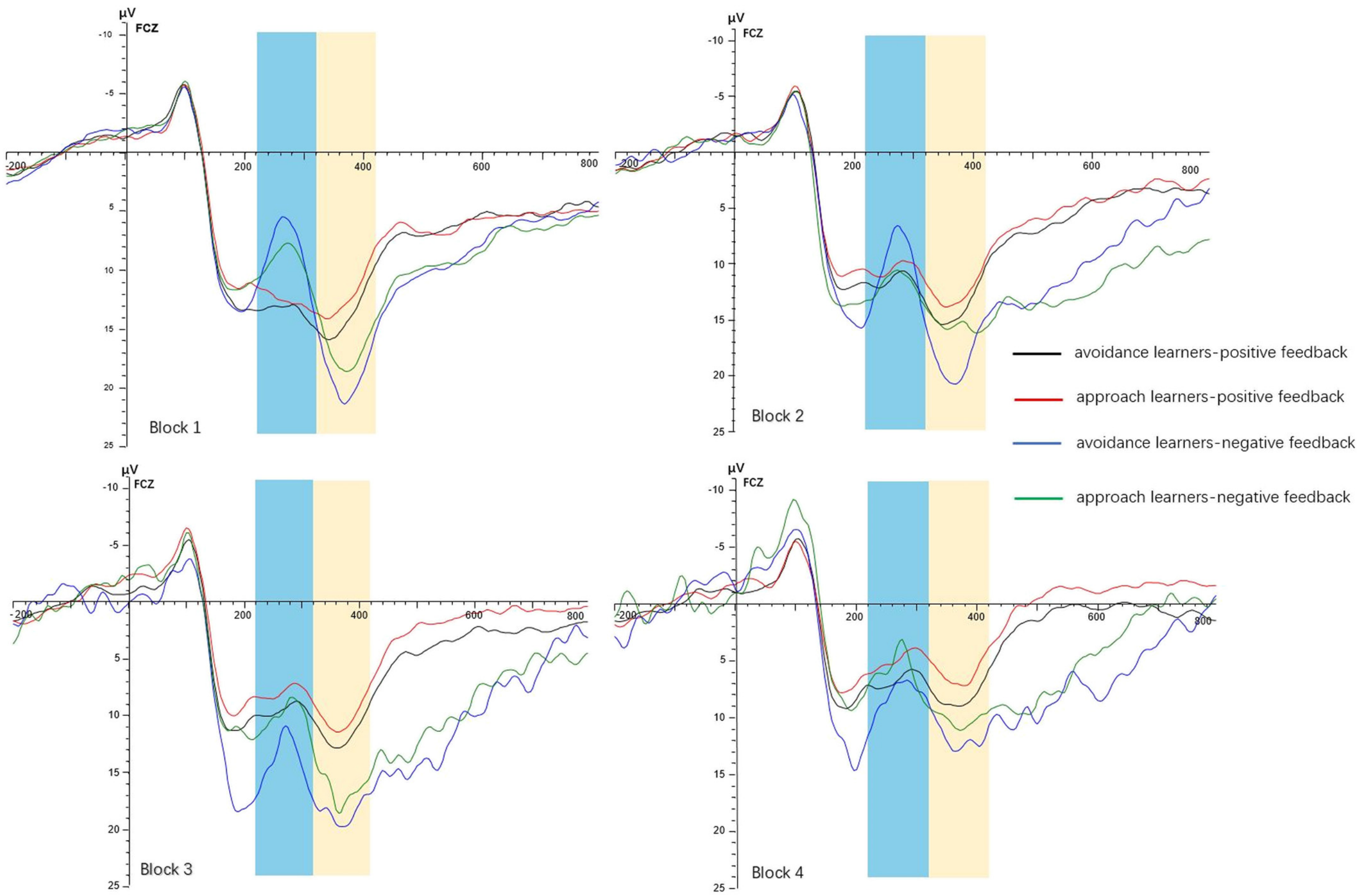

Fig. 3. Grand-averaged feedback-locked ERP waveforms at FCz in four blocks for approach learners and avoidance learners.

more positive for avoidance learners than approach learners. Meanwhile, it was more positive for negative feedback than positive feedback in four blocks and gradually positive from block1 to block4. Moreover, the amplitudes of FRN and P300 in block 4 were significantly smaller than the other three blocks.

In the following, we will discuss the individual differences, the tradeoff relationship between ERN and FRN, successful learning, and learning based on negative feedback in the early period.

4.1. Individual differences

One goal of the present study was to examine individual differences in reinforcement learning process. The response time showed that both approach learners and avoidance learners reacted gradually fast from block1 to block 4 (see Table 1) and gradually improved accuracy rate from block1 to block 4 (see Table 2). Compared to approach learners, avoidance learners speeded more time in choosing every block and had

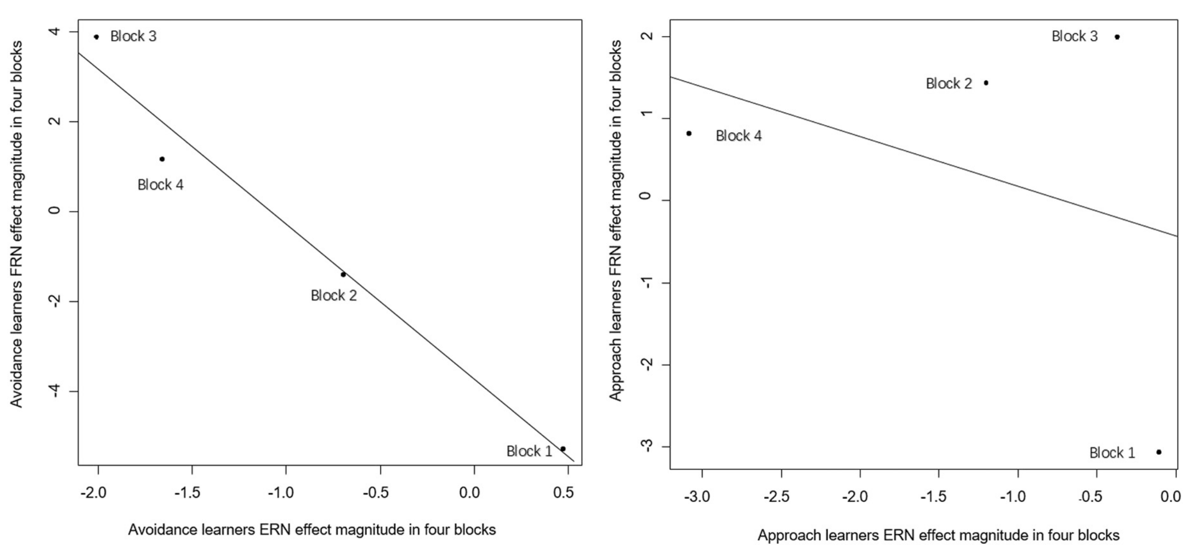

Fig. 4. Relationship between ERN and FRN effects for avoidance learners and approach learners in four blocks (ERN effect = error response amplitude − correct response amplitude; FRN effect = negative feedback amplitude − positive feedback amplitude). Larger values mean larger effects.

a higher accuracy rate in block 3 and block 4. And approach learners only showed a higher accuracy rate in block 4 (see Tables 1, 2).

For avoidance learners, the difference between correct response and error response amplitudes (ERN effect) was small in block 1 and block 2. This can be ascribed to that they did not recognize the rules and detect errors with the external feedback. However, avoidance learners gradually understood the rules and could detect errors instead of external feedback in block 3 and block 4, with large difference between correct response and error response amplitudes. As for approach learners, the ERN effect appeared until in block 4, indicating that they learned slower than avoidance learners (see scalp topographies (Fig. S1) in the Supplemental material).

On the other hand, the amplitude difference between positive feedback and negative feedback (FRN effect) was larger in block 1 than block 2 to block 4. This is consistent with the previous study (Shephard et al., 2014). The FRN effect was not significant but different between avoidance learners and approach learners in block 2. That demonstrated that avoidance learners needed more time to rely on external feedback, once avoidance learners learned the rules they could apply learned rules to improve behavior performance and recognize when making mistakes (i.e., block 3). However, approach learners needed less time to rely on external feedback, but this kind of learner could not immediately apply learned rules to improve behavior performance (see scalp topographies (Fig. S2) in the Supplemental material).

4.2. The tradeoff relationship between ERN and FRN

The central aim in the present study was to investigate "when internal monitoring and external feedback played an important role during the process of reinforcement learning respectively" and the relationship between them. The ERPs results showed that both approach learners and avoidance learners had larger FRN and smaller ERN effects in the early learning period, whereas both approach learners and avoidance learners had larger ERN and smaller FRN effects in the later learning period (see Fig. 4). This result pattern is corresponding with the increasing ERN amplitudes and decreasing FRN amplitudes from the first to the second half of experiment (Müller et al., 2005; Stahl, 2010). The foregoing results suggested that participants gradually recognized the rules, reduced errors and learned to detect errors, while they responded throughout the early learning period. Hence, the external feedback did not work any longer in the later learning period. Although the ERN effects appear slight earlier for avoidance learners than approach learners, external feedback leads to the gradually development of internal error detection (Wessel, 2012; Ullsperger et al., 2014; Hoffmann and Beste, 2015). These results support our hypothesis that the ERN plays an important role in the later learning period and the FRN plays a central part in the early learning period. Therefore, the tradeoff relationship was shown between the ERN and FRN effects during the reinforcement learning process. This inverse relationship between the amplitude of ERN and FRN reflects a shift between internal monitoring and external feedback (Hajcak et al., 2005b; Krigolson et al., 2009). The dynamic characteristic of tradeoff relationship between external feedback and internal monitoring has also been displayed in fMRI study (Gawlowska et al., 2018). The tradeoff relationship between the ERN and FRN effects indicates that ERN and FRN are quietly different two kinds of ERPs components (Cavanagh et al., 2009; Potts et al., 2011). This result is inconsistent with the Reinforcement Learning Theory (Holroyd and Coles, 2002). The Reinforcement Learning Theory postulates that ERN and FRN are reflections of the same generic level error processing system. Actually, ERN component appears in the early time window but reflects the later processing mechanism. FRN component appears in later time window but reflects the early processing mechanism.

4.3. Successful learning

The P300 mean amplitude of negative feedback was more positive than positive feedback in four blocks, and P300 mean amplitude gradually become negative from block1 to block4 for both negative feedback and positive feedback. The differences of P300 feedback valence were apparent in the early three learning periods. The context updating theory proposes that the P300 indexes the degree to which the individual's internal representation of the external context (Donchin and Coles, 1998). When a stimulus exceeding participants' expectations appears, the updating occurs (Potts, 2011). As negative feedback was a stimulus that exceeds participants' expectations, P300 was larger after negative feedback than after positive feedback in four blocks, indicating that participants wanted to perform well to finish the task. The fewer errors participants made in the later learning period, the less updating information participants needed.

The amplitudes of FRN and P300 were significantly smaller in block 4 than other three blocks. And the amplitudes of ERN in the later learning period were significantly larger than the early learning period. The aforementioned results suggest that participants gradually recognized the rules, reduced errors and learned to detect errors while they responded throughout the early learning period. Accuracy rates significantly improved after every block. Both the approach learners and avoidance learners had high accuracy rates in the last block. Taken the high accuracy rates, relatively large ERN amplitude, relatively small FRN and P300 amplitude together, it may be concluded that participants successfully mastered the learning rules in the later learning period (see scalp topographies (Fig. S3) in the Supplemental material).

4.4. Learning based on negative feedback in the early period

As mentioned above, FRN only worked in the early learning period. Participants needed more time to respond after negative feedback than after positive feedback. In order to improve performance in the task, participants had to adjust behavior after negative feedback than after positive feedback (Cavanagh et al., 2009; Danielmeier and Ullsperger, 2011; Fu et al., 2019). In addition, FRN amplitude was larger after negative feedback than after positive feedback. At the same time, the P300 was more positive after negative feedback than after positive feedback in the early period. The P300 is a necessary signal to adapt behavior for the formation of good performance (Hamamé et al., 2011). Therefore negative feedback can be seen as convincing evidence for feedback-based learning. This finding is in line with several studies (Philiastides et al., 2010; Asaad and Eskandar, 2011; Schiffler et al., 2016; Kastner et al., 2017).

Acknowledgements

This study was supported by Shandong Provincial Natural Science Foundation (ZR2015PC012), China and Graduate Dissertation Research Innovation Fund Project 2019 of Qufu Normal University, China.

Declaration of competing interest

The authors declare that no competing interests exist.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpsycho.2020.01.004.

References

Aberg, K.C., Doell, K.C., Schwartz, S., 2016. Linking individual learning styles to approach-avoidance motivational traits and computational aspects of reinforcement learning. PLoS One 11 (11), 1–16. https://doi.org/10.1371/journal.pone.0166675.

Asaad, W.F., Eskandar, E.N., 2011. Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. J. Neurosci. 31 (49), 17772–17787. https://doi.org/10.1523/JNEUROSCI.3793-11.2011.

Baker, T.E., Holroyd, C.B., 2011. Dissociated roles of the anterior or cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. Biol. Psychol. 87 (1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010.

Barch, D.M., Carter, C.S., Gold, J.M., Johnson, S.L., Kring, A.M., MacDonald, A.W., Pizzagalli, D.A., Ragland, J.D., Silverstein, S.M., Strauss, M.E., 2017. Explicit and implicit reinforcement learning across the psychosis spectrum. J. Abnorm. Psychol. 126 (5), 694–711. https://doi.org/10.1037/abn0000259.

Bartra, O., McGuire, J., Kable, J., 2013. The valuation system: a coordinate -based meta-analysis of bold fMRI experiments examining the neural correlates of subjective value. Neuroimage 76 (1), 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063.

Bultena, S., Danielmeier, C., Bekkering, H., Lemhöfer, K., 2017. Electrophysiological correlates of error monitoring and feedback processing in second language learning. Front. Hum. Neurosci. 11 (29), 1–18. https://doi.org/10.3389/fnhum.2017.00029.

Cavanagh, J.F., Cohen, M.X., Allen, J.J.B., 2009. Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. J. Neurosci. 29 (1), 98–105. https://doi.org/10.1016/j.ijpsycho.2018.02.011.

Chase, H.W., Swainson, R., Durham, L., Benham, L., Cools, R., 2010. Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. J. Cogn. Neurosci. 23 (4), 936–946. https://doi.org/10.1162/jocn.2010.21456.

Collins, A.G.E., Frank, M.J., 2018. Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory. Proc. Natl. Acad. Sci. U. S. A. 115 (10), 2502–2507. https://doi.org/10.1073/pnas.1720963115.

Crowley, M.J., Wu, J., Bailey, C.A., Mayes, L.C., 2009. Bringing in the negative reinforcements: the avoidance feedback-related negativity. Neuroreport 20 (17), 1513–1517. https://doi.org/10.1097/WNR.0b013e32832ff2f5.

Dambacher, M., Hübner, R., 2015. Time pressure affects the efficiency of perceptual processing in decisions under conflict. Psychol. Res. 79 (1), 83–94. https://doi.org/10.1007/s00426-014-0542-z.

Danielmeier, C., Ullsperger, M., 2011. Post-error adjustments. Front. Psychol. 2, 233. https://doi.org/10.3389/fpsyg.2011.00233.

Decker, J.H., Otto, A.R., Daw, N.D., Hartley, C.A., 2016. From creatures of habit to goal-directed learners: tracking the developmental emergence of model-based reinforcement learning. Psychol. Sci. 27 (6), 848–858. https://doi.org/10.1177/0956797616639301.

Donchin, E., Coles, M., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11 (3), 357–374. https://doi.org/10.1017/S0140525X00058027.

Ehlis, A.C., Deppermann, S., Fallgatter, A.J., 2018. Performance monitoring and post-error adjustments in adults with attention-deficit/hyperactivity disorder: an EEG analysis. J. Psychiatry Neurosci. 43 (6), 396–406 (http://doi.org/jpn.170118).

Eppinger, B., Kray, J., Mock, B., Mecklinger, A., 2008. Better or worse than expected? Aging, learning, and ERN. Neuropsychologia 46, 521–539. https://doi.org/10.1016/j.neuropsycholog ia.2007.09.001.

Ernst, B., Steinhauser, M., 2012. Feedback-related brain activity predicts learning from feedback in multiple-choice testing. Cogn. Affect. Behav. Neurosci. 12 (2), 323–336. https://doi.org/10.3758/s13415-012-0087-9.

Falkenstein, M., Hohnsbein, J., Hoormann, J., Blanke, L., 1991. Effects of cross modal divided attention on late ERP components. II. Error processing in choice reaction tasks. Electroencephalogr. Clin. Neurophysiol. 78 (6), 447–455. https://doi.org/10.1016/0013-4694 (91)90062-9.

Fischer, A.G., Ullsperger, M., 2013. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. Neuron 79 (6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006.

Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol. Rev. 113 (2), 300–326. https://doi.org/10.1037/0033-295X.113.2.300.

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306 (5703), 1940–1943. https://doi.org/10.1126/science.1102941.

Frank, M.J., Woroch, B.S., Curran, T., 2005. Error-related negativity predicts reinforcement learning and conflict biases. Neuron 47, 495–501. https://doi.org/10.1016/j.neuron.2005.06.02 0.

Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., Hutchison, K.E., 2007. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc. Natl. Acad. Sci. U. S. A. 104 (41), 16311–16316. https://doi.org/10.1073/pnas.0706111104.

Fu, Z.Z., Wu, D.A.J., Ross, I., Chung, J.M., Mamelak, A.N., Adolphs, R., Rutishauser, U., 2019. Single-neuron correlates of error monitoring and post-error adjustments in human medial frontal cortex. Neuron 101 (1), 165–177. https://doi.org/10.1016/j.neuron.2018.11.016.

Gawlowska, M., Domagalik, A., Beldzik, E., Marek, T., Mojsa-Kaja, J., 2018. Dynamics of error-related activity in deterministic learning - an EEG and fMRI study. Sci. Rep. 8, e14617. https://doi.org/10.1038/s41598-018-32995-x.

Gehring, W.J., Liu, Y., Orr, J.M., Carp, J., 2011. The error-related negativity (ERN/Ne). In: Luck, S.J., Kappenman, E.S. (Eds.), The Oxford Handbook of Event-related Potential Components. University Press, Oxford, pp. 231–291.

Gheza, D., Paul, K., Pourtois, G., 2017. Dissociable effects of reward and expectancy during evaluative feedback processing revealed by topographic ERP mapping analysis. Int. J. Psychophysiol. 132, 213–225. https://doi.org/10.1016/j.ijp sycho.2017.11.013.

Hajcak, G., Holroyd, C.B., Moser, J.S., Simons, R.F., 2005a. Brain potentials associated with expected and unexpected good and bad outcomes. Psychophysiology 42 (2), 161–170. https://doi.org/10.1111/j.1469-8986.2005.00278.x.

Hajcak, G., Moser, J.S., Yeung, N., Simons, R.F., 2005b. On the ERN and the significance of errors. Psychophysiology 42 (2), 151–160. https://doi.org/10.1111/j.1469-8986.2005.00270.x.

Hamamé, C., Cosmelli, D., Henriquez, R., Aboitiz, F., 2011. Neural mechanisms of human perceptual learning: electrophysiological evidence for a two-stage process. PLoS One 6 (4), e19221. https://doi.org/10.1371/journal.pone.0019221.

Hämmerer, D., Li, S.C., Müller, V., Lindenberger, U., 2010. Life span differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. J. Cogn. Neurosci. 23 (3), 579–592. https://doi.org/10.1162/jocn.2010.21475.

Hoffmann, S., Beste, C., 2015. A perspective on neural and cognitive mechanisms of error commission. Front. Behav. Neurosci. 9, eUNSP50. https://doi.org/10.3389/fnbeh.2005.00050.

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Res. 109 (4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679.

Ichikawa, N., Siegle, G.J., Dombrovski, A., Ohira, H., 2010. Subjective and model-estimated reward prediction: association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task. Int. J. Psychophysiol. 78 (3), 273–283. https://doi.org/10.1016/j.ijpsycho.2010.09.001.

Jasper, H.H., 1958. Report of the committee on methods of clinical examination in electroencephalography: 1957. Electroencephalogr. Clin. Neurophysiol. 10 (2), 370–375. https://doi.org/10.1016/0013-4694(58)90053-1.

Jessup, R.K., Busemeyer, J.R., Brown, J.W., 2010. Error effects in anterior cingulate cortex reverse when error likelihood is high. J. Neurosci. 30, 3467–3472. https://doi.org/10.1523/JNEUROSCI.4130-09.2010.

Kastner, L., Kube, J., Villringer, A., Neumann, J., 2017. Cardiac concomitants of feedback and prediction error processing in reinforcement learning. Front. Neurosci. 11 (598), 1–19. https://doi.org/10.3389/fnins.2017.00598.

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 (9), 1833–1840. https://doi.org/10.1162/jocn.2009.21128.

Lou, B., Hsu, W.Y., Sajda, P., 2015. Perceptual salience and reward both influence feedback related neural activity arising from choice. J. Neurosci. 35 (38), 13064–13075. https://doi.org/10.1523/JNEUROSCI.1601-15.2015.

Luft, C.D.B., 2014. Learning from feedback: the neural mechanisms of feedback processing facilitating better performance. Behav. Brain Res. 261, 356–368. https://doi.org/10.1016/j.bbr.2013.12.043.

Luque, D., Moris, J., Rushby, J.A., Le Pelley, M.E., 2015. Goal-directed EEG activity evoked by discriminative stimuli in reinforcement learning. Psychophysiology 52 (2), 238–248. https://doi.org/10.1111/psyp.12302.

Martin, R.S., 2012. Event-related potential studies of outcome processing and feedback-guided learning. Front. Hum. Neurosci. 6. https://doi.org/10.3389/fnhum.2012.00304.

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. Cogn. Neurosci. 9 (6), 788–799. https://doi.org/10.1162/jocn.1997.9.6.788.

Müller, S., Möller, J., Rodriguez-Fornells, A., Münte, T., 2005. Brain potentials related to self-generated and external information used for performance monitoring. Clin. Neurophysiol. 116 (1), 63–74. https://doi.org/10.1016/j.clinph.2004.07.009.

Muller-Gass, A., Duncan, M., Campbell, K., 2017. Brain states predict individual differences in perceptual learning. Personal. Individ. Differ. 118, 29–38. https://doi.org/10.1016/j.paid.2017.03.066.

Philiastides, M.G., Biele, G., Vavatzanidis, N., Kazzer, P., Heekeren, H.R., 2010. Temporal dynamics of prediction error processing during reward-based decision making. Neuroimage 53 (1), 221–232. https://doi.org/10.1016/j.neuroimage.2010.05.052.

Potts, G.F., 2011. Impact of reward and punishment motivation on behavior monitoring as indexed by the error-related negativity. Int. J. Psychophysiol. 81 (3), 324–331. https://doi.org/10.1016/j.ijpsycho.2011.07.020.

Potts, G.F., Martin, L.E., Kamp, S.M., Donchin, E., 2011. Neural response to action and reward prediction errors: comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations. Psychophysiology 48 (2), 218–228. https://doi.org/10.1111/j.1469-8986.2010.01049.x.

Sambrook, T.D., Goslin, J., 2015. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol. Bull. 141 (1), 213–235. https://doi.org/10.1037/bul0000006.

Schiffler, B.C., Almeida, R., Granqvist, M., Bengtsson, S.L., 2016. Memory-reliant post-error slowing is associated with successful learning and fronto-occipital activity. J. Cogn. Neurosci. 28 (10), 1539–1552. https://doi.org/10.1162/jocn_a_00987.

Schmid, P.C., Hackel, L.M., Jasperse, L., Amodio, D.M., 2018. Frontal cortical effects on feedback processing and reinforcement learning: relation of EEG asymmetry with the feedback-related negativity and behavior. Psychophysiology 55 (1), e12911. https://doi.org/10.1111/psyp.12911.

Shephard, E., Jackson, G.M., Groom, M.J., 2014. Learning and altering behaviours by reinforcement: neurocognitive differences between children and adults. Dev. Cogn. Neurosci. 7, 94–105. https://doi.org/10.1016/j.dcn.2013.12.001.

Stahl, J., 2010. Error detection and the use of internal and external error indicators: an investigation of the first-indicator hypothesis. Int. J. Psychophysiol. 77 (1), 43–52. https://doi.org/10.1016/j.ijpsycho.2010.04.005.

Uehara, S., Mawase, F., Celnik, P., 2017. Learning similar actions by reinforcement or sensory-prediction errors rely on distinct physiological mechanisms. Cereb. Cortex 28 (10), 2478–3490. https://doi.org/10.1093/cercor/bhx214.

Ullsperger, M., Fischer, A.G., Nigbur, R., Endrass, T., 2014. Neural mechanisms and temporal dynamics of performance monitoring. Trends Cogn. Sci. 18, 259–267. https://doi.org/10.1016/j.tics.2014.02.009.

Van den Bos, W., Gürǒglu, B., van den Bulk, B.G., Rombouts, S.A.R.B., Crone, E.A., 2009. Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. Front. Hum. Neurosci. 3, 52. https://doi.org/10.3389/neuro.09.052.2009.

Wang, L., Sun, H.Y, Li, L., Meng, L., 2018. Hey, what is your choice? Uncertainty and inconsistency enhance subjective anticipation of upcoming information in a social context. Exp. Brain Res. 236 (10), 2797–2810. https://doi.org/10.1007/s00221-018-5336-x.

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. 36, 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008.

Wessel, J.R., 2012. Error awareness and the error-related negativity: evaluating the first decade of evidence. Front. Hum. Neurosci. 6, e88. https://doi.org/10.3389/fnhum.2012.00088.

West, R., Bailey, K., Anderson, S., 2018. Transient and sustained ERP activity related to feedback processing in the probabilistic selection task. Int. J. Psychophysiol. 126, 1–12. https://doi.org/10.1016/j.ijpsycho.2018.02.011.

Yin, H., Wang, Y., Zhang, X.K, Li, P., 2018. Feedback delay impaired reinforcement learning: Principal components analysis of reward positivity. Neurosci. Lett. 685, 179–184. https://doi.org/10.1016/j.neulet.2018.08.039.

Zubarev, I., Parkkonen, L., 2018. Evidence for a general performance-monitoring system in the human brain. Hum. Brain Mapp. 1–12. https://doi.org/10.1002/hbm.24273.