Similarities and differences between the processing of immediate and delayed performance feedback

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Benjamin Maurice Weismüller

aus München

Düsseldorf, September 2018

aus dem Institut für Experimentelle Psychologie

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Christian Bellebaum

2. Prof. Dr. Tobias Kalenscher

Tag der mündlichen Prüfung: 11.12.2018

Contents

List of Abbreviations

ACC	anterior cingulate cortex
BDI	Beck depression inventory
BOLD	blood-oxygen-level-dependent
COMT	catechol-O-methyltransferase
CR	conditioned response
CRN	correct-response negativity
CS	conditioned stimulus
DA	dopamine
EEG	electroencephalography
ERN	error-related negativity
ERP	event-related potential
fERN	feedback error-related negativity
fMRI	functional magnetic resonance imaging
FRN	feedback-related negativity
GABA	gamma-amino butyric acid
GLU	glutamate
GP	globus pallidus
HC	hippocampus
L-DOPA	L-3,4-dihydroxyphenylalanin
MMST	mini mental status test
PD	Parkinson's disease
PET	positron-emission-tomography
PFC	prefrontal cortex
RewP	reward positivity
RL-theory	reinforcement learning theory
SN	substantia nigra
UPDRS	unified Parkinson's disease rating scale
UR	unconditioned response
US	unconditioned stimulus

Summary

Learning from feedback enables behavioural adaptation. Actions with beneficial outcomes are performed more frequently in the future, whereas actions leading to non-beneficial or aversive outcomes are avoided. Animal research on feedback learning showed that midbrain dopamine (DA) neuron firing rates burst in response to reward-predicting stimuli, while their firing rate dips when the expected reward is omitted, thereby coding a prediction error signal. In humans, DA neurons project from the midbrain to the striatum and anterior cingulate cortex (ACC), which have both been shown to play a role in feedback processing. However, a previous imaging study reported that the hippocampus (HC) becomes active when feedback is processed that follows a related action after a temporal delay (7 s). This finding associates the HC with delayed feedback processing, while it is usually ascribed to declarative memory. In line with this, patients suffering from Parkinson's disease (PD), which is characterised by a substantial depletion of striatal DA levels, were impaired in learning from immediate (1 s), but not delayed (7 s) feedback, possibly making use of their intact HC. Furthermore, an event-related potential (ERP) study found that the activity in the striatum/ACC decreases with increasing temporal delay between an action and the relative feedback.

It is not yet clear whether the striatum/ACC and the HC are parts of two competing neuronal systems and feedback is processed in either one or the other depending on its timing, or whether these structures interact in a cooperative manner. In a series of three studies, the present dissertation examined the modulatory effects of temporal delay on the neuro-cognitive underpinning of feedback processing by comparing effects that are well-established in learning from immediate feedback between immediate (1 s) and delayed (7 s) feedback learning.

Unlike healthy control participants, PD patients have previously been found to learn better from negative than positive immediate feedback, which was attributed to their lack of striatal DA. Study 1 compared this learning bias between two groups of patients (and their controls) that learned from immediate and delayed feedback, respectively. The feedback timing did, however, not affect the patients' tendency to

learn better from negative feedback, which indicates that striatal DA depletion affected both learning from immediate and delayed feedback. This in turn suggests a striatal contribution to delayed feedback processing.

The feedback-related negativity (FRN) is an ERP component that supposedly reflects DA-related activity in the striatum/ACC. Previous research showed that the FRN amplitude is increased for unexpected compared to expected immediate feedback mirroring a prediction error. Study 2 compared this FRN expectancy effect between immediate and delayed feedback finding larger FRN amplitudes for unexpected compared to expected feedback during both immediate and delayed feedback learning. In line with Study 1, this finding suggests striatal/ACC involvement in delayed feedback processing.

Feedback agency has previously been shown to modulate activity in the striatum/ACC as well, with decreased activity for learning from feedback for observed versus self-performed actions. Study 3 investigated the combined influence of feedback agency and delay on the FRN as well as beta and theta band oscillations. The FRN was affected by feedback agency and delay, with a combined influence only for feedback for self-generated actions. Beta and theta power were affected by feedback timing, while agency only modulated the power in the theta band. These results indicate that the mechanisms underlying the FRN may differ from those underlying beta and theta oscillations. Beta and theta oscillations have been suggested to reflect valence-specific communication signals in the need for memory consolidation or cognitive control, respectively.

Taken together, the results of all three studies indicate a contribution of the striatum/ACC to delayed feedback processing, which may suggest a cooperation of declarative and non-declarative systems during delayed feedback learning.

Zusammenfassung

Lernen aus Feedback ermöglicht Verhaltensanpassung. Handlungen mit vorteilhaften Folgen werden in der Zukunft öfter ausgeübt, während Handlungen, die zu unvorteilhaften oder aversiven Folgen führen, vermieden werden. Tierforschung zu Feedbacklernen hat gezeigt, dass die Feuerrate von Dopamin (DA)-Neuronen im Mittelhirn mit starken Ausschlägen auf belohnungsvorhersagende Stimuli reagieren, während die Feuerrate abfällt, wenn eine erwartete Belohnung ausgelassen wird. Somit kodiert die DA-Aktivität einen Vorhersagefehler. Bei Menschen projizieren DA-Neurone vom Mittelhirn in das Striatum und den anterior-zingulären Cortex (engl. *anterior-cingulate cortex,* ACC), welche beide eine Rolle bei Feedbackverarbeitung spielen. Allerdings hat eine vorhergehende bildgebende Studie gezeigt, dass der Hippokampus (engl. *hippocampus,* HC) aktiv wird, wenn ein Feedback verarbeitet wird, das zeitlich verzögert (nach 7 s) auf eine Handlung folgt. Dieser Fund assoziiert den HC mit der Verarbeitung von verzögertem Feedback, obwohl er normalerweise deklarativem Gedächtnis zugeschrieben wird. Gleichermaßen waren Patienten, die unter der Parkinson Krankheit (engl. *Parkinson's disease,* PD) leiden, die durch eine starke Verarmung striatalen DAs charakterisiert ist, beim Lernen aus unverzögertem Feedback beeinträchtigt, aber nicht beim Lernen aus verzögertem Feedback, möglicherweise, weil sie ihren intakten HC benutzten. Des Weiteren hat eine Studie zu Ereignis-korrelierten Potentialen (engl. *event-related potential,* ERP) herausgefunden, dass sich die Aktivität im Striatum/ACC mit zunehmendem zeitlichen Abstand zwischen einer Handlung und dem entsprechenden Feedback verringert.

Es ist noch nicht klar, ob das Striatum bzw. der ACC und der HC Teile zweier konkurrierender Systeme sind und Feedback abhängig vom Erscheinungszeitpunkt entweder in dem einem oder dem anderen verarbeitet wird oder, ob diese Strukturen in einer kooperativen Weise miteinander interagieren. In einer Reihe aus drei Studien untersuchte die vorliegende Dissertation Einflüsse von Feedbackverzögerung auf die neuro-kognitiven Grundlagen der Feedbackverarbeitung, indem Effekte, die für

unverzögertes Feedbacklernen etabliert sind, zwischen verzögertem (1 s) und unverzögertem (7 s) Feedbacklernen verglichen wurden.

Zuvor wurde gezeigt, dass PD-Patienten anders als gesunde Kontrollprobanden besser aus negativem als aus positivem Feedback lernen, was auf ihren striatalen DA-Mangel zurückgeführt wurde. Die Studie 1 verglich diesen Lern-Bias zwischen zwei Gruppen von Patienten (und deren Kontrollen), die aus unverzögertem beziehungsweise verzögertem Feedback lernten. Der Erscheinungszeitpunkt des Feedbacks hatte allerdings keinen Einfluss auf die Tendenz der Patienten, besser aus negativem Feedback zu lernen, was darauf hindeutet, dass striatale DA-Verarmung sowohl einen Einfluss auf das Lernen aus unverzögertem, als auch aus verzögertem Feedback hatte. Dies wiederum suggeriert einen striatalen Beitrag zur Verarbeitung von verzögertem Feedback.

Die Feedback-bezogene Negativität (engl. *feedback-related negativity,* FRN) ist eine ERP-Komponente, die vermutlich DA-bezogene Aktivität im Striatum/ACC widerspiegelt. Vorhergehende Forschung zeigte, dass die FRN Amplitude für unerwartetes Feedback verglichen mit erwartetem Feedback erhöht ist, was einen Erwartungsfehler widerspiegelt. Die Studie 2 verglich diesen FRN-Erwartungseffekt zwischen unverzögertem und verzögertem Feedback, wobei größere FRN Amplituden für unerwartetes als für erwartetes Feedback sowohl beim Lernen aus unverzögertem als auch beim Lernen aus verzögertem Feedback gefunden wurde. Wie auch Studie 1 suggeriert dieser Fund einen Beitrag des Striatums/ACCs zu die Verarbeitung verzögerten Feedbacks.

Es wurde zuvor gezeigt, dass der Selbstbezug eines Feedbacks ebenfalls die Aktivität im striatum/ACC moduliert, wobei die Aktivität beim Lernen aus Feedback für beobachtete gegenüber selbst-generierten Handlungen verringert war. Die Studie 3 untersuchte den kombinierten Einfluss des Feedbackselbstbezugs und der Feedbackverzögerung auf die FRN, sowie auf Beta- und Thetaoszillationen. Die FRN wurde durch den Selbstbezug des Feedbacks und dessen Verzögerung beeinflusst, wobei ein kombinierter Einfluss nur für Feedback für selbst-generierte Handlungen gefunden wurde. Die Beta- und Thetapower wurden durch den Feedback-Erscheinungszeitpunkt beeinflusst, während nur der Feedbackselbstbezug die Power

im Thetaband modulierte. Diese Ergebnisse deuten darauf hin, dass sich die Mechanismen, die der FRN zugrunde liegen, von denen unterscheiden, die Beta- und Thetaoszillationen unterliegen. Zu Beta- und Thetaoszillationen wurde vorgeschlagen, dass sie valenzspezifische Kommunikationssigle widerspiegeln, wenn Gedächtniskonsolidierung beziehungsweise kognitive Kontrolle notwendig ist.

Zusammengefasst deuten alle drei Studien auf einen Beitrag des Striatums/ACCs zur Verarbeitung verzögerten Feedbacks hin, was eine Kooperation des deklarativen und des non-deklarativen Systems beim Lernen aus verzögertem Feedbacks suggeriert.

1 Introduction

As resources are limited, living organisms need to act in a way that maximises desired outcomes with the least possible effort. Thus, organisms have to learn how to optimise their behaviour, which can be achieved by interactions with the environment and adapting behaviour accordingly. This implies that living organisms have to learn from their actions' outcomes. Regardless of whether an outcome is favourable or unfavourable, living organisms are able to use it as feedback and learn from it. Although many interactions with the environment cause immediate consequences, in the modern complex environment of human beings some actions may also result in temporally delayed consequences. Still, humans are able to learn from outcomes that follow their actions after several seconds and adapt their behaviour accordingly. However, recent research detected differences in the neuronal activation patterns when learning had to be achieved from feedback provided after a temporal delay (after 7 s) compared to feedback provided immediately (after 1 s; Foerde, Race, Verfaellie, & Shohamy, 2013; Foerde & Shohamy, 2011a).

The present dissertation further explores effects of a temporal delay between an action and the consequential outcome on neuronal mechanisms underlying feedback learning in humans. Behavioural performance and psychophysiological correlates were assessed during feedback learning tasks that included conditions with immediate and delayed feedback. The first section of the introduction section explains the general concept of feedback learning and summarises the current state of research on how a temporal gap between an action and the resulting outcome may influence learning. The next section provides information on how previous research investigated the neuronal underpinnings of feedback learning. Special attention is given to behavioural investigations of patients suffering from Parkinson's Disease (PD) and to the acquisition and interpretation of electrophysiological data during feedback learning.

1.1 Theoretical definition of feedback learning and underlying mechanisms

The ability to learn from feedback comprises the ability to associate a consequence with a previous action or event. Early experimental investigations of learning revealed that a behaviour can be linked to a stimulus through associative learning (Pavlov, 1927; Pavlov & Gantt, 1928). It was found that dogs learned to associate stimuli with each other that had been repeatedly presented together and to react to both in a comparable manner. In detail, an unconditioned stimulus (US) such as food resulted in an unconditioned response (UR), in this case, salivation. If the US was repeatedly presented together with a neutral stimulus such as acoustic signal, the neutral stimulus became a conditioned stimulus (CS). The CS that did previously not cause a reaction, resulted in a conditioned response (CR) after the dogs learned to associate the CS (the acoustic signal) to the US (the food; Pavlov, 1927; Pavlov & Gantt, 1928). In this first classical conditioning paradigm, repeated presentation of two paired unrelated stimuli led to the association of these stimuli.

Another important associative learning mechanism involves the consequences of a behaviour. In operant conditioning actions leading to beneficial outcomes are more frequently repeated in the future, while actions leading to unsatisfying results will be performed less frequently (Skinner, 1938; Thorndike, 1927). In detail, living organisms experience their actions' consequences when interacting with the environment. The consequential information can be used by the organism to learn about the environment, to apply the gained knowledge in future interactions with it, and eventually to adapt its behaviour with respect to a desired goal. In his experimental investigations, Thorndike (1898) provided empirical evidence for this effect. He observed cats that tried to escape from a puzzle box to receive a reward. To escape this box the cats needed to show a specific behaviour or sequence of actions. Thorndike found that the entrapped cats first tried out several behaviours. If a behaviour resulted in a satisfactory reward, it was remembered and shown sooner in following trials, while other behaviours were discarded. Thorndike (1927) termed this process of behavioural adaption based on learning from past consequences *law of effect*.

Skinner (1938) extended Thorndike's findings inventing the *operant conditioning chamber* in which animals, e.g., rats or pigeons were thought to show a certain simple

behaviour that led to a reward or, in some experimental designs, to a punishment. In contrast to Thorndike (1927), Skinner (1938) primarily measured the behaviour's frequency and investigated changes in behavioural response rates depending on its outcome. Skinner (1938) found that behaviours that lead to rewarding outcomes were conducted more frequently, while unrewarded behaviours decreased in frequency. Necessary prerequisites of this reinforcement based learning are a desired outcome or a goal, exploration, and a well-defined rule that reliably relates the reward to the previous action.

Trying out different actions and receiving feedback enables an organism to map outcomes to previous actions and eventually to generate an internal model about the environment. This internal model is then used to predict the outcomes of future actions, i.e., to generate an *internal forward model* (Sutton & Barto, 1998). Any consequence of an action is evaluated based on this internal forward model by comparing the actually received to the predicted outcome. New behaviours are adopted depending on this evaluation so that actions leading to a predicted outcome will strengthen the prediction and be executed more frequently. Conversely, actions that do not lead to a predicted outcome will weaken the prediction and be executed less likely. Given that organisms always strive for beneficial outcomes, both types of behavioural adaptations will maximise the desired outcome in future interactions with the environment (Sutton & Barto, 1998).

If feedback learning is defined by behavioural adaptation based on the evaluation of an action's outcome, the internal model of the environment is of particular importance, as it generates a specific prediction about the outcome. The more accurate this prediction is, the more optimised is the behaviour (Sutton & Barto, 1998). Consequently, feedback learning only takes place when a behaviour is not yet optimised and behaviour and predictions still have to be adapted. In this case, the critical point that facilitates learning is when a prediction about an action's outcome is violated, i.e., when a *prediction error* occurs. When a prediction error is detected, the previous prediction has to be updated and behaviour is adapted (Rescorla & Wagner, 1972).

In summary, feedback learning can be understood as updating of an internal forward model about action-outcome contingencies based on the deviation of the expected from the actual outcome. To update existing predictions, an actions' actual outcome needs to be monitored in order to detect deviations from the prediction (Ullsperger, Danielmeier, & Jocham, 2014).

1.2 Neural mechanisms of feedback processing and learning

Thorndike (1898) measured the time a cat needed to escape the puzzle box in each trial. The resulting learning curves provided insight into the cats' learning progress as they adapted their behaviour and consequentially needed less time for their escape with each trial of the experiment. Aside from learning curves, other behavioural measures such as post-error slowing or post-error accuracy improvement indicate error-detection and imply that feedback learning is underlying the observed behavioural adjustments (e.g., Danielmeier & Ullsperger, 2011; Forster & Cho, 2014; King, Korb, von Cramon, & Ullsperger, 2010; Steinhauser, Maier, & Steinhauser, 2017). However, such behavioural measures cannot directly inform about the neuronal underpinnings of post-error adjustments (Ullsperger et al., 2014). In search of the neuronal mechanisms underlying feedback learning, invasive and neurochemistry research suggested that dopamine (DA) plays a crucial role in processing and learning from reward (Wise, 2008).

Already in the late 1970s, Wise, Spindler, deWit, and Gerberg (1978) introduced the idea that DA release in the brain would have a hedonistic effect on animals. Although they rejected this hypothesis (Wise et al., 1978), reward processing and learning have been associated with DA in subsequent research. A more detailed view on the DA system during learning was provided by Schultz (1997) and Schultz, Dayan, and Montague (1997), who investigated DA neuron spiking rates in the monkey substantia nigra (SN) and ventral tegmental area (VTA). In these studies, monkeys received a reward after a reward-predicting cue. When the monkeys observed the unconditioned cue and the following reward during the first few trials, DA neuron firing remained at a baseline level when the cue was shown, but increased when the reward was provided. This is in line with the notion of a hedonistic effect of DA during

reward consumption (Wise et al., 1978). However, after some experience with the task, the monkeys learned the association between the cue and the following reward so that the cue became a CS predicting the reward. In later trials, firing rates of DA neurons already increased after cue presentation and dropped to a baseline level when the reward itself was presented. Accordingly, the DA neuron firing burst propagated back in time from the reward to the predictive cue (Schultz, 1997, 1998; Schultz et al., 1997). This finding contradicted the hedonism theory, which only related DA to the satisfactory effect of reward consumption (Wise et al., 1978), instead associating DA with the expectation of reward (for review, see Schultz, 2002).

Importantly, when a reward was omitted although the predictive cue had been presented, DA neuron firing rates dropped below baseline at the time when the reward would have been presented. As the outcome in these trials was worse than expected, this dip in DA neuron firing rates coded a negative reward prediction error. Likewise, the DA neuron firing burst after reward presentation during the first trials reflected a positive prediction error, as no reward was expected before learning the predictive quality of the CS (Schultz, 1997, 1998; Schultz et al., 1997; for review, see Schultz, 2002).

In summary, midbrain DA neurons in the monkey brain seem to code a prediction error signal if an outcome deviates from expectation. More recent work showed that this also holds true in the human brain. Single nigro-striatal DA neuron firing patterns recorded during deep-brain stimulator implantation in PD patients coded prediction error signals for unexpected rewards (Zaghloul et al., 2009). Additional animal studies directly linked DA prediction error signals to reward learning, showing that DA neuron activity after reward presentation decreased as soon as learning was established because learning necessarily improves the correctness of predictions (Schultz, Apicella, & Ljungberg, 1993).

Another line of research suggested a critical influence of the habenula on the DA neurons' predictions and thus error signals (e.g., Bromberg-Martin, Matsumoto, Hong, & Hikosaka, 2010; Hong, Jhou, Smith, Saleem, & Hikosaka, 2011; Matsumoto & Hikosaka, 2007, 2009). Matsumoto and Hikosaka (2007, 2009) found that neurons in the rat lateral habenula become more active when negative prediction errors occur and

less active during positive prediction errors. Electrical stimulation of the habenula was shown to inhibit DA neuron activation in the VTA and the SN (see also Christoph, Leonzio, & Wilcox, 1986; Ji & Shepard, 2007). In view of this complementary activation pattern, Matsumoto and Hikosaka (2007, 2009) proposed that the habenula may act as a permissive system for midbrain DA neurons. In line with this, evidence for lateral habenular activation during negative prediction errors in humans was provided in a functional magnetic resonance imaging (fMRI) study (Salas, Baldwin, de Biasi, & Montague, 2010). However, the habenula and the midbrain DA neurons may not compete but cooperate with each other. Tian and Uchida (2015) argued that both the habenula and the midbrain DA neurons seem to underlie prediction errors signals. They found that the DA neurons' normal inhibitory response was impaired for negative prediction errors when the lateral habenula was lesioned in rats performing a classical conditioning paradigm. The effect of habenula lesions was less evident for positive prediction errors, as DA neuron activation was still affected by reward expectation. The authors concluded that the habenula plays a critical but not exclusive role in prediction error coding and thus rejected the idea that prediction error signals are only relayed from the habenula to midbrain DA neurons.

DA neurons in the midbrain are most densely located in the VTA and the SN pars compacta (Moore & Bloom, 1978; Smith & Kieval, 2000). Midbrain DA neurons project from the SN pars compacta to the nucleus caudatus and the putamen in the dorsal part of the striatum, and from the VTA to the nucleus accumbens (NAcc) in the ventral striatum and the prefrontal cortex (PFC) including the anterior cingulate cortex (ACC; Bédard, Larochelle, Parent, & Poirier, 1969; Gerfen, 1992; Haber & Fudge, 1997; Joel & Weiner, 2000; Lavoie, Smith, & Parent, 1989; Lynd-Balta & Haber, 1994; S. M. Williams & Goldman-Rakic, 1998; for review, see Kandel et al., 2000). In addition, midbrain DA neurons from the VTA also project to the hippocampus (HC; Gasbarri, Sulli, & Packard, 1997; Kandel et al., 2000). Consistently, reward-related activation and also prediction error signals have been reported in the striatum (O'Doherty, 2004; O'Doherty et al., 2004; Pagnoni, Zink, Montague, & Berns, 2002) but also in the ACC (Hayden, Heilbronner, Pearson, & Platt, 2011; Jahn, Nee, Alexander, & Brown, 2014), as well as in the HC (Foerde & Shohamy, 2011a; Gasbarri

et al., 1997). Other studies generally associated the striatum with reward processing and motivation (Cardinal, Parkinson, Hall, & Everitt, 2002) and stimulus-reward and stimulus-action-reward associations (O'Doherty et al., 2004), with probably dissociable roles of the ventral and dorsal parts.

As midbrain DA signals conveyed to the striatum seem to be particularly involved in associating outcomes to stimuli or to actions, neurological disorders such as PD are of particular interest for reward learning research in humans. PD is defined by depleted levels of DA in the SN, which affects the basal ganglia and causes motoric and cognitive impairments including dysfunctional executive controls and learning (see below). Studies comparing PD patients with healthy controls can thus yield important insights into the neuronal underpinnings of learning as well as the functionality of striatal DA in humans. Furthermore, in severe cases of PD, patients are treated with a deep-brain stimulator that is implanted in the brain, enabling single neuron recording during the surgery (e.g., Zaghloul et al., 2009).

1.2.1 Parkinson's Disease

PD is a slowly progressing degenerative disorder of the central nervous system. It is a sub-category of extrapyramidal and movement disorders as it mainly affects the motor system (World Health Organisation, 1992). The pathognomonic symptoms of PD are akinesia, rigidity, tremor, and postural instability, but cognitive or affective symptoms may also occur, e.g., executive function impairments, dementia, anxiety, and depression (Hacke, 2010). In Northern-American and Middle-European countries, PD is the most prevalent neurological disease, with 0.1-0.2% of the population suffering from it. Men are affected more frequently than women and age fosters the disease so that the prevalence increases to 1.8% in the population older than 65 years (Hacke, 2010). Despite cognitive and affective impairments, PD is prominently classified in five stages based on the severity of the motor symptoms as assessed with the Hoehn and Yahr scale (Hoehn & Yahr, 1967). A multidimensional assessment of the disease is provided by the Unified Parkinson's Disease Rating Scale (UPDRS), which quantifies the symptom severity for a wide range of cognitive and motor functions (UPDRS; Goetz et al., 2008).

On a neuronal level, PD is characterised by a degeneration of DA neurons in the SN pars compacta that affects the basal ganglia (Parent, Levesque, & Parent, 2001). More specifically, depleted nigro-striatal DA projections cause less excitation of striatal D1 receptors, while D2 receptors are less inhibited by DA. Most neuronal projections within the basal ganglia are transmitted via the inhibitory synapses using gamma-amino butyric acid (GABA; Parent et al., 2001) as neurotransmitter. However, due to the differential effect of nigro-striatal DA projections on D1 and D2 receptors, neural projections are conveyed further to the cortex via two pathways – a direct excitatory and an indirect inhibitory pathway. On the direct pathway, DA binds to D1 receptors and thereby excites striatal GABA projections to the Globus pallidus (GP) internum and the SN pars reticulata. As a result, the thalamus is disinhibited as it receives less GABAergic input from the GP internum/SN pars reticulata and projects excitatory Glutamate (GLU) to the (motor) cortex. The loop closes with the thalamus in turn projecting an excitatory signal back to the striatum. On the indirect pathway, however, DA projections from the SN pars compacta inhibit the striatum via D2 receptors, which causes the striatum to project less GABA to the GP externum. As the GP externum becomes more active, it projects more GABA to the subthalamic nucleus, resulting in less GLUergic excitation of the GP internum/SN pars reticulata (Parent et al., 2001). Accordingly, the thalamus is disinhibited, which again excites the motor cortex. In summary, DA activity excites the motor cortex via both the direct and the indirect pathways, while a lack of DA inhibits it (Gerfen, 1992; Hernandez-Lopez, Bargas, Surmeier, Reyes, & Galarraga, 1997).

When DA levels in the SN pars compacta are generally diminished such as in PD, the excitatory D1 receptors in the striatum become less active, resulting in decreased disinhibition of the thalamus via the GP internum/SN pars reticulata and ultimately decreased motor cortex excitation. On the indirect pathway, the striatum becomes less inhibited by D2 receptors when DA projections are lacking so that the subthalamic nucleus is less disinhibited via the GP externum. This causes stronger GLU projections to the GP internum/SN pars reticulata, which projects more GABA to the thalamus, again resulting in reduced excitatory projection to the (motor) cortex (Parent, 2001).

This pathologic physiological process is widely considered as the most important cause of the prevalent motor symptoms in PD (Hacke, 2010). Furthermore, based on a large number of studies relating DA to learning in monkeys (see above) and neuronal plasticity (Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010), depleted levels of DA most probably also cause the cognitive deficits observable in PD (Dubois & Pillon, 1996; Sawamoto et al., 2008). For this dissertation, learning impairments in PD are of particular interest, as the prominent DA depletion in PD may help to gain specific insights into mechanisms underlying learning from feedback.

1.2.1.1 Learning in Parkinson's Disease

Studies on learning in PD reported that altered striatal function in PD negatively affects the patients' performance in reinforcement and category learning tasks (Knowlton, Mangels, & Squire, 1996; Shohamy et al., 2004). Shohamy et al. (2004) linked performance impairments in PD patients explicitly to learning from feedback for actions as opposed to learning from associated stimuli. PD patients seemed to be particularly impaired in the ability to acquire new knowledge and to adapt behaviour accordingly, but not in the ability to transfer knowledge to a new task (Myers et al., 2003). The decreased DA levels in the basal ganglia are thought to be the main cause of these cognitive impairments in PD, which suggests that DA replacement medication would improve PD patients' performance in learning tasks. However, PD specific DA replacement medication, was shown to improve some PD specific cognitive symptoms, while it worsened the performance in other tasks (e.g., Cools, Altamirano, & D'Esposito, 2006; Cools, Barker, Sahakian, & Robbins, 2001; Graef et al., 2010; Mehta, Swainson, Ogilvie, Sahakian, & Robbins, 2001; Swainson et al., 2000). For example, Cools et al. (2001) showed that DA replacement medication improved PD patients' ability to switch between naming either digits or letters in stimuli with mixed letters and digits. In contrast, DA medication impaired the patients' reversal learning performance. Participants in the same study had to choose one out of two stimuli according to its colour to receive reward. After a few trials, the stimulus-outcome contingency was changed so that the participants were thought to choose the other colour to receive reward. PD patients performed significantly worse compared to

controls only when they were ON medication. This raised the question how exactly nigro-striatal DA affects learning.

Frank, Seeberger, and O'Reilly (2004) addressed this question by comparing PD patients' performance in a reinforcement learning task ON versus OFF medication. The authors uncovered a dissociation of the PD patients' learning performance in relation to the feedback's valence and, at the same time, in relation to their DA medication status. PD patients ON their usual DA replacement medication performed better when choosing stimuli that were previously rewarded than when avoiding stimuli that were previously associated with negative feedback. The opposite pattern was found for patients OFF medication. They avoided stimuli that were previously associated with negative feedback more reliably than they chose previously rewarded stimuli (Frank et al., 2004).

This behavioural pattern was also confirmed by computational models simulating DA projections via the two separate pathways within the basal ganglia (Frank, 2005; O'Reilly & Frank, 2006). In line with the findings on reinforcement learning in animals outlined above (Schultz, 1997, 1998, 2002; Schultz et al., 1997), positive feedback caused phasic DA neuron activity bursts in the basal ganglia, while omitted reward or negative feedback caused phasic DA neuron activity dips (Frank, 2005; O'Reilly & Frank, 2006). According to these computational models, phasic DA activity bursts in the SN facilitate a Go-signal on the direct pathway via the D1 receptors that activate the frontal cortex, while they impede the D2 receptors on the indirect pathway. Vice versa, phasic DA dips send a NoGo-signal to the frontal cortex as the indirect pathway becomes facilitated and the direct pathway becomes less activated (Frank, 2005; O'Reilly & Frank, 2006). If the tonic level of DA is diminished such as in PD patients OFF medication, phasic DA bursts following positive feedback seem to sufficiently reach the necessary threshold to reach a net activation of the frontal cortex. Conversely, due to diminished DA levels, DA dips after negative feedback more easily affect the indirect pathway so that the frontal cortex is inhibited more easily (Frank, 2005; Frank, Seeberger, & O'Reilly, 2004). This model explains differential performances in feedback learning tasks in unmedicated PD patients (Frank et al., 2004; Kobza et al., 2012). DA replacement medication, however, creates a tonically

increased DA level that facilitates Go-pathway activation by binding to D1 receptors, while also inhibiting the NoGo-pathway by binding on D2 receptors (Frank et al., 2004; Hacke, 2010). Thus, additional phasic DA bursts during positive feedback seem to over-activate the Go-pathway while the NoGo-pathway is strongly inhibited, which in the end may lead to a stronger representation of positive events (Cools et al., 2001; Cools, Barker, Sahakian, & Robbins, 2003; Frank, Samanta, Moustafa, & Sherman, 2007; Frank et al., 2004; Kobza et al., 2012).

Computational models not only provided a convincing explanation for differential behavioural findings in PD patients ON and OFF medication, but crucially strengthened the link between feedback learning and DA neuron activation. Yet, as outlined above, learning from feedback implies the ability to monitor and evaluate actions and the consequential outcomes in relation to the predicted outcomes. Thus, a feedback based learning system most probably implies an additional, more complex module that critically monitors prediction errors, evaluates them, and eventually sends back a teaching signal for future interactions (Barto, 1995). Extending behavioural studies and computational models by providing a direct measure of neuronal activity in the cortex (Luck, 2014), electrophysiological research has substantially helped to understand neuronal mechanisms that possibly underlie feedback learning.

1.2.2 Electrophysiological correlates of feedback learning

Starting in the early 1990s, electrophysiological research on learning focussed on the error-related negativity (ERN) or error negativity (Ne; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993) and the feedback-related negativity (FRN) or feedback-ERN (fERN; Gehring & Willoughby, 2002; Miltner, Braun, & Coles, 1997). The ERN is a response-locked medio-frontal event-related potential (ERP) component in the electroencephalography (EEG) that is defined by a negative deflection peaking about 50 to 100 ms after erroneous responses (Falkenstein et al., 1991; Gehring et al., 1993). It is most pronounced at fronto-central electrode sites and has been associated with error monitoring and error detection (Gehring et al., 1993). When a similar negative deflection was later also found after correct responses, the correct-response negativity (CRN; Falkenstein, Hoormann,

Christ, & Hohnsbein, 2000; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000), it was suggested that the ERN is more likely to reflect response conflict rather error detection (Carter et al., 1998). In accordance with the reinforcement learning theory (RL-theory, see below), the ERN is supposably generated in the ACC, where it is elicited by error-commission (Holroyd & Coles, 2002; Holroyd, Dien, & Coles, 1998; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001). In line with this, Falkenstein et al. (2001) found diminished ERN amplitudes in PD patients compared to healthy controls during error-commission in different learning tasks. As PD is mostly caused by dysfunctions in the basal ganglia, the authors concluded that the ACC, the most probable generator of the ERN, interacts with the basal ganglia during error detection and action monitoring. The ERN may thus reflect a monitoring system that is able to detect errors based on a representation of the desired ideal action according to an internal forward model (Gehring et al., 1993). Corroborating this conclusion, the ERN was not observed when a correct action could not clearly be discriminated from an erroneous action and increased in amplitude with higher stimulus validity (Eppinger, Kray, Mock, & Mecklinger, 2008). Accordingly, the ERN is affected by the ability to build up expectations on correctness of an action being most pronounced during clearly identifiable erroneous actions (Eppinger et al., 2008).

Importantly, the ERN is a response-locked ERP component, while the FRN, sometimes also called fERN, is an ERP component defined by a relative negativity that is elicited by feedback following an action (Gehring & Willoughby, 2002; Miltner et al., 1997). It emerges between 200 and 300 ms post-feedback (Gehring & Willoughby, 2002; Miltner et al., 1997) and is quite consistently reported to be increased for negative compared to positive feedback (e.g., Hajcak, Moser, Holroyd, & Simons, 2006; Hajcak, Moser, Yeung, & Simons, 2005; Holroyd & Coles, 2002; Holroyd, Hajcak, & Larsen, 2006; Holroyd & Krigolson, 2007; Simons, 2010; Walsh & Anderson, 2012; but see Ferdinand, Mecklinger, Kray, & Gehring, 2012; Oliveira, McDonald, & Goodman, 2007). For this reason, the FRN was frequently analysed as a difference wave between ERPs for positive and negative feedback (e.g., Bellebaum & Colosio, 2014; Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd & Coles, 2002;

Holroyd & Krigolson, 2007; Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Peterburs, Kobza, & Bellebaum, 2016; Walsh & Anderson, 2012).

Aside from feedback valence, several studies reported a modulation of the FRN by expectancy, usually with increased FRN amplitudes for unexpected compared to expected feedback (Bellebaum & Daum, 2008; Ferdinand et al., 2012; Hajcak et al., 2007; Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Holroyd et al., 2009; Oliveira et al., 2007; Yasuda, Sato, Miyawaki, Kumano, & Kuboki, 2004). Along these lines, the FRN reflects a prediction error (Holroyd et al., 2009).

These findings strongly corroborate the RL-theory (Holroyd & Coles, 2002). This theory extends the actor-critic model (Barto, 1995) by linking the FRN to DA projections from the midbrain to the striatum and the ACC. The RL-theory proposes that midbrain DA neurons project prediction error signals for feedback that is better or worse than expected to the basal ganglia, including the striatum, and to the PFC, including the ACC. The striatum evaluates the prediction error signal, revises expectancy accordingly, and forwards a teaching signal to improve future predictions (Nieuwenhuis et al., 2002). The ACC acts as control system, using the information provided by the DA projections to adapt behaviour and thus improve performance. Together, the dopaminergic error signal conveyed from the mesencephalic DA neurons disinhibits the ACC, where it elicits the FRN when events are worse than expected (Holroyd & Coles, 2002).

Indeed, both reward-related and prediction error related activations have been found in the ACC (Amiez, Joseph, & Procyk, 2005, 2006; Jocham, Klein, & Ullsperger, 2011; Orr & Hester, 2012; Sallet et al., 2007; Ullsperger et al., 2014). In humans, the basal ganglia and the PFC including the ACC have generally been associated with executive functions such as planning and decision making (Cummings, 1993; Damasio, 1994; Posner & DiGirolamo, 1997). FMRI studies in human participants reported ACC activity related to reward, reward expectation, and prediction errors (Jocham, Neumann, Klein, Danielmeier, & Ullsperger, 2009; Knutson & Cooper, 2005; Knutson, Fong, Adams, Varner, & Hommer, 2001). Furthermore, a study with human participants undergoing planned surgical cingulotomy revealed crucial involvement of the dorsal ACC in adaptive behaviour

(Z. M. Williams, Bush, Rauch, Cosgrove, & Eskandar, 2004). Corroborating evidence was found in post-error performance updating, which was impaired in patients with a partly lesioned ACC (di Pellegrino, Ciaramelli, & Ladavas, 2007).

The RL-theory linked DA projections targeting the ACC to electrophysiological correlates of feedback processing integrating it to the so-called reward system. This link is further supported by studies localising the FRN in the ACC (e.g., Bellebaum & Daum, 2008; Gehring & Willoughby, 2002; Gruendler, Ullsperger, & Huster, 2011; Miltner et al., 1997; Zhou, Yu, & Zhou, 2010; for review, see Walsh & Anderson, 2012) as well as some other studies that ascribed the FRN to striatal activation (Carlson, Foti, Mujica-Parodi, Harmon-Jones, & Hajcak, 2011; Foti, Weinberg, Bernat, & Proudfit, 2015; Foti, Weinberg, Dien, & Hajcak, 2011). Becker, Nitsch, Miltner, and Straube (2014) linked striatal activation to reward-related ERPs in the FRN time window by merging simultaneously acquired EEG and fMRI data. In this study, trial-to-trial variance in amplitude explained changes in the blood-oxygen-level-dependent (BOLD) signal in the ventral striatum and the ACC. These findings link the FRN to dopaminergic activation of the medial-frontal reward system, qualifying it as a useful indicator of neuronal processing in the DA system during the course of learning.

Corroborating evidence is provided by effects of learning on the FRN that resemble learning-related changes of the prediction error. With increasing knowledge of stimulus-outcome associations, the FRN decreases (Eppinger et al., 2008; Holroyd & Coles, 2002; Walsh & Anderson, 2012). The FRN decreases in amplitude the more accurate a prediction about an action-outcome association is and thus, the more expected a feedback is (Holroyd & Coles, 2002; Walsh & Anderson, 2012). As prediction errors necessarily become less pronounced and eventually vanish with increasing prediction accuracy and feedback expectation (Schultz et al., 1993), leaning-related decreases in FRN amplitude resemble prediction error signals. Considering the RL-theory, the FRN decreases with learning as the feedback becomes less and less useful for a possible update of the internal forward model and behavioural adaptation (Holroyd & Coles, 2002). In contrast, the response-locked ERN increases in amplitude with prediction accuracy (Eppinger et al., 2008). The internal forward

model becomes more accurate and more easily detects actions deviating from the intended behaviour that is predicted to cause the best outcome. Accordingly, the FRN and the ERN are inversely related during the course of learning (Holroyd & Coles, 2002; Walsh & Anderson, 2012). This effect was, however, found to mostly result from positive, not negative events and feedback, respectively (Eppinger et al., 2008).

The RL-theory mostly links negative prediction errors to the FRN, supposedly reflecting disinhibited ACC neuron activity when events are worse than expected (Gehring & Willoughby, 2002; Holroyd & Coles, 2002). However, more recent research has highlighted that the ERP signal in the FRN time-window is actually mostly affected by positive events causing a positive deflection (Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Proudfit, 2015). Strong midbrain DA neuron firing bursts for positive prediction errors motivated this idea (Schultz, 2002). Supportive evidence was for example provided by Kujawa, Smith, Luhmann, and Hajcak (2013), who compared ERPs for monetary gains or losses to a neutral outcome, respectively, in two separate conditions. They found that feedback-locked ERPs differed only between gains and the neutral feedback, but not between losses and neutral feedback. Proudfit (2015) argued that the negative deflection following negative feedback could thus be seen as a baseline, while positive feedback causes a positive deflection warranting the label reward positivity (RewP). According to this argumentation, the signal becomes more negative for negative feedback due to the missing positive deflection (Proudfit, 2015). Carlson et al. (2011) reported positive correlations between feedback-locked ERPs for monetary wins and greater neuronal activation for wins compared to losses in the ventral striatum and the medial PFC. Likewise, Becker et al. (2014) linked increasing BOLD responses in the ventral striatum and the ACC to increasing positive amplitudes after positive feedback in the FRN time-window. Foti et al. (2015) analysed the temporal frequency of the FRN and showed that it may be composed of two independent neuronal processes relying on different frequencies, a loss-related process reflected in theta oscillatory activity and a gain-related process reflected in the delta-band activity. Also, these activation patterns were source localised in distinct neuronal structures, suggesting that the ACC processes negative feedback, while gain-related activity was associated with the basal ganglia. Accordingly, two different neuronal

processes might underlie ERPs for positive and negative feedback that may overlap in the FRN time-window. To emphasise the stronger effect of positive feedback on the feedback-locked ERPs, Holroyd et al. (2008) and Proudfit (2015) proposed to refer to the positive deflection following positive feedback as RewP.

The idea to interpret the negative deflection after negative feedback as a baseline ERP component that generally follows after feedback, was inspired by the close similarity of the FRN to the N200 (J. R. Folstein & Van Petten, 2008; Holroyd, 2004). The N200 is a fronto-central negativity that peaks after about 250 ms in response to infrequent, task-relevant events (i.e., targets; Holroyd, 2004; Towey, Rist, Hakerem, Ruchkin, & Sutton, 1980), with increasing amplitude for decreasing target stimulus frequency (Duncan-Johnson & Donchin, 1977; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003), and for feedback that does not carry information about a previous behaviour (Baker & Holroyd, 2009). Indeed, for years, the FRN was regarded as a special case of the N200. Holroyd et al., (2008), however, suggested that the N200 may rather reflect a baseline response to feedback that is suppressed by feedback that is better than expected. According to the RL-theory, DA neurons project prediction error signals to the ACC, to which the RL-theory ascribes a cognitive motor control function (Holroyd & Coles, 2002). Along these lines, ACC activity is affected by response conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004), which is reflected in the N200 (Yeung, Botvinick, & Cohen, 2004). Considering this, Holroyd et al. (2008) proposed that phasic DA neuron bursts signalling unexpectedly positive outcomes would inhibit conflict-related ACC activity and thereby reduce the N200 (Holroyd, 2004). During learning, the ACC would use mesencephalaic DA prediction error signals as a teaching signal to adapt behaviour so that response conflict is minimal (Holroyd et al., 2008). Along these lines, uninformative feedback elicits a baseline negative deflection, the N200 (Holroyd, 2004), while positive feedback causes a positive deflection, the RewP, that possibly reflects ACC activity when no conflict has to be processed (Holroyd et al., 2008; Proudfit, 2015).

Irrespective of whether one refers to the FRN as a negative deflection after unfavourable feedback or to the RewP as a positive deflection after favourable

feedback, feedback-locked EPRs are usually measured as the difference between positive and negative feedback. Accordingly, many studies quantified the FRN based on a positive-negative difference wave (e.g., Hajcak et al., 2007; Holroyd et al., 2009; Proudfit, 2015; Walsh & Anderson, 2012). However, the appropriateness of this approach is still being discussed (Proudfit, 2015). Difference waves can be logically interpreted only if the original waveforms differ in merely one parameter (e.g., Luck, 2014). As outlined above, opinions differ on whether the FRN/RewP is mainly caused by signal variations after negative or positive feedback (Proudfit, 2015). Furthermore, the influence of feedback expectancy is still being debated (Ferdinand et al., 2012; see below).

Although the RL-theory has dominated theoretical considerations concerning the link between ACC activity (and thus the FRN) on the one hand and learning on the other hand, another approach also demands attention. The predicted response-outcome (PRO) model links the FRN to an expectancy coding system rather than to a reward learning system (Alexander & Brown, 2010, 2011; Ferdinand, Mecklinger, Kray, & Gehring, 2012). This approach states that neurons in the ACC learn to predict actions outcomes and become active when an outcome is expected. Activity in the ACC is thought to increase until the predicted outcome actually occurs, which inhibits the prediction signal. When a predicted outcome is omitted, however, activity in the ACC becomes maximal. Accordingly, the PRO-model suggests ACC activation irrespective of the outcome's valence or value, which distinguishes it from the RL-theory. It proposes that ACC activity is affected by unexpected events, a notion that has been supported by some studies (Ferdinand et al., 2012; Hayden et al., 2011; Matsumoto, Matsumoto, Abe, & Tanaka, 2007; Wessel, Danielmeier, Morton, & Ullsperger, 2012).

In line with this, recent studies linked the FRN to a salience prediction error (SPE) that is independent of valence. According to this line of research, SPEs elicit an FRN because rewards are more salient compared to punishments. Indeed, evidence for this idea was provided by studies that compared feedback-locked ERPs for expected versus unexpected feedback (Soder & Potts, 2017; Talmi, Atkinson, & El-Deredy, 2013). Importantly, both studies investigated expectancy effects on the FRN in two separate

trial conditions for positive and negative feedback, respectively. Both studies found evidence that both positive (unexpected punishment omission) as well as negative prediction errors (unexpected reward omission) are reflected by the FRN. This made them suggest that the FRN reflects an SPE rather than a reward prediction error. Recently, however, Mulligan and Hajcak (2017) criticised these studies for having used passive paradigms to induce reward expectancy. As the participants did not make choices themselves, the feedback was not relevant for behavioural adaptation and may thus not maximise the FRN (Holroyd et al., 2009; Walsh & Anderson, 2012). When participants were tested in a gambling task with equally distributed reward-punishment probabilities, evidence was again found for the FRN's sensitivity to feedback valence rather than for feedback salience (Mulligan & Hajcak, 2017).

Taken together, electrophysiological correlates of feedback-based learning have been a fruitful measure for investigating the neuronal underpinnings and mechanisms involved in learning. However, while actor-critic based models emphasise the involvement of dopaminergic error signals, other models do not build on mesocortical and mesolimbic DA projections. Also, it is still a matter of debate, to what extent striatal and ACC activity, as reflected in the FRN, is affected by an outcome's valence, expectancy, and value. To further elucidate the neuronal mechanisms and communications underlying feedback learning, EEG data has not only been investigated regarding ERPs but also regarding oscillatory activity. For example, two separate processes that are both reflected by the FRN can be distinguished by separate frequency bands, i.e., theta band activation for negative and delta band activation for positive feedback (Foti et al., 2015). Like the FRN, this oscillatory activity was source localised in the ACC. Oscillatory activity during feedback processing may thus provide additional information about the neuronal underpinnings of feedback learning.

1.2.3 Oscillations in the EEG

Over the past decades, another approach to analyse EEG data has received increasing attention in the field of electrophysiological research. Cohen (2011b) argued that ERP analyses miss out on important information that may reliably reflect neuronal activation. According to this line of reasoning, the temporal dynamic of

oscillatory electrophysiological activity, i.e., the deflections per time, indicate neuronal activation and thus hold important information such as neuronal communication that cannot be assessed by ERPs (Cohen, 2011b). Furthermore, electrophysiological oscillations may directly reflect neuronal activation as they are known to reflect rhythmic fluctuations in excitability (Luck, 2014; X. J. Wang, 2010). Also, oscillatory activity was suggested to underlie synaptic plasticity, which fortifies its relevance for research on learning (Cohen, Wilmes, & Vijver, 2011). Just like ERPs, electrophysiological oscillations have been associated with different underlying neuronal processes depending on their temporal dynamics and spatial distribution with widely varying frequencies (Buzsaki & Draguhn, 2004). Cognitive processes have been found to be mostly reflected by oscillations in the delta (1-4 Hz), theta (4-8 Hz), beta (20-30 Hz), and gamma (30-80 Hz) frequency bands (e.g., Buzsaki & Draguhn, 2004; Cavanagh, Figueroa, Cohen, & Frank, 2012; Cavanagh & Frank, 2014; Cohen, 2011a, 2014; Cohen et al., 2011; HajiHosseini & Holroyd, 2015; HajiHosseini, Rodriguez-Fornells, & Marco-Pallares, 2012; Marco-Pallares, Cucurell, et al., 2008; Mas-Herrero, Ripolles, HajiHosseini, Rodriguez-Fornells, & Marco-Pallares, 2015).

Furthermore, electrophysiological oscillations carry multidimensional information that may provide insights in neuronal involvement and inter- and intra-structural communication. More precisely, the above mentioned frequency is just one dimension reflecting the speed of variations in neuronal excitability. Oscillations also carry information about the energy of a frequency at a certain time (i.e., power) and its onset (measured by its position along the sine wave, i.e., phase; Cohen, 2011b).

Another important point is that distinct oscillatory dynamics can specifically differentiate separate processes. Cohen (2011b) pointed out that several cognitive processes were spatially linked to the same structure. For instance, as also partly described above, the PFC including the ACC has been associated with monitoring of actions, conflicts, feedback processing, and behavioural adaptation. These processes may be distinguishable by the frequency of electrophysiological oscillations (Cohen, 2011b; Cohen et al., 2011). It seems plausible that the ACC integrates several functional networks that are not distinguishable by their location, but possibly by oscillatory dynamics (see also Cohen et al., 2011).

In the field of reward processing, research focussed on two frequency bands, the theta and the beta frequencies. While oscillations in the theta band were linked to processing of negative feedback (Janssen, Poljac, & Bekkering, 2016) and prediction error coding (Cavanagh, Figueroa, et al., 2012; Cavanagh & Frank, 2014; Cavanagh, Frank, Klein, & Allen, 2010; Cohen, 2011a), beta and low gamma oscillations were reported for reward processing (Cohen et al., 2011; Marco-Pallares, Cucurell, et al., 2008; Mas-Herrero et al., 2015). Both frequency bands were attributed to medial frontal networks (Cohen et al., 2011; Foti et al., 2015) and may thus, similar to medial-frontal ERPs such as the FRN/RewP (Foti et al., 2015), reflect striatal phasic and decreased prefrontal tonic DA activity (Marco-Pallares, Cucurell, et al., 2008). In line with this, theta oscillations are thought to underlie several ERP components such as the FRN, the N2, but also the ERN (Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012).

However, despite its functional similarity to ERP components elicited by negative feedback (Cavanagh & Frank, 2014), it is still unclear whether theta band activation reflects prediction error coding or rather the allocation of attentional resources driven by a need for cognitive control (Cavanagh & Frank, 2014). On the one hand, theta power has been shown to increase in generally uncertain environments and for prediction errors without a differential effect of their valence (Cavanagh, Figueroa, et al., 2012), but on the other hand also for errors (Cohen, 2011a), negative feedback (Cohen et al., 2011), and conflict (Cohen, 2014). Thus, an increase in theta power has been proposed to reflect a rather general top-down alarm signal in the need of cognitive control when events are generally “bad” or unexpected (Cavanagh & Frank, 2014; Cavanagh et al., 2010; Cavanagh & Shackman, 2015; Cohen, 2014). In line with this, increased theta frequency oscillations have been suggested to reflect increased structural connectivity of the ACC and pre-frontal cortical structures during error processing (Cohen, 2011a). In contrast, feedback-related ERPs have been associated with positive and negative feedback (see above) and possibly DA neuron projections to the striatum and ACC (Holroyd & Coles, 2002). Accordingly, theta band osciallatory power may provide persuing information on the processing and the transmission of prediction error and error signals across structures involved in action

monitoring and feedback based learning (Cavanagh & Frank, 2014; Cohen et al., 2011).

Similarly, but more specifically, beta band oscillations have been hypothesised to reflect a DA-driven motivational signal that drives mnemonic consolidation of beneficial events such as rewards (Feingold, 2011; Marco-Pallares, Munte, & Rodriguez-Fornells, 2015; Mas-Herrero et al., 2015). In a study that combined fMRI and EEG, Mas-Herrero et al. (2015) found reward-related mid-frontal beta power increases associated with fronto-striatal-hippocampal activation. In line with findings in non-human primates suggesting a frontal-hippocampal synchronisation during reward processing reflected by beta oscillations (Feingold, 2011), Mas-Herrero et al. (2015) proposed that beta oscillations may signal rewards to the HC which becomes active when an action leading to a rewarding outcome has to be memorised. This reward signal probably relies on DA projections (Gasbarri et al., 1997; Mas-Herrero et al., 2015) in line with studies relating beta-gamma oscillations to striatal phasic and decreased prefrontal tonic DA activity (Marco-Pallares, Cucurell, et al., 2008).

1.3 Conditions with altered feedback learning mechanisms

Over the course of the past two decades, research on feedback-based learning has excessively investigated different manipulations affecting the underlying neuronal mechanisms. As outlined above, debates are still not completely abated on whether expectancy or feedback valence are the primary force affecting feedback-related ERPs during learning. Furthermore, feedback processing may depend on a variety of factors, such as the feedback's usability for performance optimisation (Holroyd et al., 2009) or personal preferences (Chib, Rangel, Shimojo, & O'Doherty, 2009). Although direct evidence for a manipulation of the FRN by subjective reward value is as yet lacking, increased activation in the ventral-medial PFC was reported for preferred compared to non-preferred rewards in human participants (Chib et al., 2009). Reward magnitude, on the other hand, has been shown to affect the FRN (Bellebaum, Polezzi, & Daum, 2010; Polezzi, Sartori, Rumiati, Vidotto, & Daum, 2010; Wu & Zhou, 2009), in line with a study that found monkey DA neurons in the SN and the VTA to code reward-

magnitudes, with the strongest firing rate increases for the most preferred rewards and decreased firing for the least preferred rewards (Lak, Stauffer, & Schultz, 2014).

It becomes increasingly evident that neuronal mechanisms in the medial-frontal reward system valuate feedback, link it to previous actions, and update plans for future actions accordingly. However, learning cannot only be achieved by trial-and-error, but also by observing feedback that refers to another person's actions. This raises the question whether the same neuronal mechanisms are involved, given that feedback for an observed action does not necessarily link to one's own previous behaviour that needs to be adapted in order to improve performance.

1.3.1 Observational learning

One line of research addressed this question by investigating neuronal mechanisms of learning from feedback that refers to observed as opposed to own actions. From an evolutionary perspective, it makes sense for living organisms to learn from feedback that conspecifics receive for their actions instead of trying several (potentially risky) actions themselves. This way, organisms can associate a beneficial outcome with an action without taking the risks to receive a negative, possibly harmful outcome. Accordingly, feedback that is not related to one's own action still seems to be used to update an internal forward model that predicts action outcomes. The ACC has been proposed to receive meso-cortical DA prediction error signals during learning that train it to select the best possible action relative to a desired goal (Holroyd & Coles, 2002). Thus, the ACC might be less involved when the feedback is self-relevant, but no action has to be selected.

This idea was investigated by Bellebaum, Kobza, Thiele, and Daum (2010). The authors found diminished FRN amplitudes in participants learning from observed feedback as compared to participants who actively performed the same learning task. Participants were asked to learn from the feedback in both groups and learning performance was assessed in test trials in which no feedback was provided and in which both observational and active learners had to actively select response options. Thus, feedback was also relevant for the observationally learning participants as they needed to update their reward predictions, although they did not need to adapt

behaviour on a trial-by-trial basis. Interestingly, although the FRN amplitudes for negative feedback differed, learning performance was comparable between groups. This result strengthens the view that the ACC is important for behavioural adaptation, as particularly negative feedback induced a larger electrophysiological response when feedback could be used to adapt the participants' own behaviour. For positive feedback, ACC activation is supposedly less involved because it does not imply the necessity to change behaviour. In line with these findings, a later study found that actively performing participants learned as well from feedback as participants that observed others' performances, while the FRN was reduced for observational compared to active learning (Bellebaum & Colosio, 2014). This study also reported that the FRN elicited by feedback for the participants' own actions decreased when the action-outcome contingencies were successfully learned. Interestingly, the FRN for feedback that referred to somebody else's feedback was not modulated by learning, indicating that the neuronal mechanisms underlying feedback processing may differ between active and observational learning.

In line with these results, an imaging study reported greater prediction error related activity in the dorsal striatum when human participants learned stimulus-action-outcome associations than when they learned stimulus-outcome associations. The ventral striatum, however, was found to become active for prediction error coding during both association learning types (O'Doherty et al., 2004). Bellebaum, Jokisch, Gizewski, Forsting, and Daum (2012) later suggested that especially the anterior caudate nucleus in the dorsal striatum might be involved in linking evaluative feedback to own actions. In their study, participants again learned from either feedback for own or observed actions, with comparable performance in both conditions. Prediction error related activation differed between groups with greater activation in the dorsal striatum in participants that received feedback for their own actions compared to participants receiving feedback for an observed action. Resembling O'Doherty et al.'s (2004) findings, this group-difference was not found in the ventral striatum. As the striatum has also been associated with the FRN (Carlson et al., 2011; Foti et al., 2015; Foti et al., 2011), less striatal activity may also contribute to a diminished FRN amplitude when feedback refers to observed actions.

Additional insight was provided by a clinical study investigating active and observational feedback learning in PD patients (Kobza et al., 2012). As described above, PD patients OFF medication learn better from negative than positive feedback, which has been related to depleted striatal DA levels (Frank, 2005; Frank et al., 2007; Frank et al., 2004). Kobza et al. (2012) compared this learning bias in patients performing an active learning task to patients performing an observational learning task. They found that the negative learning bias in PD patients during active learning changed into a rather positive bias, which was comparable to the healthy controls' performance. This result indicates diminished striatal involvement when feedback does not refer to own actions and thus corroborates its supposed role in linking outcomes to own actions.

1.3.2 Learning from delayed feedback

To this date, most studies investigating the neuronal mechanisms of feedback learning used tasks, in which the feedback followed the related action immediately, i.e., after less than a second. However, in everyday life, feedback for an action may be temporally delayed by several seconds. Therefore, some studies investigated the impact of feedback timing on the involved underlying mechanisms during learning finding hints for dissociable neural systems involved in the processing of immediate and delayed feedback (e.g., Foerde et al., 2013; Foerde & Shohamy, 2011a). Apart from the dorsal striatum's role in ascribing outcomes to self-generated actions, a different line of research hinted at possible striatal involvement in linking outcomes only to recent actions (Foerde & Shohamy, 2011a). In contrast, possibly the medial-temporal lobe (MTL) including the HC may bind elements across time. Investigating this hypothesis, a study using fMRI reported increased activation in the HC when feedback was processed that followed an action only after several seconds, while activity in the dorsal striatum was diminished (Foerde & Shohamy, 2011a). In a later study, the same researchers reported a double-dissociation in patients with brain damage, with PD patients being impaired in learning from immediate but not delayed feedback, while amnestic patients with suspected hippocampal lesions suffered from impaired learning from delayed, but not immediate feedback (Foerde et al., 2013). The

authors suggested that neuronal prediction error coding may shift from the basal ganglia to the HC if there is a temporal gap between an action and the related outcome. Interestingly, however, the authors also found comparable prediction error related activation in the ventral striatum during immediate and delayed feedback learning, while the dorsal part was exclusively active during immediate feedback (Foerde & Shohamy, 2011a).

In line with the notion of a shift away from striatal prediction error coding with increasing feedback delay, striatal/ACC activation as reflected in the FRN was found to be diminished for delayed compared to immediate feedback processing (Opitz, Ferdinand, & Mecklinger, 2011; Peterburs et al., 2016; Weinberg, Luhmann, Bress, & Hajcak, 2012). In contrast, J. Wang, Chen, Lei, and Li (2014) failed to find differences in the FRN caused by feedback timing. As the tasks and procedures strongly differed between these ERP studies and the FRN is sensitive to a variety of factors such as subjective reward relevance (see above), it is difficult to draw firm conclusions from these studies. Furthermore, the FRN was not explicitly linked to prediction error signals in any of these studies so that a direct influence of feedback timing on the neuronal mechanisms underlying reward predictions cannot be inferred.

The effect of a temporal delay has also been investigated at the level of single midbrain DA neuron firing patterns in non-human primates (e.g., Fiorillo, Newsome, & Schultz, 2008; Kobayashi & Schultz, 2008). It was found that the subjective value of reward predicting stimuli decreased with increasing temporal delay of the following reward, although the stimuli did not differ in their reward-prediction probability. More precisely, strong DA neuron firing bursts were observed for stimuli that were rewarded 2 s after the stimulus. These neuron firing bursts diminished the more time passed between the predictive stimulus and the actual reward (Fiorillo et al., 2008; Kobayashi & Schultz, 2008). DA neuron activity has previously been found to code a CS' value reflecting the associated reward's magnitude (Tobler, Fiorillo, & Schultz, 2005). In line with diminishing DA neuron firing rates, the monkeys preferred stimuli that predicted small rewards after a small temporal delay over larger rewards after a longer delay (Kobayashi & Schultz, 2008). Similar behavioural preferences have been reported in pigeons and humans (Rodriguez & Logue, 1988), which was explained by

an economic model stating that a reward's value is discounted the longer it is temporally delayed because more time may hold an increased risk of not receiving the reward at all (Samuelson, 1937).

However, the opposite DA neuron firing pattern was observed following the reward itself, when the monkeys engaged in a subsequent task that did not require an active choice. When the reward was received, DA neuron firing activity increased with increasing temporal delay, although the reward's value did not differ (Kobayashi & Schultz, 2008). The authors suggested that a temporal prediction error may cause increased DA neuron firing as the reward's timing may become increasingly unexpected with increasing temporal delay. In line with this, the dopaminergic response was highest for rewards that were both temporally unpredictable and delayed (Fiorillo et al., 2008). As an alternative explanation Kobayashi and Schultz (2008) suggested diminished associative strength for rewards that follow the predictive stimulus after a temporal delay. Animals may struggle to associate stimuli to delayed rewards (Holland, 1980), which is why temporally delayed stimuli may cause a positive prediction error, as the predictive stimulus did not build up a reward expectation in the animal.

The findings by Kobayashi and Schultz (2008) are of particular interest as they seem contradictive to findings in humans regarding feedback-locked ERPs (see above) and hemodynamic responses in fMRI (Foerde & Shohamy, 2011a). Accordingly, midbrain DA neuron activation patterns seem to differ between humans and non-human primates due to the monkeys' inability to link rewards to stimuli that lay back in time. Another possibility is that midbrain DA neurons also become more active for delayed compared to immediate rewards, but project more to the HC than the striatum. Evidence for this proposal is provided by hippocampal activation when delayed feedback is processed (Foerde & Shohamy, 2011a), and by midbrain DA projections to the HC (Gasbarri et al., 1997).

As immediate and delayed feedback seem to be processed differently in any case, further research on the influence of feedback timing on reward processing would advance the understanding of feedback based learning in general. On one hand, if delayed feedback is not processed in the ACC/striatum but only in the HC, the FRN

should not be manipulated by prediction errors from delayed feedback. This could be assumed due to findings of increasingly diminished FRN amplitudes with an increasing temporal gap between action and the consequential outcome (e.g., Peterburs et al., 2016). Moreover, theoretical considerations are in line with the idea of an anatomical separation of immediate and delayed feedback, given that declarative (including reward-based) and non-declarative learning have been associated with the basal ganglia and the HC, respectively (Delgado, 2007; Sherry & Schacter, 1987). On the other hand, the HC might be involved in processing of delayed feedback as mesolimbic DA projections affect neuronal plasticity and episodic memory formation (Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010) in the HC. The HC might thus enable participants to link outcomes, processed by DA neurons, to actions across a temporal gap. However, a direct comparison of medial-frontal prediction error correlates during learning from immediate versus delayed feedback remains elusive.

1.4 Open questions and research objectives

Findings from behavioural, EEG and fMRI studies show that a temporal delay of 7 s significantly changes neuronal activity patterns during feedback learning in probabilistic feedback learning tasks (Foerde et al., 2013; Foerde & Shohamy, 2011a; Peterburs et al., 2016). These findings suggest that striatal/ACC activity is decreased during delayed feedback processing (Foerde et al., 2013; Foerde & Shohamy, 2011a; Peterburs et al., 2016; Weinberg et al., 2012), which other structures like the HC may compensate for (Foerde & Shohamy, 2011a). In line with this, the distinct non-declarative and declarative memory systems have been associated to the basal ganglia and the HC, respectively (Ashby & O'Brien, 2005; Delgado & Dickerson, 2012; Sherry & Schacter, 1987; Squire, 1992; Squire & Zola, 1996). This motivates the idea of a dichotomous involvement of either the medial-frontal reward system or the HC during immediate and delayed feedback, respectively (see Foerde et al., 2013; Foerde & Shohamy, 2011a).

However, this strict dichotomy between the striatum/ACC and HC does not seem to hold as empirical evidence suggested hippocampal involvement in learning from probabilistic feedback that was provided after 2 s (Dickerson, Li, & Delgado, 2011) to

4 s (Dickerson & Delgado, 2015). Also, midbrain DA projections from the VTA to the HC (Gasbarri et al., 1997; Kandel et al., 2000) where they modulate neuronal plasticity and adaptive memory (Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010) support the idea of a hippocampal contribution to feedback learning. Vice versa, striatal involvement was reported in declarative memory tasks (Scimeca & Badre, 2012) and when delayed feedback was processed (Foerde & Shohamy, 2011a).

This dissertation aims to investigate neuronal processes involved in feedback processing when the outcome of an action is delayed by several seconds, with particular focus on similarities and differences between immediate and delayed feedback processing in the context of probabilistic feedback learning. Importantly, in order to reliably relate behavioural and electrophysiological effects to striatal and ACC activity, the studies in this dissertation explore modulations that are well established for immediate feedback processing in the context of delayed feedback.

Study 1 investigates the notion of a selective impairment of PD patients in learning from immediate but not delayed feedback. This notion is based on striatal DA involvement in immediate but not delayed feedback processing (Foerde et al., 2013). If PD patients do not make use of their striatum during delayed feedback processing, the negative learning bias that is usually observed when PD patients learn from immediate feedback (Frank et al., 2004) should be absent when feedback follows the actions only after 7 s.

Study 2 aims at finding prediction error signals during delayed feedback processing. Striatal activity coding prediction errors may reflect an FRN when an action is followed by unexpected feedback (Carlson et al., 2011; Foti et al., 2011). Similar to Study 1, if the striatum is not involved in delayed feedback processing, expectancy effects should not be observable in the FRN when feedback is delayed.

Study 3 extends Study 2 by investigating the effect of feedback timing and feedback agency (defined as feedback referring to one's own versus somebody else's action) on the FRN and theta and beta band power. Frequency analyses may provide additional information about mechanisms underlying feedback processing that may differ from what the FRN analysis elucidates (Cohen, 2011b). The additional factor of

agency may furthermore hint at differential striatal involvement (Bellebaum et al., 2012; Kobza & Bellebaum, 2015; Kobza, Thoma, Daum, & Bellebaum, 2011).

2 Study 1: Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication

Study 1 investigated whether patients suffering from PD in the OFF medication state learned differentially from positive versus negative immediate and delayed feedback in a probabilistic feedback learning task. PD is characterised by DA neuron depletion resulting in a negative learning bias for learning from immediate feedback. If the striatum is less involved in learning from delayed feedback due to a potentially reduced role of the DA system, this negative learning bias in PD patients should not be evident during delayed feedback processing. This study was published in Neuropsychologia, 117, 46-54 (Weismüller et al., 2018).

2.1 Introduction

Altered DA levels differentially affect learning from positive and negative feedback (Frank et al., 2007; Frank et al., 2004). More specifically, a tonically depleted nigro-striatal DA level such as in PD causes a tendency to learn better from negative relative to positive feedback, while DA replacement medication causes the opposite pattern (Frank et al., 2007; Frank et al., 2004). Foerde et al. (2013) and Foerde and Shohamy (2011a) found that PD patients OFF medication performed comparable to controls in a task that did not differentiate between learning from positive and negative feedback. The authors suggested that delayed feedback processing seems to rely less on the DA system or at least the striatum, but rather appears to recruit the MTL, more specifically the HC. Kobza et al. (2012) reported comparable learning performance for positive and negative feedback in PD patients OFF medication for feedback that did not refer to the patient's own actions and may thus recruit mechanisms other than the striatal DA system. Given that PD patients OFF medication seem to make use of their intact HC when feedback follows an action after a temporal delay, it could be reasoned that delayed feedback alleviates the negative learning bias.

However, a recent ERP study by our group suggests a contribution of the striatum/ACC to delayed feedback processing as the FRN was still observable for delayed feedback (Peterburs et al., 2016). This is in line with literature proposing a cooperation of the striatum and the HC during reward learning (Dickerson & Delgado, 2015; Dickerson et al., 2011). Consequently, the present study investigated whether feedback delay reduces the negative learning bias in PD patients.

2.2 Method

2.2.1 Study participants

Two groups of PD patients (Hoehn and Yahr scale 1-3; Hoehn & Yahr, 1967), and two groups of healthy age-matched control participants were recruited for this study. One group of 12 PD patients and their 24 healthy controls learned from immediate feedback, while the other group of 12 patients and their 24 healthy controls were engaged in learning from delayed feedback. For the participants learning from immediate feedback, we reused data from 10 PD patients and 20 controls that had been acquired for a previous study by our group (Kobza et al., 2012). All patients attended testing in the OFF medication state. Patients were asked to withdraw from their Parkinson specific medication at least 12 hours prior to testing, which was conducted in the morning. Symptom severity was assessed with the UPDRS (Goetz, LeWitt, & Weidenman, 2003), which also helped to verify the medication status. UPDRS scores were higher for the OFF than the ON state for patients from both groups, those learning from immediate and delayed feedback. Patients were screened for the possibly occurring comorbidities dementia and depression using the Mini Mental Status Test (MMST; M. F. Folstein, Folstein, & McHugh, 1975) and the Beck Depression Inventory (BDI; Hautzinger, Keller, & Kühner, 2006), respectively. All patients scored in the normal range for both measures and thus entered the analysis.

2.2.2 Experimental task and procedure

All participants completed a modified version of the probabilistic feedback learning task introduced by Frank et al. (2004), which consisted of three phases. First, in the learning phase, participants were asked to choose one of two previously unknown symbols in each trial to receive a binary feedback ("richtig", the German word for correct, or "falsch", the German word for incorrect). In this phase, three different symbol pairs (symbols A vs. B, C vs. D, and E vs. F) were randomly presented 20 times. Unknown to the participants, a certain reward-probability was locked to each symbols so that the choice of symbol A resulted in positive feedback in 80% of the trials and symbol B in 20% of the trials. Likewise, reward probabilities for the other pairs were 70% versus 30% (C/D pair) and 60% versus 40% (E/F pair).

The learning phase was followed by a test phase, which was identical to the learning phase with the exception that the participants did not receive feedback. The test phase was used to acquire each participant's response accuracy in the absence of trial-to-trial feedback. If the participants chose the symbol with the higher reward probability in 80% of the A/B pairs and 70% of the C/D pairs, the third task phase, the transfer phase, was initiated. If the participants failed to reach one of these criteria, another learning phase was initiated, followed by another test phase. Learning and test phases were repeated up to four times. After the fifth learning phase, the transfer phase was initiated even if participants had not reached the learning criterion.

In the transfer phase, the symbols A and B were paired with all symbols they had not been paired with previously (resulting in the possible combinations A/C, A/D, A/E, and A/F and B/C, B/D, B/E, and B/F), and participants were asked to choose one of the symbols without receiving feedback. Choices of symbol A indicated that participants previously learned from positive feedback during the learning phase, while avoidances of symbol B indicated learning from negative feedback. Thus, by comparing the number of trials, in which participants chose the A symbol and in which they avoided B, it could be inferred whether participants learned better from positive feedback or negative feedback (see Bellebaum, Rustemeier, & Daum, 2011; Frank et al., 2004; Kobza et al., 2012 for a similar approach).

Note that one group of PD patients and their healthy controls received feedback immediately (i.e., 500 ms after the chosen symbol had disappeared) during the learning phase, while the other PD group and their healthy controls received temporally delayed feedback (i.e., 6500 ms after the chosen symbol had disappeared). Test and transfer phases did not differ between these groups.

2.2.3 Data analysis

In order to investigate general learning performance, the mean number of correct choices during the first learning phase was compared between PD patients and controls, between feedback timing groups, and between symbol pairs. Another measure of performance was obtained by the number of phases participants needed to reach a learning criterion of 70% correct choices for the A/B pair during the test phase, which was also compared between groups and feedback timings. The number of choices of symbol A and avoidances of symbol B in the transfer phase were considered as measures for learning from positive and negative feedback, respectively, and were thus the dependent variables of main interest. Accordingly, the ratio of choices of A versus avoidances of B mirrored the learning bias and was compared between groups and feedback timings.

2.3 Results and discussion

Results showed that the number of correct responses during the learning phase differed neither between PD patients and controls, nor between feedback timings. However, participants scored generally higher for pairs with larger relative to lower differences in reward probabilities (e.g., A/B vs. E/F). Participants also needed comparable numbers of learning phases to reach the learning criterion, regardless of feedback timing and of whether or not they suffered from PD. Thus, all participants were similarly well able to learn which symbols most probably resulted in reward irrespective of the feedback timing and the DA level.

Feedback timing differentially affected performance in PD patients and healthy controls: both groups performed comparably when they learned from delayed

feedback, while controls performed better than PD patients when feedback was provided immediately. This finding was at least partially in line with previous studies (Foerde et al., 2013; Foerde & Shohamy, 2011a), which suggested that a diminished striatal DA level, such as in PD patients OFF medication, impaired learning from immediate feedback. Along these lines, learning from delayed feedback may depend less on the dorsal striatum, and rather on the MTL, so that PD patients' learning performance was not affected when feedback was delayed.

Replicating earlier results (Frank, 2005; Frank et al., 2004; Kobza et al., 2012), the transfer phase results showed differential learning from positive and negative feedback in patients and controls: while PD patients avoided symbol B (indicating negative feedback learning) more frequently than they chose symbol A (indicating positive feedback learning), the opposite pattern emerged in healthy controls. As postulated before (Frank, 2005; Frank et al., 2004; Moustafa, Cohen, Sherman, & Frank, 2008; Samson, Frank, & Fellous, 2010), this may reflect the differential effect of striatal DA on learning from positive and negative feedback.

However, in the present study, feedback timing did not affect the learning bias depending on whether or not a participant suffered from PD. This hampers the before mentioned interpretation of less striatal involvement in processing delayed feedback. If striatal DA levels affected learning from positive and negative immediate but not delayed feedback, feedback timing should have differentially influenced the learning bias in PD patients and controls. Accordingly, the current results suggest an effect of striatal DA on learning from both immediate and delayed feedback. Additional separate comparisons of differential learning between the PD patients and controls for each feedback timing corroborated this conclusion. Different learning biases were revealed for the two groups in both feedback timing conditions: PD patients who had previously received immediate feedback, learned better from negative compared to positive feedback during the transfer phase, while the respective controls did not show such a learning bias. When receiving delayed feedback, however, PD patients performed similarly for positive and negative feedback, while a positive learning bias was revealed in the controls. Interestingly, PD patients also learned descriptively better

from negative feedback than controls when feedback was previously provided after a temporal delay.

The result pattern for immediate feedback closely resembles previous studies (Frank, 2005; Frank et al., 2004; Kobza et al., 2012). Also, the tendency to learn better from positive than negative feedback in healthy older adults has been reported before in learning from feedback for observed actions (Bellebaum et al., 2011), although this was found for delayed feedback for the participants' own actions here. When considering findings by Foerde et al. (2013) and Foerde and Shohamy (2011a), delayed feedback processing may rely less on the basal ganglia, but on the HC, which has been associated with declarative information processing (Squire, 1992; Squire & Zola, 1996). The positive bias in the healthy delayed feedback learners is in line with this notion due to a rather declarative nature of the positivity effect (Lind, Visentini, Mantyla, & Del Missier, 2017; Mather & Carstensen, 2003). Accordingly, a positive bias may be considered as the baseline for learning from delayed feedback, with the PD patients' similar performance for learning from positive and negative feedback thus representing a deviation from this baseline. This, in turn, suggests that striatal DA also affects learning from delayed feedback.

In line with this, activation in the ventral striatum did not differ between learning from immediate and delayed feedback (Foerde & Shohamy, 2011a). Also, recent EEG studies suggested that comparable mechanisms underlie immediate and delayed feedback processing (Arbel, Hong, Baker, & Holroyd, 2017; Peterburs et al., 2016). These studies reported a diminished, but still present FRN during delayed feedback processing. Taken together, it is feasible that structures like the HC may cooperate with the DA system/striatum during delayed feedback processing (Dickerson & Delgado, 2015; Dickerson et al., 2011), possibly with the HC bridging the temporal gap between actions and the relative outcomes linked by the DA system.

2.4 Conclusion

In summary, the results suggest a contribution of the striatum/DA system to delayed feedback processing, possibly together with other structures. The relative tendency in PD patients to learn better from negative compared to positive feedback, which has been linked to diminished striatal DA levels, was observed for both immediate and delayed feedback. On one hand, normal overall learning performance in PD patients indicated that striatal DA seems to be involved in both immediate and delayed feedback processing. On the other hand, delayed feedback processing seems to rely less on striatal DA as transfer performance was comparable for PD patients and controls learning from delayed but not immediate feedback.

3 Study 2: Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning

Study 2 investigated electrophysiological differences between expected and unexpected positive and negative feedback following choice actions, provided either immediately or after a temporal delay. Midbrain DA signals probably code a prediction error signal that is projected to the striatum and ACC where it is reflected by the FRN. Thus, expectancy effects in the FRN during learning from delayed feedback would suggest a contribution of the DA system during delayed feedback learning. This study was published in Psychophysiology, 53(11), 1739-1750 (Weismüller & Bellebaum, 2016).

3.1 Introduction

The RL-theory (Holroyd & Coles, 2002) associates the FRN (Gehring & Willoughby, 2002; Miltner et al., 1997) with dopaminergic prediction error signals projected from the midbrain to the striatum and ACC (Ferdinand et al., 2012; Hajcak et al., 2007; Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Oliveira et al., 2007; Walsh & Anderson, 2012). In line with the RL-theory, the effect of outcome expectancy on the FRN is well established, with increased FRN amplitudes for unexpected compared to expected events (Bellebaum & Daum, 2008; Ferdinand et al., 2012; Hajcak et al., 2007; Holroyd & Krigolson, 2007; Holroyd et al., 2009; Oliveira et al., 2007; Yasuda et al., 2004).

Feedback timing appears to affect which structures are recruited during feedback processing, with the HC becoming involved when an outcome follows an action after several seconds delay (Foerde et al., 2013; Foerde & Shohamy, 2011a). However, the FRN has only been sparsely investigated when feedback was delayed, with different results (Opitz et al., 2011; Peterburs et al., 2016; J. Wang et al., 2014; Weinberg et al., 2012). Peterburs et al. (2016) demonstrated that the FRN amplitude decreases but does not completely vanish with increasing feedback delay (up to 6,500 ms). In line with this, the ventral striatum is also active when delayed feedback is processed (Foerde &

Shohamy, 2011), and the FRN has been associated with activation in the striatum and ACC (Becker et al., 2014; Foti et al., 2011). Thus, for the present study, we hypothesised that expectancy would modulate the FRN even when feedback follows an action after a temporal delay.

3.2 Method

3.2.1 Study participants

Fifty volunteers were randomly assigned to one of two groups. One group learned from immediate feedback, while the other group learned from feedback that was temporally delayed. All participants were students of the Heinrich-Heine-University Düsseldorf, aged between 18 and 38 years, and were rewarded by reimbursement or course credit additional to a monetary gain based on their performance during the experimental task.

3.2.2 Experimental task and procedure

In this experiment, participants were asked to choose one of two randomly paired symbols in order to receive feedback on a trial-to-trial basis. Six different symbols were used overall. Additionally, the participants had to indicate their subjective reward probability on a scale (ranging from 0 to 100 %) that was displayed between the symbols. Participants were instructed that only their choice of a symbol, not their subjective reward probability, would result in monetary reward (+20¢) or punishment (-10¢) in each trial. It was also made clear that the aim was to always choose the better symbol (i.e., the symbol with the higher reward probability) and thus to maximise total monetary gain. Each participant's intermediate sum score of the monetary gain was presented after each phase. The six symbols' objective reward probabilities (0%, 20%, 40%, 60%, and 80%), however, were unknown to the participants and had to be learned from the feedback the participants received for their choices.

In total, the task consisted of 500 trials separated into 5 blocks that were separated by short breaks. All 10 possible combinations of the six symbols appeared equally

often. While participants were engaged in the task, EEG was continuously recorded from 32 nose referenced electrodes.

3.2.3 Data analysis

We assessed learning performance separately for each group and each block by averaging correct choices (i.e., choices of symbols with higher reward probability) across all trails in each block. As we intended to investigate ERPs in response to expected and unexpected positive and negative feedback and since learning is a necessary prerequisite for the development of reward expectations, only the data of those participants who reached and maintained a 65% accuracy criterion was analysed.

After standard pre-processing, EEG data was segmented into four conditions relative to the combination of the chosen symbol's objective reward probability and the actually received feedback in each trial. If a participant chose a symbol with a comparably low objective reward probability, that is, the 20% or 40% symbol, and received positive feedback, this event was considered an unexpected positive feedback. A choice of a symbol with a comparably high reward probability (60% and 80%) that was not rewarded was considered an unexpected negative feedback. Likewise, unrewarded choices of symbols with a low probability resulted in expected negative feedback, while rewarded choices of symbols with a high probability resulted in expected positive feedback. The 0% symbol was not considered for analysis but only included in the task so that participants would in some trials have to choose the 20% symbol as the better of two bad options. The FRN was defined as the maximum negative peak in the negative-positive difference wave between 180 and 350 ms post-feedback at FCz.

3.3 Results and discussion

Behavioural data analysis showed that participants from the immediate learning and the delayed feedback learning group learned similarly well. Irrespective of feedback timing, participants scored higher for later than for earlier blocks, which indicates learning progress in both groups. Also, the subjective reward expectancy

approached the symbols' objective reward probabilities in both groups. Hajcak et al. (2007) postulated that the participants' own reward expectation would not necessarily reflect the stimuli's objective reward probability so that different objective reward probabilities would possibly not elicit different FRN amplitudes. Therefore, in the present study, subjective reward expectancy was assessed for symbols with high versus low objective reward probabilities, analogous to ERP analysis. This analysis showed that subjective reward expectancy was comparable between feedback timing groups, but differed for high and low reward probabilities. It is noteworthy that participants overestimated reward probabilities, particularly for symbols with low reward probabilities, which can be explained by a general overoptimistic bias in reward expectation (Miller & Ross, 1975; Oliveira et al., 2007; Radhakrishnan, Arrow, & Sniezek, 1996). Still, subjective reward expectancies were higher for high compared to low objective reward probabilities, which verifies the expectancy effect on the FRN.

The FRN was generally higher in amplitude for unexpected compared to expected feedback and for immediate compared to delayed feedback. The first result is in line with the RL-theory (Holroyd & Coles, 2002), which associates increased difference wave FRN amplitudes for unexpected events with midbrain dopaminergic prediction error signals projected to the striatum/ACC (Gehring & Willoughby, 2002; Holroyd et al., 2009; Miltner et al., 1997; Oliveira et al., 2007; Walsh & Anderson, 2012).

The latter finding corroborates results from previous studies using fMRI (Foerde & Shohamy, 2011) or EEG (Peterburs et al., 2016) that suggest diminished striatal/ACC involvement during processing of feedback that is temporally delayed relative to an action. Other EEG studies using different experimental parameters and tasks reported controversial results. While Opitz et al. (2011) reported feedback timing effects on the FRN when comparing 0 s to 1 s feedback delay, Wang et al. (2014) failed to find a difference in the FRN for immediate versus delayed feedback in a gambling task. Also using a gambling task, Weinberg et al. (2012) found an FRN for immediate, but not delayed feedback. The FRN might be more pronounced in the present study, because participants were able to use the feedback in order to learn stimulus-outcome associations and adapt their behaviour accordingly. Gambling tasks,

however, may not be suited to evoke prediction errors, as error information cannot be used for behavioural adaptation (Holroyd et al., 2009).

Importantly, the difference FRN was increased for unexpected compared to expected feedback during both learning from immediate and from delayed feedback. In accordance with the RL-theory (Holroyd & Coles, 2002), this indicates striatal/ACC activity during delayed feedback processing. At first sight, this contradicts previous findings (Foerde et al., 2013; Foerde & Shohamy, 2011a) that propose a functional dissociation of the basal ganglia and HC during feedback processing (Myers et al., 2003; Poldrack & Packard, 2003), with a more pronounced role of the HC when feedback is delayed (Foerde et al., 2013). However, Foerde et al. (2013) also report at least some striatal activity in the delayed feedback condition, which, together with the present findings, suggest striatal/ACC contributions to processing of delayed feedback. In line with this, Peterburs et al., (2016) reported a diminished, but not absent FRN when feedback is delayed, with decreasing FRN amplitudes for increasing time between action and outcome. It is thus conceivable that the striatal/ACC system and HC cooperate when feedback after an action is temporally delayed. This idea is in line with literature that proposes a cooperation of declarative and non-declarative memory systems, which have been associated to the HC and the basal ganglia, respectively (Dickerson et al., 2011; Sherry & Schacter, 1987; Squire, 1992; Squire & Zola, 1996).

As some recent studies argued that the signal in the FRN time window is mainly manipulated by positive, not negative feedback (Holroyd et al., 2008; Proudfit, 2015; Weinberg, Riesel, & Proudfit, 2014), additional analyses were conducted to investigate whether the modulation of the difference wave was driven by positive or negative feedback or both. To do so, feedback-locked ERPs following positive and negative feedback were analysed around each participant's individual FRN difference wave peak. FRN amplitudes were increased for unexpected compared to expected positive feedback, but did not differ between expected and unexpected negative feedback, which corroborates previous results (Holroyd, 2004; Holroyd et al., 2008). In line with the FRN difference wave amplitudes, this expectancy effect of positive feedback was found in the immediate and delayed feedback learning group.

3.4 Conclusion

Although the FRN was diminished relative to immediate feedback, expectancy affected it during delayed feedback learning. The current study thus provides evidence for a contribution of the mesolimbic/mesocortical DA system to the processing of feedback that follows an action after a temporal delay.

4 Study 3: Effects of feedback delay and agency on feedback-locked beta and theta power during reinforcement learning

The neuronal mechanisms seem to differ between processing of feedback that refers to a self-generated action and feedback that refers to an observed action. Moreover, the impact of feedback timing on learning from observed feedback has not yet been clarified. Study 3 investigated the combined influence of feedback timing and agency on feedback processing. This study has been submitted to Psychophysiology for revision (Weismüller, B., Kullmann, J., Hoenen, M., & Bellebaum, C., 2018).

4.1 Introduction

Action-outcome contingencies can not only be learned from own actions, but also when observed actions result in an outcome. However, the neuronal underpinnings of feedback processing differ between feedback that is relevant for one's own actions and feedback referring to somebody else's actions (Morelli, Sacchet, & Zaki, 2015). Especially the striatum seems to be differentially active during reward for a self-generated versus an observed action with a selective role of the dorsal part in processing of feedback for own actions (Bellebaum et al., 2012; Kobza & Bellebaum, 2015). Electrophysiological studies support this idea by reporting increased FRN amplitudes for feedback referring to own versus observed actions (Bellebaum & Colosio, 2014; Bellebaum, Kobza, et al., 2010). Accordingly, diminished striatal DA levels in PD patients have been shown to affect learning from feedback for self-generated actions, but not from feedback for observed actions (i.e., feedback referring to another person's action; Kobza et al., 2012). Together, striatal activity may contribute more to active than to observational feedback processing.

Feedback timing has been shown to affect neuronal activity with a striatal contribution to immediate, but not delayed feedback processing, as revealed in an fMRI study (Foerde & Shohamy, 2011a). However, recent ERP research found at least some medial-frontal activation during delayed feedback processing that was linked to striatal/ACC activity (Peterburs et al., 2016; Weismüller & Bellebaum, 2016). In line

with this, fMRI data by Foerde & Shohamy (2011) suggests some activation in the ventral striatum for delayed feedback processing.

Recent literature pointed at the benefits time-frequency analysis of EEG data that might provide information that extend ERP analyses (Cohen, 2011b). Predominantly oscillations in the theta and beta and low gamma frequency have been associated with negative and positive feedback processing, respectively (Cohen et al., 2011). Larger medial frontal power in the theta band has been reported for negative feedback (Janssen et al., 2016) as well as for negative prediction errors (Cavanagh & Frank, 2014; Cavanagh et al., 2010; Cohen, 2011a; Li, Baker, Warren, & Li, 2016). Because medial frontal oscillations in the theta band seem to underlie ERP components that have been associated with novelty detection, conflict processing, error detection, conflict processing, and error processing, it was suggested to rather reflect a general signal communicating the need for cognitive control (Cavanagh & Frank, 2014; Cohen, 2014). Positive feedback processing, however, has been associated with oscillations in the beta and low gamma range, possibly caused by DA activity in the ventral striatum and PFC (Marco-Pallares, Camara, Munte, & Rodriguez-Fornells, 2008; Marco-Pallares et al., 2015; Mas-Herrero et al., 2015). Furthermore, synchronisation in the beta band was suggested to reflect communication across distinct structures (Marco-Pallares et al., 2015). In line with studies observing beta band oscillations after actions that were associated with rewarding outcomes (Feingold, 2011), beta oscillations may reflect a motivational signal, that drives mnemonic consolidation of beneficial actions (Marco-Pallares et al., 2015).

Although both feedback agency (i.e., the feedback's relevance for one's own action) and feedback timing seem to similarly affect striatal activation, a combined influence of feedback timing and agency has not yet been investigated. It seems as if the dorsal striatum links actions to self-generated outcomes (Bellebaum et al., 2012) and to recent events (Foerde & Shohamy, 2011a), which is why it becomes less active during both observational and delayed feedback. Theta power was hypothesised for self-relevant negative feedback, as a need for cognitive controls has to be communicated only for self-generated erroneous actions. No clear hypothesis was formulated for delayed feedback. Power increases in the beta range were associated

with striato-frontal activation (Marco-Pallares et al., 2015) after actions associated with reward (Feingold, 2011; Mas-Herrero et al., 2015) so that diminished beta power was predicted for delayed and observed feedback, independently but also in for delayed observational feedback. Since the FRN has been specifically related to striatal/ACC activation, a decrease in FRN was hypothesised for both, observational and delayed feedback, independently.

4.2 Method

4.2.1 Study participants

Two groups of 20 healthy students participated in this study. In one group, participants were asked to complete an active feedback learning task, while the other group completed an observational version of the same task (see below). All participants gave informed written content and were reimbursed with 15 €.

4.2.2 Experimental task and procedure

The active group engaged in a modified version of the task used in a previous study (Weismüller & Bellebaum, 2016). In each trial, participants were asked to choose one of two symbols presented on the left and right side of the screen in order to receive positive or negative monetary feedback (+20 ¢ or -10 ¢). This feedback could be used to learn that some stimulus-locked reward probabilities were higher than others (six symbols with probabilities of 0%, 20%, 40%, 60%, and 80%), which helped participants to maximise their reward during the task.

In the observational group, participants were also presented with two symbols in each trial. However, they did not choose between them, but rather watched another participant's choice indicated by a picture of a hand pointing at one of the symbols. After the choice was made, the observing participant was asked to confirm it by pressing the respective left or right button.

Importantly, all participants completed two versions of the task. In one version, feedback was delivered immediately (after 1000 ms) after the participant had chosen

a symbol or confirmed a symbol choice by an observed person, while in the other version, feedback was provided after a temporal delay (after 7000 ms).

After each block of 100 trials with feedback, a test phase followed with 60 trials in which the same symbol pairs were presented but no feedback was given after a choice. During the test phases, participants of both the active and the observational group were asked to choose the symbols according to their learned reward probabilities. This procedure was used to assess and compare learning in active and observational learners. Also, to maintain motivation during the time course of the experiment, the performance reflected in the total monetary gain/loss was presented after every test phase.

In total, all participants (active and observational) completed three learning phases with immediate and three learning phases with delayed feedback, each followed by the respective test phase. While doing so, 28 active electrodes were used to record the EEG.

4.2.3 Data analysis

Correct choices (i.e., choices of symbols with higher reward probability) were summed across all trials in a block for the participant of each group and for each feedback timing.

For each group (the active and the observational learning group), EEG data was first down-sampled, pre-analysed, and segmented into four conditions according to the feedback timing (immediate and delayed feedback) and feedback valence (positive and negative feedback). Then, the data in each condition was transformed into time-frequency data using separate continuous complex Morlet wavelets with spectral bandwidths that were optimised for theta and beta frequency analyses. For each frequency band and each condition, total power was computed by averaging single-trial spectral power. Afterwards, evoked power (the spectral power of the feedback locked ERPs) was computed and subtracted from the total power to obtain the induced power. Mean induced spectral power in the theta frequency was extracted in the range between 4.12 and 7.75 Hz between 200 and 500 ms at electrode FCz, while mean

induced spectral power in the beta frequency was extracted between 19.62 and 35 Hz at FCz.

For ERP analysis, EEG data was analysed separately. After standard pre-processing, data was segmented and averaged relative to the type of the feedback for each group, feedback timing, and feedback valence. Then, the maximum negative peak in the negative-positive difference wave between 200 and 370 ms following the feedback stimulus at FCz defined the FRN and was extracted for statistical analysis.

4.3 Results and discussion

Behavioural analysis revealed a linear increase in learning accuracy that did not differ between active and observational groups or between feedback timings. Over the course of the experiment, all participants learned to choose the symbols with a higher reward probability irrespective of the feedback timing or its relevance for their own actions.

As hypothesised, mean power in the theta band was generally more increased for negative than positive feedback, but this difference was also modulated by feedback agency: negative feedback for one's own actions increased theta power more than positive feedback, while this was not found for feedback in observational learning. Similarly, negative immediate feedback increased theta power more than positive immediate feedback, which was not found when feedback was delayed. These findings are in line with previous research relating theta power to negative feedback processing (Janssen et al., 2016). Furthermore, the current data supports the idea that theta oscillations reflect a global top-down alarm signal when cognitive control is needed (Cavanagh & Frank, 2014; Cohen, 2014) for behavioural adaptation (Cavanagh et al., 2010). As observed actions cannot be adapted, theta power did not differ between positive and negative observed feedback.

In this context, differential theta power for positive and negative immediate and delayed feedback might appear counter-intuitive, because feedback timing does not necessarily affect the need for behavioural adjustment. In accordance with evidence from previous research, this finding may reflect the involvement of different neuronal mechanisms when delayed feedback is processed (Arbel et al., 2017; Opitz et al., 2011;

Peterburs et al., 2016; Weismüller & Bellebaum, 2016). Indeed, medial prefrontal neurons were reported to synchronise with hippocampal theta oscillations in rats (Siapas, Lubenov, & Wilson, 2005), while theta frequency synchronised over frontal and occipital electrode sites after unconscious errors in humans (Cohen, van Gaal, Ridderinkhof, & Lamme, 2009). In line with evidence suggesting a rather hippocampal than striatal contribution to delayed feedback processing (Foerde et al., 2013; Foerde & Shohamy, 2011a), the HC may communicate theta oscillatory signals during negative delayed feedback.

Beta power was generally larger for positive relative to negative feedback, confirming the relation between beta oscillations and prefrontal and ventral striatal reward processing (HajiHosseini & Holroyd, 2015; Mas-Herrero et al., 2015). Similar to theta power, feedback timing affected this difference, with larger beta power for positive immediate compared to negative immediate feedback, but no difference between positive and negative delayed feedback. Beta oscillatory activity has been linked to DA-driven memory consolidation of rewarded behaviour (Feingold, 2011; HajiHosseini & Holroyd, 2015; Mas-Herrero et al., 2015). The current data appears to support this, adding the possible interpretation of reduced striatal/prefrontal reward-related activation when feedback is delayed, in line with previous studies (Arbel et al., 2017; Foerde et al., 2013; Foerde & Shohamy, 2011a; Peterburs et al., 2016; Weismüller & Bellebaum, 2016).

Feedback agency, however, did not affect beta band oscillatory power, in contrast to theta power. This is in line with results associating beta oscillations with mnemonic consolidation of positive outcomes (Feingold, 2011), which also has to take place during observational reward processing to facilitate observational learning. Results also show that neuronal mechanisms differ between feedback processing in active and observational learning (Bellebaum et al., 2012; Kobza et al., 2012; Morelli et al., 2015), although activation may also overlap possibly in the dorsal striatum (Cooper, Dunne, Furey, & O'Doherty, 2012).

The FRN was larger for both immediate compared to delayed feedback, and for active compared to observational learning. Importantly, the FRN was also differentially affected by feedback timing and agency: it was increased for immediate

compared to delayed feedback referring to own actions, while this timing effect was not found for feedback referring to somebody else's actions. These results replicate previous studies suggesting reduced striatal/ACC activity during delayed (Peterburs et al., 2016; Weismüller & Bellebaum, 2016) and observational (Bellebaum & Colosio, 2014; Bellebaum, Kobza, et al., 2010) feedback processing. It furthermore adds to these findings the combined effect of feedback agency and timing, probably pointing to diminished striatal/ACC activation during observational feedback processing that is not further reduced by feedback delay.

4.4 Conclusion

Taken together, the time-frequency analyses of the present study revealed frequency-specific systems possibly signalling either the need for behavioural adaptation after negative events (Cavanagh & Frank, 2014; Cavanagh et al., 2010; Cohen, 2014) or a motivational signal for facilitating memory consolidation for positive events (Mas-Herrero et al., 2015). The FRN, showing a different activation pattern, seems to reflect another more specific mechanism.

5 General Discussion and Conclusion

All three studies of this dissertation investigated the influence of feedback timing on feedback learning and/or feedback processing. Behavioural and electrophysiological data was acquired from human subjects participating in forced choice tasks in which probabilistic feedback was the only information that could be used to improve performance. Importantly, the delay between the participants' actions and the related feedback was varied. The studies included in this dissertation focussed on the comparison of immediate and delayed feedback. The aim was to reveal neuronal mechanisms that underlie learning from temporally delayed feedback and thereby learning from feedback in general.

The results of Study 1 demonstrated that the tendency to learn better from positive or negative feedback differed between PD patients and healthy controls for both immediate and delayed feedback. The PD patients' bias to learn better from negative feedback was attributed to diminished striatal DA levels as these are the defining feature of PD. Because this learning bias did not differ between immediate and delayed feedback in PD patients or controls, feedback timing does not seem to modulate how striatal DA affects feedback-based learning.

Study 2 focused on electrophysiological correlates of feedback processing, that is, the FRN, which is thought to reflect DA prediction error signals that are conveyed from the midbrain to the striatum and ACC. The FRN in Study 2 was sensitive to feedback expectancy during both learning from immediate and delayed feedback. This finding suggests at least a partial striatal/ACC contribution to delayed feedback processing.

Study 3 investigated how feedback agency together with feedback timing modulate feedback processing. This study used a spectral power analysis of theta and beta frequency ranges, which revealed feedback valence-specific systems that may be involved in signalling the need for cognitive control and memory consolidation, respectively. Both measures differentiated less between feedback valences when feedback was delayed, while behavioural task performance remained comparable. In

line with Studies 1 and 2, this may hint at the involvement of different neuronal mechanisms when delayed feedback is processed.

Foerde and Shohamy (2011a) provided evidence for hippocampal rather than striatal activation during delayed feedback processing, which at first glance contradicts the present studies' findings. Furthermore, in a later study, the authors linked impaired learning performance from immediate feedback to the striatal dysfunctions in PD, while impaired learning from delayed feedback was related to hippocampal dysfunctions in amnestic patients (Foerde et al., 2013). The studies included in this dissertation draw a more complex picture than this double-dissociation, accumulating evidence for a reduced, but still present nigro-striatal DA involvement when delayed feedback is processed compared to immediate feedback. First, depleted striatal DA levels caused a negative learning bias in PD patients relative to healthy controls when feedback was delayed in Study 1, similar to the condition when feedback was given immediately (Weismüller et al., 2018). Second, in Study 2, the FRN was sensitive to feedback expectancy for immediate and delayed feedback (Weismüller & Bellebaum, 2016), which also suggests an involvement of the dopaminergic medial-frontal reward system for delayed feedback processing. Likewise, Study 3 revealed medial frontal valence-specific communication signals reflected by oscillatory activation during feedback processing mostly for immediate feedback. These signals are probably communicated along valence-specific systems involving structures that support the medial-frontal learning system during learning from delayed feedback.

5.1 Cooperating memory systems in delayed feedback learning

A candidate structure that might contribute to delayed feedback learning is the HC (Foerde et al., 2013; Foerde & Shohamy, 2011a). Foerde and Shohamy (2011a) reported activation in the HC in a feedback learning task in which the outcome was provided 7 s after the participants' action. The authors proposed that the HC might bind outcomes to the related action across time (Staresina & Davachi, 2009). The HC is commonly associated with declarative (conscious) memory (Eichenbaum, 2004; Squire, 1992; Squire & Zola, 1996). It was also suggested to underlie classical conditioning when the US was temporally delayed (Cheng, Disterhoft, Power, Ellis,

& Desmond, 2008). In contrast, non-declarative memory has been associated with the basal ganglia (Packard & Knowlton, 2002; Seger, 2006; Yin, Ostlund, & Balleine, 2008; for review, see Foerde & Shohamy, 2011b). Similarly, another model associated the striatal DA system to a habitual and the PFC and HC to a goal-directed system (Corbit & Balleine, 2000; Cosman & Vecera, 2013; Daw & Shohamy, 2008; Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995).

Myers et al. (2003) provided clinical evidence for a functional dissociation of the striatum and HC. In their study, participants with hippocampal atrophy, PD patients, and healthy controls learned two pre-defined stimulus-target associations to test "initial" associative learning. In a later phase, one antecedent stimulus was re-associated with a new target, and the participants had to learn this new association. Lastly, the participants were supposed to associate the other antecedent stimulus to a new target equivalent to the first newly learned stimulus-target pairing. This was thought to represent autonomous knowledge transfer. Interestingly, the study found a double-dissociation between striatum and HC: PD patients, but not patients with hippocampal atrophy, performed worse than controls in the initial-learning phase, while hippocampal atrophy, but not PD, caused impaired performance in the transfer phase. In line with this, previous research reported that amnestic patients were exclusively impaired in recalling tasks compared to controls, while they performed comparably to controls when learning artificial grammar (Knowlton, Ramus, & Squire, 1992). Likewise, Knowlton et al. (1996) emphasised the striatal involvement in habitual learning. The authors used a weather-prediction task in which the presence or non-presence of different cues out of four cues in total predicted an outcome, i.e., rain or sunshine. The participants' task was to associate different cue combinations with the outcome in order to predict the resulting weather. This associative learning task was thought to reflect non-declarative habitual learning, while the authors additionally tested the participants' declarative memory of the task by eight multiple-choice questions about the previously learned cues. PD patients were significantly impaired compared to controls and amnestic patients when learning to predict the weather, with worse learning performance for more severe symptoms. However, no impairments were found in the PD patients' performance in the declarative memory

task irrespective of the symptom severity, while amnestic patients were significantly impaired here. More recent behavioural findings from Foerde et al. (2013) and Foerde and Shohamy (2011a) functionally dissociated striatal and hippocampal involvement depending on the feedback timing.

Feedback timing seemed to modulate whether the striatum or the HC was recruited during the same stimulus-action-outcome learning task with a clear dissociation of the two systems (Foerde et al., 2013). The results of Study 2 in this dissertation partially support this idea indicating diminished striatal/ACC involvement for delayed feedback. Also, the results of Study 3 suggest diminished medial-frontal activation for both delayed monetary reward and delayed monetary punishment. Conversely, Study 2 also found correlates of medial-frontal prediction error coding during delayed feedback processing that indicate the involvement of the striatal DA reward learning system. This idea contradicts Foerde and Shohamy's (2011a) interpretation of dichotomous striatal and hippocampal activation, as it proposes activation in both structures rather than exclusively in one of them. However, when investigating healthy controls in their fMRI study, Foerde and Shohamy (2011a) reported activation of the ventral striatum also during delayed feedback. This restricts the dissociation of striatum and HC activation during immediate and delayed feedback processing, respectively, to the dorsal part of the striatum. In reinforcement learning, the dorsal striatum was suggested to maintain information about beneficial action-outcome-contingencies and the initiation of behavioural adaptation, while the ventral part was suggested to learn to predict the future (O'Doherty et al., 2004). In line with this, the ventral striatum was associated with prediction error coding (Pagnoni et al., 2002). In PD patients, the nigro-striatal pathway and thus the dorsal striatum is usually affected earlier and more severely compared to the ventral striatum (Fearnley & Lees, 1990; Jellinger, 1999; MacDonald et al., 2011; McRitchie, Cartwright, & Halliday, 1997). Thus it is conceivable that, depending on the severity and state of the disease, PD patients could make use of their intact HC when processing delayed feedback, while being impaired for immediate feedback (Foerde et al., 2013; Foerde & Shohamy, 2011a; Weismüller et al., 2018). On the other hand, clinical findings suggesting hippocampal involvement during delayed feedback processing do not necessarily

demonstrate that delayed feedback is exclusively processed in the HC in the healthy brain.

Considering theoretical approaches on the healthy brain, the actor-critic-model (Barto, 1995) states that the ventral striatum functions as an adaptive critic evaluating prediction error signals, while the dorsal part, the actor, learns action preferences and adjusts actions in order to maximise the outcome (Joel & Weiner, 2000; Montague, Dayan, & Sejnowski, 1996; O'Doherty et al., 2004; for reviews, see Schultz, 2002; Walsh & Anderson, 2012). In view of the present and previous findings, it may be speculative but conceivable that the HC might compensate for the dorsal striatum's role as the actor when feedback is delayed, while the ventral striatum still evaluates prediction errors (Foerde & Shohamy, 2011a). Shohamy and Adcock (2010) proposed that the HC is affected by tonic rather than phasic DA signals. Possibly, delayed feedback triggers low tonic DA changes, which then affect the HC rather than the dorsal striatum.

However, healthy participants seem to make less use of their medial-frontal reward system when processing delayed feedback, rather than not using it at all (see Study 2; Weismüller & Bellebaum, 2016). In line with these findings, Peterburs et al. (2016) suggested that neuronal systems underlying feedback learning become involved in a more-or-less rather than an all-or-nothing fashion. Also, in these studies, participants performed as well during delayed feedback as they did during immediate feedback. Although electrophysiological findings are not capable of detecting activity in sub-cortical structures (Luck, 2014), the HC is a prominently suggested structure that might become involved in learning from delayed feedback in a supportive rather than competitive manner. Cooperation between the HC and the striatum, which are commonly associated with competing memory systems (Daw & Shohamy, 2008; Myers et al., 2003; Squire, 1992; Squire & Zola, 1996), has previously been suggested to occur during feedback learning (Dickerson & Delgado, 2015; Dickerson et al., 2011). Evidence for midbrain DA projection to the HC (Gasbarri et al., 1997; Otmakhova & Lisman, 1998) that facilitate neuronal plasticity and memory formation (Shohamy & Adcock, 2010) support this notion. Prediction error signals conveyed

from the SN to the HC (Gasbarri et al., 1997) may enhance long-term consolidation mostly of reward-predicting memory (Wittmann et al., 2005).

Considering single-neuron activation patterns recorded in monkeys, midbrain DA neurons seem to increase their firing activity when delayed rewards are received compared to the receipt of an equal immediate reward (Kobayashi & Schultz, 2008). At first sight, these results contradict findings from feedback-related ERPs suggesting diminished amplitudes for delayed compared to immediate feedback processing (Peterburs et al., 2016; Weismüller & Bellebaum, 2016). Despite the ERPs' inability to reflect sub-cortical neuron activity, the FRN has been repeatedly linked to mesolimbic DA projections to the striatum and ACC (see above). Kobayashi and Schultz (2008) suggested that the monkeys were unable to link the rewards to stimuli that lay back in time (Holland, 1980) so that the reward reception was unexpected and therefore caused larger prediction errors coded midbrain DA neurons. Interestingly, however, Peterburs et al. (2016) reported larger FRN amplitudes with increasing feedback delay in the original positive and negative waveforms (labelled FRN_{peak}). These increases were thought to reflect a declarative type of feedback processing that might also involve the HC (Eichenbaum, 2004). When combining these results, it is conceivable that also in humans delayed feedback processing induces stronger midbrain DA neuron firing, which is however not measurable in the difference wave FRN. Reward prediction errors for delayed feedback are projected from midbrain DA neurons to the striatum and ACC where they are reflected in the FRN_{peak} (Cavanagh & Frank, 2014; Ferdinand et al., 2012). However, the difference wave FRN probably mostly reflects the influence of positive feedback processing (Becker et al., 2014; Holroyd et al., 2008; for review, see Proudfit, 2015). Although speculative, delayed rewards for actions may cause midbrain DA neurons to project more (but not exclusively) to the HC to facilitate adaptive memory of recently rewarded actions (Feingold, 2011; Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010). Concluding, as both the difference wave FRN and the FRN_{peak} are generated in the ACC (Holroyd & Coles, 2002), the difference wave FRN decreases with increasing feedback delay despite possibly stronger midbrain neuron activation, while the FRN_{peak} increases (Peterburs et al., 2016).

Evidence from frequency analyses revealed cortical networks that possibly involve the MTL. Medial prefrontal neurons phase locked to hippocampal theta oscillations in rats (Siapas et al., 2005), which possibly reflects inter-cortical communication (Benchenane et al., 2010; Buzsaki & Draguhn, 2004). Likewise, theta synchrony was enhanced between FCz and lateral frontal (Cavanagh, Cohen, & Allen, 2009) as well as occipital electrode sites (Cohen et al., 2009) in human participants during error processing. More specifically, Cohen (2011a) found indicators for inter-cortical communication reflected by error-related theta band synchronisation. Similar to the theta band, power increases in the beta frequency range were also suggested to reflect communicational signals (Marco-Pallares et al., 2015). Beta power was interpreted to reflect DA projections from the VTA to the ventral striatum (Mas-Herrero et al., 2015) and PFC (HajiHosseini & Holroyd, 2015) that possibly convey a motivational signal in order to consolidate memory for a beneficial stimulus-action-outcome association (Feingold, 2011; Marco-Pallares et al., 2015; Mas-Herrero et al., 2015).

Merging results from both frequency bands, Study 3 found supportive evidence for medial-frontal error processing reflected by increased theta power and medial-frontal reward processing reflected by increased beta power. Both frequency bands were less distinguished in power for positive and negative outcomes when feedback was delayed. Accordingly, it is conceivable that other structures become involved in the processing of feedback after a temporal delay that are not reflected by medial-frontal theta/beta activity. In support of this notion, Arbel et al. (2017) recently also reported decreased FRN amplitudes during delayed compared to immediate feedback processing, and increased parietal N170 amplitude for delayed compared to immediate feedback. Interestingly, the N170 has been linked to MTL activation (Grippo, Pelosi, Mehta, & Blumhardt, 1996). However, oscillations in the theta band were also affected by feedback agency in Study 3. In line with this, the FRN was affected by the combinational influence of feedback agency and delay. Kobza et al. (2012) showed that neuronal mechanisms differ between active and observational feedback processing, with probably less striatal involvement in processing observational compared to active feedback. This may explain why feedback delay did not further diminish the FRN for observational feedback in Study 3, as the FRN has been linked

to striatal and ACC activity. On the other hand, considering the results by Kobza et al. (2012), neuronal mechanisms inducing oscillations in the beta and theta band are differently affected by feedback agency and may thus overlap for feedback processing to some extent.

It has to be kept in mind, however, that the current studies are methodologically restricted, which is why inferences to specific sub-cortical activity in one or the other structure are speculative and need empirical verification. In summary, the results of Study 3 support the notion of collaborating medial-frontal and medial-temporal structures, previously associated with distinct memory systems in feedback processing. Importantly, these systems do not seem to compete with each other, but overlap and become more or less active depending on the feedback's timing (and, in case of negative feedback, its self-relevance).

5.2 Limitations and alternative interpretations

The FRN has been associated with non-declarative feedback processing in a large body of literature. However, the FRN has, at least partially, also been linked to declarative learning. For instance, Arbel, Goforth, and Donchin (2013) demonstrated that the FRN predicted the successful recall of associations between stimulus pairs. Their results conjoin with an fMRI study by Tricomi and Fiez (2012), who found increased striatal activity in a declarative learning task, but only for memories linked to rewards, which suggests that the striatum may access and strengthen rewarded action-outcome associations in the declarative memory. Following this interpretation, the FRN may at least partially reflect declarative memory processing, because it has been linked to striatal and ACC activity (Foti et al., 2011). A similar idea was formulated by Peterburs et al. (2016), who found increasing FRN_{peak} amplitudes with increasing feedback delay. The authors argued that the FRN_{peak} as opposed to the FRN in the difference signal, may reflect the violation of outcome expectations that have become more explicit and thus declarative. This explanation is in line with increasing involvement of declarative structures such as the HC (Eichenbaum, 2004) when delayed feedback is processed, which is potentially processed more explicitly.

Interestingly, in the study by Peterburs et al. (2016), the FRN was differently affected by feedback timing depending on whether its amplitude was extracted from the negative-positive difference wave or from the original waveforms. This underscores the necessity of methodological consideration for measurement of the FRN (see above). Peterburs et al. (2016) pointed out that the two differently extracted FRN amplitudes seem to reflect different processes with a possible link of the FRN_{peak} to explicit information processing.

An important methodological consideration in this dissertation refers to a difference in the paradigms between our EEG studies (i.e., Study 2 and Study 3) and the studies by Foerde et al. (2013) and Foerde and Shohamy (2011a). While Foerde et al. (2013) and Foerde and Shohamy (2011a) visually presented the participant's choices during the whole time between the action and the consequential outcome, the choice was only presented for 500 ms in our paradigm and then cleared off the screen. This way, we could confidently avoid sensory processing in the pre-stimulus interval, which we could thus consider as a neutral baseline. Although just a small methodological change, this may significantly increase the participant's working memory demands. Higher working memory demands could possibly have caused diminished FRN amplitudes for delayed feedback (see Opitz et al., 2011). Although we considered this unlikely as learning performance did not differ between feedback timings in either study, an effect of working memory cannot completely be excluded. However, higher working memory demands cannot be held responsible for FRN expectancy effects in delayed learning, which is the major finding of Study 1.

Another important criticism refers to the general theoretical link of the FRN to phasic striatal DA neuron activity. To date, no direct evidence has been provided that explicitly linked the FRN to nigro-striatal DA. Rather, Holroyd and Coles (2002) theoretically linked phasic DA activity in monkeys (Schultz, 1997, 1998, 2002; Schultz et al., 1993; Schultz et al., 1997) to electrophysiological findings in humans (Gehring & Willoughby, 2002; Miltner et al., 1997). The FRN has yet only been shown to be unspecifically affected by dopaminergic genotypes (Heitland et al., 2012; Mueller et al., 2014). The ERN that is probably generated by the same mechanisms as the FRN, was shown affected by the DA precursors L-3,4-dihydroxyphenylalanin (L-

DOPA) in PD patients (Seer et al., 2017; Volpato et al., 2016). A direct link between the FRN and DA remains elusive, however. Apart from these general considerations, some study-specific limitations have to be discussed.

5.2.1 Limitations of Study 1

Study 1 investigated valence-specific learning in PD patients and healthy controls by comparing their choice performances for novel combinations of symbols whose reward probabilities were previously learned. Smittenaar et al. (2012) and Shiner et al. (2012) related DA effects in this task to action selection rather than to feedback learning. Although, this interpretation finds some theoretical support (see Guitart-Masip et al., 2012; Guitart-Masip, Duzel, Dolan, & Dayan, 2014), DA has repeatedly been linked to learning instead of action selection in humans (Frank et al., 2007; Frank et al., 2004) and animals (Schultz, 2002; Schultz et al., 1993; Schultz et al., 1997). Of course, it is difficult to disentangle learning and action selection in a within-group design due to the medication's decomposition time. Possibly, deep-brain stimulated PD patients could represent a promising participant sample for such an investigation. On the other hand, studies on the effect of deep-brain stimulation on feedback learning revealed a different learning pattern compared to the effect of DA medication (Frank et al., 2007).

To counter the effect of DA depletion in the basal ganglia, PD is usually treated with DA-receptor agonists or DA precursors such as L-DOPA often together with catechol-O-methyltransferase (COMT) blockers (Hacke, 2010). PD patients' medication is highly individual and differs strongly in its combination due to the severity of specific symptoms or due to the agent's ability to cross the blood-brain-barrier. Also, the medications' half-lives differ, with decomposition times of up to 24 hours for long lasting medication. As this medication is usually consumed once per day in the mornings, we tested all participants in the morning before they received their usual medication.

Aside from DA, other neurotransmitter systems are also implicated in PD, as for example the nucleus basalis meynert and the nucleus coeruleus are affected (Hacke, 2010). Thus, depression is a common co-morbidity of PD that may also cause a

negative learning bias due to a generally increased cognitive and attentional focus on negative events. To control for a possible confounding effect of depression on learning behaviour, we screened for and excluded patients with comorbid depression in Study 1.

Lastly, Study 1 built its argumentation on a null-result, namely on a non-significant three-way-interaction between the groups (PD patients vs. healthy controls), the feedback timing (immediate vs. delayed feedback), and the learning type (better learning from positive or negative feedback). In order to ascertain that we did not commit a β-error (by falsely accepting the null-hypothesis), we computed the statistical power of this null-result. We also examined all existing studies that used the same paradigm (and thus cite Frank et al. (2004)) and found a comparable, but significant three-way-interaction. We then used the effect size that was previously needed to receive a statistically significant result in a comparable three-way-interaction together with the current sample size to compute the statistical power. As this was found to be considerably high, we felt save to draw conclusions from the null-result. However, despite the large power, a significant three-way interaction with more participants is plausible as the result patterns descriptively differed between feedback timings. To rule out this possibility, following studies could compare larger patient samples in a comparable study.

Furthermore, we used Bayesian modelling approach to additionally corroborate the conclusion that only the DA level (represented by the comparison of PD patients and healthy controls) but not the feedback timing influenced the bias to learn from positive versus negative feedback (Wagenmakers et al., 2017). In the Bayesian analysis the predictive adequacy of four statistical models were compared to the null-hypothesis. Each model assumed a different main effect or their interaction to cause the strongest change in the difference between choices of symbol A and avoidances of symbol B. The strongest evidence for a predictive influence of the current data was provided by the model assuming a main effect of DA levels, which further confirmed our interpretation.

5.2.2 Limitations of Study 2

An important critical point in this study concerns the suitability of using the difference wave. As mentioned above, the FRN has been considered as reflecting two different components, with unexpected positive feedback causing a positive deflection in the feedback-locked ERP, while generally unexpected task-relevant feedback causes a negative deflection. An ERP difference signal can be interpreted only if the original signals differ in only one affecting parameter (Luck, 2014). In Study 2, the original FRN signals for positive and negative were compared within the same expectancy condition so that the only difference between the original waves was valence, rendering the difference wave approach appropriate.

In a similar study, Arbel et al. (2017) failed to find an effect of feedback valence for delayed feedback, which they explained by the declarative nature of their word-object association learning task. Although this is a plausible explanation, also the subjectivity of the participants' feedback expectation could hold for it (Hajcak et al., 2007). In Study 2, the participants had to explicitly state their subjective reward expectation, which enabled us to directly verify their reward prediction and prediction errors. Furthermore, this method controlled for a subjective participation in the task that may have maximised the FRN for prediction errors (Hajcak et al., 2007).

5.2.3 Limitations of Study 3

Although the electrophysiological oscillations in the human brain can hint at neuronal communication within a functional network during learning (Cohen et al., 2011), the current frequency analysis does not fully exhaust its options. As discussed by Cohen (2011b), a variety of information is included in oscillatory fluctuations of neuronal excitability such as the time of the onset and end, the frequency, the power, and the phase. This variety of information is one of the aspects, rendering a frequency analysis a valuable tool that possibly extends the ERP analysis according to Cohen (2011b). The author argues that these aspects all rely on the information that is missed in ERPs, i.e., the fluctuating excitability of underlying neuronal populations. Phase-locked analyses use oscillatory synchronisations across distinct neuronal structures to

infer functional connectivity (Buzsaki & Draguhn, 2004; Cavanagh et al., 2009; Cohen, 2011a).

Although frequency analyses have been found to provide more information that classical EPR analyses in terms of information gain, ERP analyses are more commonly used in psychophysiological research. Therefore, much more is known and theorised about modulators of certain ERP components resulting in a vast literature and thus great comparability of findings (for example see Luck, 2014). In contrast, literature on frequency analyses are sparse and vary methodically so that results may differ due to methodological aspects rather than the involvement of different underlying neural mechanisms. As we have shown, the mechanisms eliciting feedback-locked ERP components such as the FRN and oscillatory activity in the theta and beta band seem to differ, but are to some extent similarly affected by feedback timing and agency. Also, oscillatory activation may underlie several ERP components as Cavanagh and Frank (2014) pointed out. This emphasises the necessity of further research on the differentiation and attribution of oscillatory activation to functional processes.

5.3 Implications for future research

As outlined above, investigations of phase-locking across distinct neuronal structures during reward processing could reveal cortical communication (Cavanagh et al., 2009; Cohen, 2011b; Cohen et al., 2011). This could, of course, further corroborate the idea that medial-frontal structures communicate with other structures during feedback processing (Siapas et al., 2005), possibly as prediction error signals are conveyed from the midbrain to the HC (Gasbarri et al., 1997). A more precise spatial localisation of neuronal activity can be achieved by measuring changes in the BOLD signal measured by fMRI during feedback processing. This can even be done simultaneously with EEG (Becker et al., 2014) or separately (Carlson et al., 2011). Also, positron-emission-tomography (PET) has been used to identify and localise dopaminergic activity in the human brain (Volkow et al., 1996) and could detect possible DA projections to the HC during feedback processing. Future investigations could compare participants receiving immediate and delayed feedback in a learning task, while they are scanned by fMRI or PET.

Another approach that is possibly easier to accomplish compared to fMRI and PET scans due to practical and economic reasons, is to test declarative memory for stimuli that were previously presented together with immediate and delayed feedback without any relation to the feedback. It is conceivable that stimuli paired with delayed feedback are memorised better than those compared to immediate feedback due to a higher activation of the HC when delayed feedback is processed. However, in a very recent study investigating this idea, Höltje and Mecklinger (2018) found out that subsequent memory performance for stimuli that had previously been linked to feedback was not affected by the feedback's timing. Interestingly, however, when the authors investigated ERPs during subsequent recall, they found that successful remembering was associated with a positive going medial-frontal deflection in the FRN time window that was larger for stimuli locked to positive, not negative feedback. In line with the studies in this dissertation, the authors suggested that positive feedback processing and memory encoding work in parallel. This idea is supported by hippocampal activation during feedback learning in humans (Foerde & Shohamy, 2011a) and striatal-hippocampal synchronisation for positive feedback in rats (Feingold, 2011). It does therefore seem promising to combine non-declarative feedback learning tasks with declarative subsequent encoding tasks in combination with EEG and other methods (see above) that may provide more insights into the involved mechanisms.

It would furthermore be interesting to investigate longer temporal gaps between an action and the consequential outcome. The studies presented here may only have made the first move towards the investigation of the neuronal mechanisms during the processing of delayed feedback. Feedback in everyday life may also follow the performance after several hours, days, or even longer. For example, pupils receive feedback for a test after several weeks. As shown here, neuronal mechanisms seem to differ already after a few seconds, although humans are still able to use these long delayed feedbacks for learning. In line with this, Peterburs et al. (2016) suggested a gradual shift of neuronal activity from the striatum/ACC to other structures, possibly the MTL (Arbel et al., 2017). It is thus conceivable, that longer feedback delays may even increase the differences in the underlying neuronal mechanisms and that the medial-frontal activation vanishes after enough time has passed. Future research may

thus investigate, how feedback is processed that refers to performances that lay back in time more than only a few seconds, although this might be less practicably.

Lastly, in the studies included in this dissertation, different learning behaviours and electrophysiological deflections were always linked to DA. However, more recently, serotonin has been found to affect learning as well as prefrontal activity (for review, see Ullsperger et al., 2014). Future research could therefore either try to isolate DA effects, or investigate the effect of serotonin on the FRN during feedback learning.

To conclude, the studies in this dissertation provide accumulating evidence for a reduced but still present striatal/ACC activation when feedback is processed that is delayed relative to an action by 7 s. These findings hint at the collaborative involvement of other structures, possibly the HC, because behavioural learning performance was not affected by the delay. In the modern environment humans face every day, feedback does not always follow an action immediately, but temporally delayed by a few seconds. Still, humans are able to learn and adapt their behaviour in order to cope with their increasingly complex environment. This emphasises the need to investigate the neuronal underpinnings of delayed feedback processing, which to this date is sparse. The insights gained from the studies presented in this dissertation thus contribute to an intensive ongoing discussion about the neuronal underpinnings of reward learning in cognitive neuroscience.

6 References

Alexander, W. H., & Brown, J. W. (2010). Computational models of performance monitoring and cognitive control. *Top Cogn Sci, 2*(4), 658-677. doi:10.1111/j.1756-8765.2010.01085.x

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14*(10), 1338-1344. doi:10.1038/nn.2921

Amiez, C., Joseph, J. P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *Eur J Neurosci, 21*(12), 3447-3452. doi:10.1111/j.1460-9568.2005.04170.x

Amiez, C., Joseph, J. P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cereb Cortex, 16*(7), 1040-1055. doi:10.1093/cercor/bhj046

Arbel, Y., Goforth, K., & Donchin, E. (2013). The Good, the Bad, or the Useful? The Examination of the Relationship between the Feedback-related Negativity (FRN) and Long-term Learning Outcomes. *Journal of Cognitive Neuroscience, 25*(8), 1249-1260. doi:Doi 10.1162/Jocn_a_00385

Arbel, Y., Hong, L., Baker, T. E., & Holroyd, C. B. (2017). It's all about timing: An electrophysiological examination of feedback-based learning with immediate and delayed feedback. *Neuropsychologia, 99*, 179-186. doi:10.1016/j.neuropsychologia.2017.03.003

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends Cogn Sci, 9*(2), 83-89. doi:10.1016/j.tics.2004.12.003

Baker, T. E., & Holroyd, C. B. (2009). Which way do I go? Neural activation in response to feedback and spatial processing in a virtual T-maze. *Cereb Cortex, 19*(8), 1708-1722. doi:10.1093/cercor/bhn223

Barto, A. G. (1995). Adaptive Critics and the Basal Ganglia. In J. C. Houk, J. Davis, & D. Beiser (Eds.), *Models of Information Processing in the Basal Ganglia* (pp. 215-232). Cambridge, MA: MIT Press.

Becker, M. P., Nitsch, A. M., Miltner, W. H., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience, 34*(8), 3005-3012. doi:10.1523/JNEUROSCI.3684-13.2014

Bédard, P., Larochelle, L., Parent, A., & Poirier, L. J. (1969). The nigrostriatal pathway: a correlative study based on neuroanatomical and neurochemical criteria in the cat and the monkey. *Exp Neurol, 25*(3), 365-377.

Bellebaum, C., & Colosio, M. (2014). From feedback- to response-based performance monitoring in active and observational learning. *J Cogn Neurosci, 26*(9), 2111-2127. doi:10.1162/jocn_a_00612

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *Eur J Neurosci, 27*(7), 1823-1835. doi:10.1111/j.1460-9568.2008.06138.x

Bellebaum, C., Jokisch, D., Gizewski, E. R., Forsting, M., & Daum, I. (2012). The neural coding of expected and unexpected monetary performance outcomes: dissociations between active and observational learning. *Behav Brain Res, 227*(1), 241-251. doi:10.1016/j.bbr.2011.10.042

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cereb Cortex, 20*(12), 2874-2883. doi:10.1093/cercor/bhq038

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: the feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia, 48*(11), 3343-3350. doi:10.1016/j.neuropsychologia.2010.07.023

Bellebaum, C., Rustemeier, M., & Daum, I. (2011). Positivity effect in healthy aging in observational but not active feedback-learning. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 19*(3), 402-420. doi:10.1080/13825585.2011.629289

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron, 66*(6), 921-936. doi:10.1016/j.neuron.2010.05.013

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci, 8*(12), 539-546. doi:10.1016/j.tics.2004.10.003

Bromberg-Martin, E. S., Matsumoto, M., Hong, S., & Hikosaka, O. (2010). A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J Neurophysiol, 104*(2), 1068-1076. doi:10.1152/jn.00158.2010

Buzsaki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304*(5679), 1926-1929. doi:10.1126/science.1099745

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews, 26*(3), 321-352. doi:https://doi.org/10.1016/S0149-7634(02)00007-6

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage, 57*(4), 1608-1616. doi:10.1016/j.neuroimage.2011.05.037

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior Cingulate Cortex, Error Detection, and the Online Monitoring

of Performance. *Science, 280*(5364), 747-749. doi:10.1126/science.280.5364.747

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience, 29*(1), 98-105. doi:10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cereb Cortex, 22*(11), 2575-2586. doi:10.1093/cercor/bhr332

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci, 18*(8), 414-421. doi:10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage, 49*(4), 3198-3209. doi:10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *J Physiol Paris, 109*(1-3), 3-15. doi:10.1016/j.jphysparis.2014.04.003

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49*(2), 220-238. doi:10.1111/j.1469-8986.2011.01293.x

Cheng, D. T., Disterhoft, J. F., Power, J. M., Ellis, D. A., & Desmond, J. E. (2008). Neural substrates underlying human delay and trace eyeblink conditioning. *Proceedings of the National Academy of Sciences, 105*(23), 8108-8113. doi:10.1073/pnas.0800374105

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience, 29*(39), 12315-12320. doi:10.1523/JNEUROSCI.2575-09.2009

Christoph, G. R., Leonzio, R. J., & Wilcox, K. S. (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *Journal of Neuroscience, 6*(3), 613-619.

Cohen, M. X. (2011a). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage, 55*(3), 1373-1383. doi:10.1016/j.neuroimage.2010.12.072

Cohen, M. X. (2011b). It's about Time. *Frontiers in Human Neuroscience, 5*, 2. doi:10.3389/fnhum.2011.00002

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends Neurosci, 37*(9), 480-490. doi:10.1016/j.tins.2014.06.004

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., & Lamme, V. A. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Frontiers in Human Neuroscience, 3*, 54. doi:10.3389/neuro.09.054.2009

Cohen, M. X., Wilmes, K., & Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends Cogn Sci, 15*(12), 558-566. doi:10.1016/j.tics.2011.10.004

Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia, 44*(10), 1663-1673. doi:10.1016/j.neuropsychologia.2006.03.030

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex, 11*(12), 1136-1143.

Cools, R., Barker, R. A., Sahakian, B. J., & Robbins, T. W. (2003). L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia, 41*(11), 1431-1441.

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *J Cogn Neurosci, 24*(1), 106-118. doi:10.1162/jocn_a_00114

Corbit, L. H., & Balleine, B. W. (2000). The role of the hippocampus in instrumental conditioning. *Journal of Neuroscience, 20*(11), 4233-4239

Cosman, J. D., & Vecera, S. P. (2013). Learned control over distraction is disrupted in amnesia. *Psychol Sci, 24*(8), 1585-1590. doi:10.1177/0956797613475632

Cummings, J. L. (1993). Frontal-subcortical circuits and human behavior. *Arch Neurol,* 50(8), 873-880.

Damasio, A. R. (1994). Descartes' error: Emotion, rationality and the human brain. *New York: Putnam,* 352.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology, 2*, 233. doi:10.3389/fpsyg.2011.00233

Daw, N. D., & Shohamy, D. (2008). The Cognitive Neuroscience of Motivation and Learning. *Social Cognition, 26*(5), 593-620.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann N Y Acad Sci, 1104*, 70-88. doi:10.1196/annals.1390.002

Delgado, M. R., & Dickerson, K. C. (2012). Reward-related learning via multiple memory systems. *Biol Psychiatry, 72*(2), 134-141. doi:10.1016/j.biopsych.2012.01.023

di Pellegrino, G., Ciaramelli, E., & Ladavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *J Cogn Neurosci, 19*(2), 275-286. doi:10.1162/jocn.2007.19.2.275

Dickerson, K. C., & Delgado, M. R. (2015). Contributions of the hippocampus to feedback learning. *Cogn Affect Behav Neurosci, 15*(4), 861-877. doi:10.3758/s13415-015-0364-5

Dickerson, K. C., Li, J., & Delgado, M. R. (2011). Parallel contributions of distinct human memory systems during probabilistic learning. *Neuroimage, 55*(1), 266-276. doi:10.1016/j.neuroimage.2010.10.080

Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior, 23*(2), 197-206.

Dubois, B., & Pillon, B. (1996). Cognitive deficits in Parkinson's disease. *J Neurol, 244*(1), 2-8. doi:10.1007/pl00007725

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology, 14*(5), 456-467.

Eichenbaum, H. (2004). Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron, 44*(1), 109-120. doi:10.1016/j.neuron.2004.08.028

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia, 46*(2), 521-539. doi:https://doi.org/10.1016/j.neuropsychologia.2007.09.001

Falkenstein, M., Hielscher, H., Dziobek, I., Schwarzenau, P., Hoormann, J., Sunderman, B., & Hohnsbein, J. (2001). Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport, 12*(1), 157-161.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78*(6), 447-455.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biol Psychol, 51*(2-3), 87-107.

Fearnley, J. M., & Lees, A. J. (1990). Striatonigral degeneration. A clinicopathological study. *Brain, 113 (Pt 6)*, 1823-1842.

Feingold, J. (2011). *Beta oscillations in frontal cortex and striatum represent post-processing of successful behavior.* Massachusetts Institute of Technology.

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *Journal of Neuroscience, 32*(35), 12087-12092. doi:10.1523/JNEUROSCI.1410-12.2012

Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience, 11*(8), 966-973. doi:10.1038/nn.2159

Foerde, K., Race, E., Verfaellie, M., & Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: evidence from amnesia. *Journal of Neuroscience, 33*(13), 5698-5704. doi:10.1523/JNEUROSCI.5217-12.2013

Foerde, K., & Shohamy, D. (2011a). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience, 31*(37), 13157-13167. doi:10.1523/JNEUROSCI.2701-11.2011

Foerde, K., & Shohamy, D. (2011b). The role of the basal ganglia in learning and memory: insight from Parkinson's disease. *Neurobiol Learn Mem, 96*(4), 624-636. doi:10.1016/j.nlm.2011.08.006

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170. doi:10.1111/j.1469-8986.2007.00602.x

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res, 12*(3), 189-198.

Forster, S. E., & Cho, R. Y. (2014). Context specificity of post-error and post-conflict cognitive control adjustments. *Plos One, 9*(3), e90281. doi:10.1371/journal.pone.0090281

Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clin Neurophysiol, 126*(7), 1338-1347. doi:10.1016/j.clinph.2014.08.025

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. *Hum Brain Mapp, 32*(12), 2207-2216. doi:10.1002/hbm.21182

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci, 17*(1), 51-72. doi:10.1162/0898929052880093

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science, 318*(5854), 1309-1312. doi:10.1126/science.1146157

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. doi:10.1126/science.1102941

Gasbarri, A., Sulli, A., & Packard, M. G. (1997). The dopaminergic mesencephalic projections to the hippocampal formation in the rat. *Prog Neuropsychopharmacol Biol Psychiatry, 21*(1), 1-22.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychol Sci, 4*(6), 385-390. doi:10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279-2282. doi:10.1126/science.1066893

Gerfen, C. R. (1992). The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu Rev Neurosci, 15*, 285-320. doi:10.1146/annurev.ne.15.030192.001441

Goetz, C. G., LeWitt, P. A., & Weidenman, M. (2003). Standardized training tools for the UPDRS activities of daily living scale: newly available teaching program. *Mov Disord, 18*(12), 1455-1458. doi:10.1002/mds.10591

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., . . . Movement Disorder Society, U. R. T. F. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord, 23*(15), 2129-2170. doi:10.1002/mds.22340

Graef, S., Biele, G., Krugel, L. K., Marzinzik, F., Wahl, M., Wotka, J., . . . Heekeren, H. R. (2010). Differential influence of levodopa on reward-based learning in Parkinson's disease. *Frontiers in Human Neuroscience, 4*, 169. doi:10.3389/fnhum.2010.00169

Grippo, A., Pelosi, L., Mehta, V., & Blumhardt, L. (1996). Working memory in temporal lobe epilepsy: an event-related potential study. *Electroencephalography and Clinical Neurophysiology, 99*(3), 200-213.

Gruendler, T. O., Ullsperger, M., & Huster, R. J. (2011). Event-related potential correlates of performance-monitoring in a lateralized time-estimation task. *Plos One, 6*(10), e25591. doi:10.1371/journal.pone.0025591

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Action controls dopaminergic enhancement of reward representations. *Proc Natl Acad Sci U S A, 109*(19), 7511-7516. doi:10.1073/pnas.1202229109

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends Cogn Sci, 18*(4), 194-202. doi:10.1016/j.tics.2014.01.003

Haber, S. N., & Fudge, J. L. (1997). The primate substantia nigra and VTA: integrative circuitry and function. *Crit Rev Neurobiol, 11*(4), 323-342.

Hacke, W. (2010). *Neurologie*: Springer.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biol Psychol, 71*(2), 148-154. doi:10.1016/j.biopsycho.2005.04.001

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*(6), 905-912. doi:10.1111/j.1469-8986.2007.00567.x

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

HajiHosseini, A., & Holroyd, C. B. (2015). Reward feedback stimuli elicit high-beta EEG oscillations in human dorsolateral prefrontal cortex. *Sci Rep, 5*, 13021. doi:10.1038/srep13021

HajiHosseini, A., Rodriguez-Fornells, A., & Marco-Pallares, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *Neuroimage, 60*(3), 1678-1685. doi:10.1016/j.neuroimage.2012.01.125

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck depressions-inventar (BDI-II)*: Harcourt Test Services Frankfurt.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward

prediction errors driving adjustment in behavior. *Journal of Neuroscience, 31*(11), 4178-4187. doi:10.1523/JNEUROSCI.4652-10.2011

Heitland, I., Oosting, R. S., Baas, J. M., Massar, S. A., Kenemans, J. L., & Bocker, K. B. (2012). Genetic polymorphisms of the dopamine and serotonin systems modulate the neurophysiological response to feedback and risk taking in healthy humans. *Cogn Affect Behav Neurosci, 12*(4), 678-691. doi:10.3758/s13415-012-0108-8

Hernandez-Lopez, S., Bargas, J., Surmeier, D. J., Reyes, A., & Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *Journal of Neuroscience, 17*(9), 3334-3342.

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: onset, progression and mortality. *Neurology, 17*(5), 427-442.

Holland, P. C. (1980). CS-US interval as a determinant of the form of Pavlovian appetitive conditioned responses. *J Exp Psychol Anim Behav Process, 6*(2), 155-174.

Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. *Neurophysiology, 78*, 447-455.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev, 109*(4), 679-709.

Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neurosci Lett, 242*(2), 65-68

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Res, 1105*(1), 93-101. doi:10.1016/j.brainres.2005.12.015

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology, 44*(6), 913-917. doi:10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cogn Affect Behav Neurosci, 9*(1), 59-70. doi:10.3758/CABN.9.1.59

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport, 14*(18), 2481-2484. doi:10.1097/01.wnr.0000099601.41403.a5

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*(5), 688-697. doi:10.1111/j.1469-8986.2008.00668.x

Höltje, G., & Mecklinger, A. (2018). Electrophysiological Reward Signals Predict Episodic Memory for Immediate and Delayed Positive Feedback Events. *Brain Res*. doi:https://doi.org/10.1016/j.brainres.2018.07.011

Hong, S., Jhou, T. C., Smith, M., Saleem, K. S., & Hikosaka, O. (2011). Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *Journal of Neuroscience, 31*(32), 11457-11471. doi:10.1523/JNEUROSCI.1384-11.2011

Jahn, A., Nee, D. E., Alexander, W. H., & Brown, J. W. (2014). Distinct regions of anterior cingulate cortex signal prediction and outcome evaluation. *Neuroimage, 95*, 80-89. doi:10.1016/j.neuroimage.2014.03.050

Janssen, D. J., Poljac, E., & Bekkering, H. (2016). Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. *Soc Cogn Affect Neurosci, 11*(8), 1280-1289. doi:10.1093/scan/nsw033

Jellinger, K. A. (1999). Post mortem studies in Parkinson's disease--is it possible to detect brain areas for specific symptoms? *J Neural Transm Suppl, 56*, 1-29.

Ji, H., & Shepard, P. D. (2007). Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *Journal of Neuroscience, 27*(26), 6923-6930. doi:10.1523/JNEUROSCI.0958-07.2007

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *Journal of Neuroscience, 31*(5), 1606-1613. doi:10.1523/JNEUROSCI.3904-10.2011

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *Journal of Neuroscience, 29*(23), 7489-7496. doi:10.1523/JNEUROSCI.0349-09.2009

Joel, D., & Weiner, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience, 96*(3), 451-474.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Biochemistry, D. o., Jessell, M. B. T., Siegelbaum, S., & Hudspeth, A. (2000). *Principles of neural science* (Vol. 4): McGraw-hill New York.

King, J. A., Korb, F. M., von Cramon, D. Y., & Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *Journal of Neuroscience, 30*(38), 12759-12769. doi:10.1523/JNEUROSCI.3274-10.2010

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399-1402.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact Artificial Grammar Learning in Amnesia: Dissociation of Classification Learning and Explicit Memory for Specific Instances. *Psychol Sci, 3*(3), 172-179. doi:10.1111/j.1467-9280.1992.tb00021.x

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr Opin Neurol, 18*(4), 411-417.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport, 12*(17), 3683-3687.

Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *Journal of Neuroscience, 28*(31), 7837-7846. doi:10.1523/JNEUROSCI.1600-08.2008

Kobza, S., & Bellebaum, C. (2015). Processing of action- but not stimulus-related prediction errors differs between active and observational feedback learning. *Neuropsychologia, 66*, 75-87. doi:10.1016/j.neuropsychologia.2014.10.036

Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., Sudmeyer, M., & Bellebaum, C. (2012). Dissociation between active and observational learning from positive and negative feedback in Parkinsonism. *Plos One, 7*(11), e50250. doi:10.1371/journal.pone.0050250

Kobza, S., Thoma, P., Daum, I., & Bellebaum, C. (2011). The feedback-related negativity is modulated by feedback probability in observational learning. *Behav Brain Res, 225*(2), 396-404. doi:10.1016/j.bbr.2011.07.059

Kujawa, A., Smith, E., Luhmann, C., & Hajcak, G. (2013). The feedback negativity reflects favorable compared to nonfavorable outcomes based on global, not local, alternatives. *Psychophysiology, 50*(2), 134-138. doi:10.1111/psyp.12002

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences, 111*(6), 2343-2348. doi:10.1073/pnas.1321596111

Lavoie, B., Smith, Y., & Parent, A. (1989). Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase

immunohistochemistry. *J Comp Neurol, 289*(1), 36-52. doi:10.1002/cne.902890104

Li, P., Baker, T. E., Warren, C., & Li, H. (2016). Oscillatory profiles of positive, negative and neutral feedback stimuli during adaptive decision making. *Int J Psychophysiol, 107*, 37-43. doi:10.1016/j.ijpsycho.2016.06.018

Lind, M., Visentini, M., Mantyla, T., & Del Missier, F. (2017). Choice-Supportive Misremembering: A New Taxonomy and Review. *Frontiers in Psychology, 8*, 2062. doi:10.3389/fpsyg.2017.02062

Luck, S. J. (2014). *An introduction to the event-related potential technique*: MIT press.

Lynd-Balta, E., & Haber, S. N. (1994). The organization of midbrain projections to the ventral striatum in the primate. *Neuroscience, 59*(3), 609-623

MacDonald, P. A., MacDonald, A. A., Seergobin, K. N., Tamjeedi, R., Ganjavi, H., Provost, J. S., & Monchi, O. (2011). The effect of dopamine therapy on ventral and dorsal striatum-mediated cognition in Parkinson's disease: support from functional MRI. *Brain, 134*(Pt 5), 1447-1463. doi:10.1093/brain/awr075

Marco-Pallares, J., Camara, E., Munte, T. F., & Rodriguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *J Cogn Neurosci, 20*(9), 1595-1610. doi:10.1162/jocn.2008.20117

Marco-Pallares, J., Cucurell, D., Cunillera, T., Garcia, R., Andres-Pueyo, A., Munte, T. F., & Rodriguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia, 46*(1), 241-248. doi:10.1016/j.neuropsychologia.2007.07.016

Marco-Pallares, J., Munte, T. F., & Rodriguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neurosci Biobehav Rev, 49*, 1-7. doi:10.1016/j.neubiorev.2014.11.014

Mas-Herrero, E., Ripolles, P., HajiHosseini, A., Rodriguez-Fornells, A., & Marco-Pallares, J. (2015). Beta oscillations and reward processing: Coupling

oscillatory activity and hemodynamic responses. *Neuroimage, 119*, 13-19. doi:10.1016/j.neuroimage.2015.05.095

Mather, M., & Carstensen, L. L. (2003). Aging and attentional biases for emotional faces. *Psychol Sci, 14*(5), 409-415. doi:10.1111/1467-9280.01455

Matsumoto, M., & Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature, 447*(7148), 1111-1115. doi:10.1038/nature05860

Matsumoto, M., & Hikosaka, O. (2009). Representation of negative motivational value in the primate lateral habenula. *Nature Neuroscience, 12*(1), 77-84. doi:10.1038/nn.2233

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience, 10*(5), 647-656. doi:10.1038/nn1890

McRitchie, D. A., Cartwright, H. R., & Halliday, G. M. (1997). Specific A10 dopaminergic nuclei in the midbrain degenerate in Parkinson's disease. *Exp Neurol, 144*(1), 202-213. doi:10.1006/exnr.1997.6418

Mehta, M. A., Swainson, R., Ogilvie, A. D., Sahakian, J., & Robbins, T. W. (2001). Improved short-term spatial memory but impaired reversal learning following the dopamine D(2) agonist bromocriptine in human volunteers. *Psychopharmacology (Berl), 159*(1), 10-20. doi:10.1007/s002130100851

Miller, D. T., & Ross, M. (1975). Self-Serving Biases in the Attribution of Causality: Fact or Fiction? *Psychological Bulletin, 82*(2), 213-225.

Miltner, W. H., Braun, C. H., & Coles, M. G. (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), 788-798. doi:10.1162/jocn.1997.9.6.788

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*(5), 1936-1947.

Moore, R., & Bloom, F. (1978). Central catecholamine neuron systems: anatomy and physiology of the dopamine systems. *Annu Rev Neurosci, 1*(1), 129-169.

Morelli, S. A., Sacchet, M. D., & Zaki, J. (2015). Common and distinct neural correlates of personal and vicarious reward: A quantitative meta-analysis. *Neuroimage, 112*, 244-253. doi:10.1016/j.neuroimage.2014.12.056

Moustafa, A. A., Cohen, M. X., Sherman, S. J., & Frank, M. J. (2008). A role for dopamine in temporal decision making and reward maximization in parkinsonism. *Journal of Neuroscience, 28*(47), 12294-12304. doi:10.1523/jneurosci.3116-08.2008

Mueller, E. M., Burgdorf, C., Chavanon, M. L., Schweiger, D., Hennig, J., Wacker, J., & Stemmler, G. (2014). The COMT Val158Met polymorphism regulates the effect of a dopamine antagonist on the feedback-related negativity. *Psychophysiology, 51*(8), 805-809. doi:10.1111/psyp.12226

Mulligan, E. M., & Hajcak, G. (2017). The electrocortical response to rewarding and aversive feedback: The reward positivity does not reflect salience in simple gambling tasks. *Int J Psychophysiol*. doi:10.1016/j.ijpsycho.2017.11.015

Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Kluger, A., Ferris, S., . . . Schwartz, R. (2003). Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci, 15*(2), 185-193. doi:10.1162/089892903321208123

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology, 38*(5), 752-760.

Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G., Holroyd, C. B., Kok, A., & van der Molen, M. W. (2002). A computational account of altered error

processing in older age: dopamine and the error-related negativity. *Cogn Affect Behav Neurosci, 2*(1), 19-36.

O'Doherty, J. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol, 14*(6), 769-776. doi:10.1016/j.conb.2004.10.016

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454. doi:10.1126/science.1094285

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation, 18*(2), 283-328. doi:10.1162/089976606775093909

Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J Cogn Neurosci, 19*(12), 1994-2004. doi:10.1162/jocn.2007.19.12.1994

Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience, 5*, 8. doi:10.3389/fnhum.2011.00008

Orr, C., & Hester, R. (2012). Error-related anterior cingulate cortex activity and the prediction of conscious error awareness. *Frontiers in Human Neuroscience, 6*, 177. doi:10.3389/fnhum.2012.00177

Otmakhova, N. A., & Lisman, J. E. (1998). D1/D5 dopamine receptors inhibit depotentiation at CA1 synapses via cAMP-dependent mechanism. *Journal of Neuroscience, 18*(4), 1270-1279.

Packard, M. G., & Knowlton, B. J. (2002). Learning and Memory Functions of the Basal Ganglia. *Annu Rev Neurosci, 25*(1), 563-593. doi:10.1146/annurev.neuro.25.112701.142937

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience, 5*, 97. doi:10.1038/nn802

Parent, A., Levesque, M., & Parent, M. (2001). A re-evaluation of the current model of the basal ganglia. *Parkinsonism Relat Disord, 7*(3), 193-198.

Pavlov, I. P. (1927). Conditional reflexes: an investigation of the physiological activity of the cerebral cortex.

Pavlov, I. P., & Gantt, W. (1928). Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals.

Peterburs, J., Kobza, S., & Bellebaum, C. (2016). Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). *Psychophysiology, 53*(2), 209-215. doi:10.1111/psyp.12560

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia, 41*(3), 245-251.

Polezzi, D., Sartori, G., Rumiati, R., Vidotto, G., & Daum, I. (2010). Brain correlates of risky decision-making. *Neuroimage, 49*(2), 1886-1894. doi:10.1016/j.neuroimage.2009.08.068

Posner, M. I., DiGirolamo, G. J., & Fernandez-Duque, D. (1997). Brain mechanisms of cognitive skills. *Consciousness and cognition, 6*(2-3), 267-290.

Proudfit, G. H. (2015). The reward positivity: from basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449-459. doi:10.1111/psyp.12370

Radhakrishnan, P., Arrow, H., & Sniezek, J. A. (1996). Hoping, Performing, learning, and Predicting: Changes in the Accuracy of Self-Evaluations of Peformance. *Human Performance, 9*(1), 23-49.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory, 2*, 64-99.

Rodriguez, M. L., & Logue, A. W. (1988). Adjusting delay to reinforcement: comparing choice in pigeons and humans. *Journal of Experimental Psychology: Animal Behavior Processes, 14*(1), 105.

Salas, R., Baldwin, P., de Biasi, M., & Montague, P. R. (2010). BOLD Responses to Negative Reward Prediction Errors in Human Habenula. *Frontiers in Human Neuroscience, 4*, 36. doi:10.3389/fnhum.2010.00036

Sallet, J., Quilodran, R., Rothe, M., Vezoli, J., Joseph, J. P., & Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. *Cogn Affect Behav Neurosci, 7*(4), 327-336.

Samson, R. D., Frank, M. J., & Fellous, J. M. (2010). Computational models of reinforcement learning: the role of dopamine as a reward signal. *Cogn Neurodyn, 4*(2), 91-105. doi:10.1007/s11571-010-9109-x

Samuelson, P. A. (1937). Some Aspects of the Pure Theory of Capital. *The Quarterly Journal of Economics, 51*(3), 469-496. doi:10.2307/1884837

Sawamoto, N., Piccini, P., Hotton, G., Pavese, N., Thielemans, K., & Brooks, D. J. (2008). Cognitive deficits and striato-frontal dopamine release in Parkinson's disease. *Brain, 131*(5), 1294-1302. doi:10.1093/brain/awn054

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol, 7*(2), 191-197.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J Neurophysiol, 80*(1), 1-27.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*(2), 241-263.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience, 13*(3), 900-913.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599.

Scimeca, Jason M., & Badre, D. (2012). Striatal Contributions to Declarative Memory Retrieval. *Neuron, 75*(3), 380-392. doi:https://doi.org/10.1016/j.neuron.2012.07.014

Seer, C., Lange, F., Loens, S., Wegner, F., Schrader, C., Dressler, D., . . . Kopp, B. (2017). Dopaminergic modulation of performance monitoring in Parkinson's disease: An event-related potential study. *Sci Rep, 7*, 41222. doi:10.1038/srep41222

Seger, C. A. (2006). The Basal Ganglia in Human Learning. *The Neuroscientist, 12*(4), 285-290. doi:10.1177/1073858405285632

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychol Rev, 94*(4), 439-454.

Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K. P., Dayan, P., & Dolan, R. J. (2012). Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease. *Brain, 135*(Pt 6), 1871-1883. doi:10.1093/brain/aws083

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends Cogn Sci, 14*(10), 464-472. doi:10.1016/j.tics.2010.08.002

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain, 127*(Pt 4), 851-859. doi:10.1093/brain/awh100

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal phase locking to hippocampal theta oscillations. *Neuron, 46*(1), 141-151. doi:10.1016/j.neuron.2005.02.028

Simons, R. F. (2010). The way of our errors: theme and variations. *Psychophysiology, 47*(1), 1-14. doi:10.1111/j.1469-8986.2009.00929.x

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. Oxford, England: Appleton-Century.

Smith, Y., & Kieval, J. Z. (2000). Anatomy of the dopamine system in the basal ganglia. *Trends Neurosci, 23*, S28-S33.

Smittenaar, P., Chase, H. W., Aarts, E., Nusselein, B., Bloem, B. R., & Cools, R. (2012). Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection--learning or performance? *Eur J Neurosci, 35*(7), 1144-1151. doi:10.1111/j.1460-9568.2012.08043.x

Soder, H. E., & Potts, G. F. (2017). Medial frontal cortex response to unexpected motivationally salient outcomes. *Int J Psychophysiol*. doi:10.1016/j.ijpsycho.2017.11.003

Squire, L. R. (1992). Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *J Cogn Neurosci, 4*(3), 232-243. doi:10.1162/jocn.1992.4.3.232

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proc Natl Acad Sci U S A, 93*(24), 13515-13522.

Staresina, B. P., & Davachi, L. (2009). Mind the gap: binding experiences across space and time in the human hippocampus. *Neuron, 63*(2), 267-276. doi:10.1016/j.neuron.2009.06.024

Steinhauser, R., Maier, M. E., & Steinhauser, M. (2017). Neural signatures of adaptive post-error adjustments in visual search. *Neuroimage, 150*, 270-278. doi:10.1016/j.neuroimage.2017.02.059

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1): MIT press Cambridge.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia, 38*(5), 596-612. doi:https://doi.org/10.1016/S0028-3932(99)00103-7

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience, 33*(19), 8264-8269. doi:10.1523/JNEUROSCI.5695-12.2013

Thorndike, E. L. (1898). Some Experiments on Animal Intelligence. *Science, 7*(181), 818-824. doi:10.1126/science.7.181.818

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology, 39*(1/4), 212-222. doi:10.2307/1415413

Tian, J., & Uchida, N. (2015). Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. *Neuron, 87*(6), 1304-1316. doi:10.1016/j.neuron.2015.08.028

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307*(5715), 1642-1645. doi:10.1126/science.1105370

Towey, J., Rist, F., Hakerem, G., Ruchkin, D. S., & Sutton, S. (1980). N250 latency and decision time. *Bulletin of the Psychonomic Society, 15*(6), 365-368. doi:10.3758/bf03334559

Tricomi, E., & Fiez, J. A. (2012). Information content and reward processing in the human striatum during performance of a declarative memory task. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 361-372. doi:10.3758/s13415-011-0077-3

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev, 94*(1), 35-79. doi:10.1152/physrev.00041.2012

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biol Psychol, 51*(2), 109-128. doi:https://doi.org/10.1016/S0301-0511(99)00032-0

Volkow, N. D., Fowler, J. S., Gatley, S. J., Logan, J., Wang, G. J., Ding, Y. S., & Dewey, S. (1996). PET evaluation of the dopamine system of the human brain. *J Nucl Med, 37*(7), 1242-1256.

Volpato, C., Schiff, S., Facchini, S., Silvoni, S., Cavinato, M., Piccione, F., . . . Birbaumer, N. (2016). Dopaminergic Medication Modulates Learning from Feedback and Error-Related Negativity in Parkinson's Disease: A Pilot Study. *Frontiers in Behavioral Neuroscience, 10*, 205. doi:10.3389/fnbeh.2016.00205

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . Morey, R. D. (2017). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychon Bull Rev*. doi:10.3758/s13423-017-1343-3

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*(8), 1870-1884. doi:10.1016/j.neubiorev.2012.05.008

Wang, J., Chen, J., Lei, Y., & Li, P. (2014). P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *Neuroreport, 25*(13), 1044-1048. doi:10.1097/WNR.0000000000000226

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev, 90*(3), 1195-1268. doi:10.1152/physrev.00035.2008

Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (2012). Better late than never? The effect of feedback delay on ERP indices of reward processing. *Cogn Affect Behav Neurosci, 12*(4), 671-677. doi:10.3758/s13415-012-0104-z

Weinberg, A., Riesel, A., & Proudfit, G. H. (2014). Show me the Money: the impact of actual rewards and losses on the feedback negativity. *Brain Cogn, 87*, 134-139. doi:10.1016/j.bandc.2014.03.015

Weismüller, B., & Bellebaum, C. (2016). Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. *Psychophysiology, 53*(11), 1739-1750. doi:10.1111/psyp.12738

Weismüller, B., Ghio, M., Logmin, K., Hartmann, C., Schnitzler, A., Pollok, B., . . . Bellebaum, C. (2018). Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication. *Neuropsychologia, 117*, 46-54. doi:https://doi.org/10.1016/j.neuropsychologia.2018.05.010

Weismüller, B., Kullmann, J., Hoenen, M., & Bellebaum, C. (2018). *Effects of feedback delay and agency on feedback-locked beta and theta power during reinforcement learning.* Manuscript submitted for publication.

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *Journal of Neuroscience, 32*(22), 7528-7537. doi:10.1523/JNEUROSCI.6352-11.2012

Williams, S. M., & Goldman-Rakic, P. S. (1998). Widespread origin of the primate mesofrontal dopamine system. *Cereb Cortex, 8*(4), 321-345.

Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R., & Eskandar, E. N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience, 7*(12), 1370-1375. doi:10.1038/nn1354

Wise, R. A. (2008). Dopamine and reward: the anhedonia hypothesis 30 years on. *Neurotox Res, 14*(2-3), 169-183. doi:10.1007/BF03033808

Wise, R. A., Spindler, J., deWit, H., & Gerberg, G. J. (1978). Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. *Science, 201*(4352), 262-264. doi:10.1126/science.566469

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Duzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron, 45*(3), 459-467. doi:10.1016/j.neuron.2005.01.010

World Health Organization. (1992). The international statistical classification of diseases and related health problems (10th ed.). Geneva: World Health Organization.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Res, 1286*, 114-122. doi:10.1016/j.brainres.2009.06.032

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *Neuroreport, 15*(16), 2561-2565.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959. doi:10.1037/0033-295X.111.4.939

Yin, H., Ostlund, S., & Balleine, B. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *European Journal of Neuroscience, 28*(8), 1437-1448. doi:doi:10.1111/j.1460-9568.2008.06422.x

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science, 323*(5920), 1496-1499. doi:10.1126/science.1167342

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*(12), 3606-3613. doi:10.1016/j.neuropsychologia.2010.08.010

7 Eidesstattliche Erklärung

Eidesstattliche Erklärung gemäß § 5 der Promotionsordnung vom 06.12.2013 der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf:

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorliegenden oder ähnlichen Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den _____ _____

Datum Benjamin Weismüller

8 Acknowledgements

I would like to use this opportunity to officially thank the whole department of Biological Psychology for making the last four-and-a-half years a wonderful experience that enabled me to develop scientifically and personally.

First of all, I am especially very thankful for Prof. Dr. Christian Bellebaum's professional advice, guidance, and supervision, but also for his vast patience and kind understanding that always comforted me during my past years in research.

Also, I would like to thank my co-supervisor Prof. Dr. Tobias Kalenscher and my mentor Prof. Dr. Jochen Musch for their support and for making this possible.

I furthermore would like to particularly thank Laura Bechtold for her lovely companionship during all struggle and scientific insanity that made us become something like siblings. Thank you also for always finding an answer to whatever question asked instead of googleing it myself. And of course, thank you for bringing Murphy every day.

Also, I am very grateful to Dr. Jutta Peterburs for supporting me professionally (e.g., by providing "a second brain" on some thoughts or by proof-reading) and taking special care of my sportive behaviour and thus my physical and eventually psychological health.

Grazie, to Dr. Marta Ghio for being such a nice colleague and for never having to drink a coffee alone or to spend a weekend in the office alone. Thank you also for your contribution to Study 1.

Of course, I would like to thank Janna Kullmann, Dr. Matthias Hoenen, Kazimierz Logmin, Dr. Christian Hartmann, Prof. Dr. Alfons Schnitzler, Prof. Dr. Bettina Pollok, and Prof. Dr. Martin Südmeyer for their very nice and friendly cooperation.

I am also thankful for everything my family did to support me through all struggle and for persevering.

Last, I would like to thank everybody who participated in the studies described here.

9 Appendix

Original article of Study 1

Weismüller, B., Ghio, M., Logmin, K., Hartmann, C., Schnitzler, A., Pollok, B., Südmeyer, M., & Bellebaum, C. (2018). Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication. *Neuropsychologia, 117*, 46-54.

I was the main author of this article. To prepare the manuscript, I compiled the study materials to implement the empirical study reported in the article. I assessed the medical tests (with the patients) and psychometric tests, and acquired, analysed, and interpreted the data resulting from those tests and the behavioural paradigm.

Original article of Study 2

Weismüller, B., & Bellebaum, C. (2016). Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. *Psychophysiology, 53*(11), 1739-1750.

I was the main author of this article. To prepare the manuscript, I was responsible for the study materials and an operational study paradigm. I implemented the empirical study reported in the article by acquiring, analysing, and interpreting the resulting electrophysiological and behavioural data.

Original manuscript of Study 3

Weismüller, B., Kullmann, J., Hoenen, M., & Bellebaum, C. (2018). *Effects of feedback delay and agency on feedback-locked beta and theta power during reinforcement learning.* Manusicript submitted for publication.

I was the main author of this manuscript. To prepare the manuscript, I implemented the empirical study reported in the article. I supervised the data acquisition, and analysed and interpreted the resulting electrophysiological and behavioural data.

Neuropsychologia 117 (2018) 46–54

Contents lists available at ScienceDirect

Neuropsychologia

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

Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication

Benjamin Weismüller[a,*], Marta Ghio[a], Kazimierz Logmin[b], Christian Hartmann[b], Alfons Schnitzler[b,c], Bettina Pollok[c], Martin Südmeyer[b], Christian Bellebaum[a]

[a] Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf, Germany
[b] Department of Neurology, University Hospital Düsseldorf, Germany
[c] Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich-Heine University Düsseldorf, Germany

ARTICLE INFO

Keywords:
Feedback learning
Parkinson's disease
Dopamine
Feedback delay

ABSTRACT

Phasic dopamine (DA) signals conveyed from the substantia nigra to the striatum and the prefrontal cortex crucially affect learning from feedback, with DA bursts facilitating learning from positive feedback and DA dips facilitating learning from negative feedback. Consequently, diminished nigro-striatal dopamine levels as in unmedicated patients suffering from Parkinson's Disease (PD) have been shown to lead to a negative learning bias. Recent studies suggested a diminished striatal contribution to feedback processing when the outcome of an action is temporally delayed. This study investigated whether the bias towards negative feedback learning induced by a lack of DA in PD patients OFF medication is modulated by feedback delay. To this end, PD patients OFF medication and healthy controls completed a probabilistic selection task, in which feedback was given immediately (after 800 ms) or delayed (after 6800 ms). PD patients were impaired in immediate but not delayed feedback learning. However, differences in the preference for positive/negative learning between patients and controls were seen for both learning from immediate and delayed feedback, with evidence of stronger negative learning in patients than controls. A Bayesian analysis of the data supports the conclusion that feedback timing did not affect the learning bias in the patients. These results hint at reduced, but still relevant nigro-striatal contribution to feedback learning, when feedback is delayed.

1. Introduction

Most living beings learn from the outcomes of their actions and adapt their behaviour accordingly, which defines reinforcement learning. In everyday-life, outcomes can vary not only in their valence, but also in their delay following an action. Often they occur immediately, as for example when making an error in driving your car and causing an accident. They can, however, also follow after a couple of seconds like when pushing a button on a coffee dispenser, or after a very long delay, for example in financial investments.

Reinforcement learning means gaining the knowledge of both which action previously resulted in a profitable outcome and which action previously resulted in a negative outcome. Animal studies associated outcomes that are better or worse than predicted with phasic increases and decreases in midbrain dopamine (DA) neuron activity, respectively (Schultz, 1997, 2000; Schultz and Dayan, 1997). Neural network models consider projections of this DA prediction error signal to the basal ganglia and prefrontal cortex (PFC, including the anterior cingulate cortex, ACC; Bédard and Larochelle, 1969; Haber and Fudge, 1997; Lavoie and Smith, 1989; Lehéricy et al., 2004; Lynd-Balta and Haber, 1994) as the neuronal underpinnings of reinforcement learning (Frank, 2005; Frank et al., 2004), underlying the adaptation of behaviour (Sheth et al., 2012). Based on DA effects on two separate so called Go and NoGo pathways within the basal ganglia (Aubert and Ghorayeb, 2000; Frank, 2005; Frank et al., 2004; Gerfen, 1992; Hernandez-Lopez et al., 1997, 2000), chronically increased and decreased DA levels have been linked to better learning from positive and negative feedback, respectively, which has indeed been shown in Parkinson's disease (PD) patients ON and OFF DA replacement medication (Frank et al., 2004; Frank and Samanta, 2007). Generally diminished DA baseline levels reduce the chance of DA bursts and increase the chance of DA dips reaching a certain threshold level, resulting in a more dominant NoGo-pathway during learning, whereas DA replacement medication seems to lead to a DA overdose in the ventral striatum (Cools and Barker, 2001, 2003; Frank, 2005) so that the Go pathway is selectively strengthened (see Frank et al., 2007).

* Correspondence to: Institute of Experimental Psychology, Department of Biological Psychology, Heinrich-Heine University Düsseldorf, Germany.
E-mail address: Benjamin.Weismueller@hhu.de (B. Weismüller).

https://doi.org/10.1016/j.neuropsychologia.2018.05.010
Received 11 December 2017; Received in revised form 27 April 2018; Accepted 10 May 2018
Available online 11 May 2018
0028-3932/

Kobza et al. (2012) showed, however, that a lack of DA does not always lead to a negative learning bias. They found learning to be unaffected in PD patients OFF medication when they learned from the choices of another person and the accompanying outcomes, suggesting that the mechanisms in observational learning differ. Another condition, in which feedback processing seems to be altered, relates to learning from delayed feedback. When comparing the feedback-based acquisition of stimulus-outcome associations, PD patients were significantly impaired in learning from immediate, but not from delayed feedback appearing seven seconds after a choice response (Foerde and Shohamy, 2011). While activity in the dorsal striatum appeared to underlie immediate feedback processing, the hippocampus was more strongly involved during learning from delayed feedback, as was shown via functional magnetic resonance imaging (fMRI) in healthy subjects (Foerde and Shohamy, 2011). The importance of the hippocampus for delayed feedback processing was further corroborated by deficits in amnestic patients with suspected hippocampal damage (Foerde et al., 2013). Recent studies using electroencephalography (EEG) added to the impression of different neural mechanisms for immediate and delayed feedback processing. They reported that the feedback-related negativity (FRN) was diminished for delayed compared to immediate feedback (Arbel et al., 2017; Peterburs et al., 2016; Weinberg et al., 2012; Weismüller and Bellebaum, 2016). The FRN is a feedback-locked event-related potential (ERP) component that has been linked to DA effects on the ACC (Holroyd, 2004; Holroyd and Coles, 2002, 2008). Reduced FRN amplitudes thus appear to suggest reduced DA system involvement with increasing temporal delay between action and outcome, so that overall a pattern of findings emerges that suggests a weaker or even absent role of DA in delayed feedback processing. As the bias for better learning from negative than positive action outcomes found for immediate feedback has directly been linked to the lack of DA in unmedicated PD patients, one might hypothesize that learning from delayed feedback should not be affected. The negative learning bias in this patient group should thus appear exclusively for immediate feedback.

On the other hand, the mentioned ERP studies also suggest similarities in the processing of immediate and delayed feedback. Irrespective of feedback delay, negative feedback elicited a larger FRN amplitude than positive feedback (Peterburs et al., 2016; Weismüller and Bellebaum, 2016). Moreover, even for delayed feedback the FRN reflected feedback expectations and was thus influenced by the reward prediction error (Weismüller and Bellebaum, 2016), suggesting that the DA system did indeed contribute to delayed feedback processing, at least to some extent. Striatum and hippocampus might work together in associating responses to outcomes (Dickerson and Delgado, 2015; Dickerson et al., 2011). Based on these considerations, it might thus also be possible that a lack of DA as in unmedicated PD patients has comparable effects on learning from delayed and learning from immediate feedback, leading to similar negative learning biases.

In this study, we applied variants of the probabilistic selection task first described by Frank et al. (2004) to explore whether the effect of reduced DA levels on the preference for learning to avoid a non-beneficial stimulus (learning from negative feedback) over learning to choose a beneficial stimulus (learning from positive feedback) is modulated by feedback delay. For this purpose, we compared the performance of two groups of PD patients OFF medication completing an immediate (see Kobza et al., 2012) or delayed feedback version of the probabilistic selection task with each other and with the performance of corresponding groups of healthy control subjects.

2. Material and methods

2.1. Participants

Four groups of subjects participated in the present study, two groups of PD patients OFF medication and two groups of healthy control subjects. With 12 participants in each patient group and 24 participants in each control group the sample sizes were slightly larger than in a previous study of our group applying variants of the same experimental paradigm and addressing a related research question (Kobza et al., 2012). One patient and one control group each completed an immediate and delayed feedback version of the probabilistic selection task, respectively. The patient group for the immediate feedback condition had a mean age of 56.8 years (SD = 9.8; 7 men). For ten of these patients we reused data from a sample of PD patients who had already been tested for a previous study by our group (see Kobza et al., 2012; the group of subjects learning actively from their own choices). To match the delayed feedback group (see below) two additional PD patients were recruited. Similarly, data for 20 control subjects were taken from our old data set for the immediate feedback group (Kobza et al., 2012) and four more control subjects were tested. The control group learning from immediate feedback had a mean age of 55.5 years (SD = 10.0; 14 men). The patient (9 men) and control groups (16 men) in the delayed feedback condition were on average 57.9 (SD = 8.5) and 59.1 years (SD = 6.6) old. All PD patients were listed for regular attendance at the Centre for Movement Disorders and Neuromodulation of the University Hospital Düsseldorf and were diagnosed by medical staff according to the UK Brain Bank criteria (Hughes and Daniel, 1992). Symptom severity in all PD patients was between stages I and III according to the Hoehn and Yahr classification (Hoehn and Yahr, 1967) and all patients had normal or corrected-to-normal vision. To compare symptom severity between ON and OFF medication states, the Unified Parkinson's Disease Rating Scale (UPDRS; Goetz et al., 2008; Movement Disorder Society Task Force on Rating Scales for Parkinson's, 2003) was administered twice for each patient, once in the OFF state and a second time 20 min after the intake of the regular medication after testing in the ON state. The average scores amounted to 21.8 (SD = 5.8) in the OFF state and 15.5 (SD = 6.0) in the ON state for patients in the immediate feedback condition. For patients in the delayed feedback condition the average scores were 30.2 (SD = 9.3) and 18.0 (SD = 10.2) for OFF and ON state, respectively. For both groups, the scores with and without medication differed significantly ($t(11) = 5.637$; $p < .001$; $d = 1.627$ for immediate feedback and $t(11) = 8.512$; $p < .001$; $d = 2.457$ for delayed feedback). The scores were obtained with different versions of the scale. For the 10 participants that entered analysis and were tested for our previous study (Kobza et al., 2012) the version from 2003 was used (Movement Disorder Society Task Force on Rating Scales for Parkinson's, 2003), whereas for the patients tested for the present study a newer version was used (Goetz et al., 2008), which yields higher symptom scores on average for lower stages of PD (Goetz et al., 2008). A direct comparison of the UPDRS scores between the two patient groups may thus be confounded and was therefore not conducted.

Exclusion criteria for patients were psychiatric or neurological diseases (other than PD), atypical PD, traumatic brain injury with sustained unconsciousness, suspected or documented drug or alcohol abuse, and regular psychotropic medication other than DA agonists. Finally, all PD patients were screened for comorbid depression and dementia after the experimental learning task was applied but still in the OFF medication state using the Beck Depression Inventory (BDI; Hautzinger et al., 2006) and the Mini Mental Status test (MMS; Folstein and Folstein, 1975), respectively. BDI scores were assessed in the patients in order to exclude that a negative affective bias could influence performance in the probabilistic selection task. None of the patients scored above 18, which would indicate clinically relevant depressive symptoms. More specifically, patients learning from immediate feedback had a mean BDI score of 7.3 (SD = 4.4), while those in the delayed feedback group had a mean score of 7.6 (SD = 4.2). The scores did not differ significantly between the two patient groups ($p = .870$). In the MMS all the patients scored above 27 (patients immediate feedback: mean = 28.0, SD = 1.5; patients delayed feedback: mean = 28.9, SD = .8), indicating that none of the patients showed signs of dementia or clinically relevant cognitive impairment. Scores for the MMS did not differ between patient groups ($p = .080$). BDI and MMS scores were

also obtained in the control subjects learning from delayed feedback. Also in this group, none of the participants reached clinically relevant values (mean BDI score = 5.2, *SD* = 3.9; mean MMS score = 28.4, *SD* = 1.1). As the 24 participants of this group were comparable with respect to age and gender distribution to both patient groups, we compared the BDI and MMS scores between the healthy controls and each of the two groups of patients separately. The BDI and MMS scores of the control group did neither differ from the scores of the patients learning from immediate (both *ps* > .150), nor from those learning from delayed feedback (both *ps* > .099).

Healthy controls were recruited via advertisements in local newspapers. Exclusion criteria were a history of psychiatric or neurological disease, traumatic brain injury with sustained unconsciousness, suspected or documented drug or alcohol abuse, and psychotropic medication.

All subjects voluntarily participated in the study and gave informed written consent before testing. The study was approved by the ethics committee of the Medical Faculty of Heinrich-Heine University Düsseldorf (study no. 2849). PD patients as well as healthy controls received monetary reimbursement for their participation.

2.2. Probabilistic selection task

In this study we used modified versions of the probabilistic selection task that has first been described by Frank et al. (2004). This task consisted of three different phases: at least one learning and test phase and a final transfer phase. In the learning phase, on each trial, one out of three different symbol pairs was randomly presented to the participants (symbols A/B, C/D, and E/F). Each symbol pair appeared 20 times per learning phase, resulting in 60 trials in total. The participants were asked to choose one of the two symbols to receive a positive ("correct") or negative ("incorrect") feedback (Fig. 1A, left column, shows the sequence of events in an example trial of the learning phase). The participants could use the feedback to learn, which symbol was associated with what type of feedback. Unknown to the participants, each symbol was locked to a particular probability for positive feedback, and the pairs differed with respect to the probabilities: For symbol pair A/B the likelihood of receiving positive feedback was 80% when choosing symbol A and 20% when choosing B. For the other pairs, probabilities were 70% vs. 30% (C/D pair) and 60% vs. 40% (E/F pair). If the participants failed to respond within 3500 ms, a reminder was presented asking them to respond faster and the trial was scored as a miss.

After the learning phase was completed, the test phase started, in which each symbol pair was presented 10 times resulting in a total of 30 trials. In these trials, no feedback was presented so that the participants had to apply the stimulus-response associations they had learned based on the outcomes during the learning phases. Test phases thus served to examine, in how far participants continued to react according to the knowledge they gained during the learning phase in the absence of trial-by-trial feedback (Foerde and Knowlton, 2006; Foerde et al., 2013; Foerde and Shohamy, 2011; Kobza et al., 2012). Fig. 1A (right column) shows the sequence of events in an example trial of the test phase. As soon as the participants reached a certain performance criterion in the test phase (choosing the "correct" symbol in 80% of the A/B pair trials and in 70% of the C/D pair trials), the transfer phase started (see below). Choices were considered correct, when participants chose the symbol that was associated with the higher probability for positive feedback during the learning phase. If the participants did not manage to reach this learning criterion, learning and test phases were repeated. After a maximum of four repetitions, the transfer phase was initiated regardless of the participant's choice accuracy (see Fig. 1B).

In accordance with previous studies by Frank and colleagues (2004) and by our group (Bellebaum and Rustemeier, 2011; Kobza et al., 2012), new symbol combinations in the transfer phase (pairs A/C, A/D, A/E, A/F and B/C, B/D, B/E, B/F) allowed to disentangle learning from

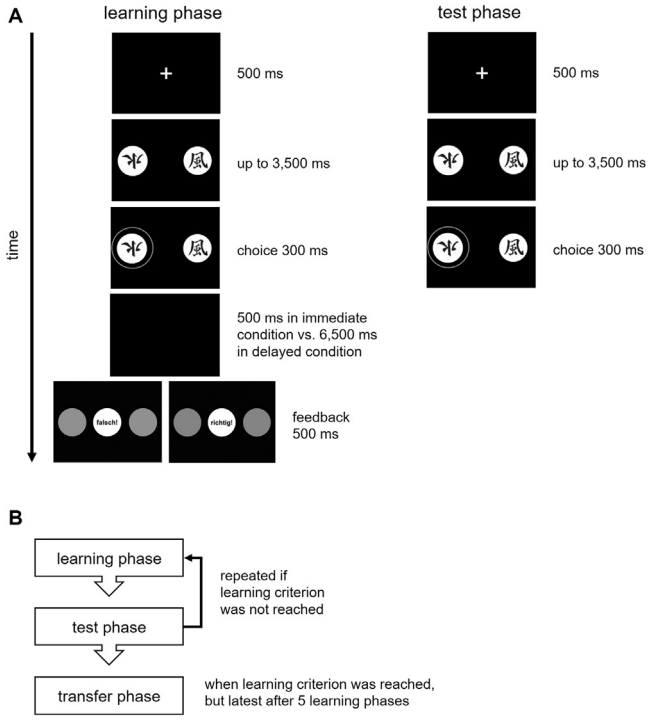

Fig. 1. A) Time course of events in a single trial during the learning phase (on the left side) and the test phase without feedback (on the right side). The words "richtig" and "falsch" are the German words for "correct" and "incorrect" and served as positive and negative feedback, respectively. Note that the sequence of events in a single transfer trial was identical to the sequence in test trials. However, different stimulus pairs were presented (see Section 2.2). B) Test phases followed learning phases until participants reached the learning criterion and the transfer phase was initiated. If participants failed to reach the learning criterion, the transfer phase started after the 5th test phase.

positive and negative feedback during the learning phase. "Positive learners" were expected to more reliably choose symbol "A" over all other symbols compared to avoiding symbol B in pairs in which "B" appeared. For "negative learners" the opposite pattern was expected. The transfer phase consisted of 40 trials (each pair being presented 5 times), 20 involving stimulus "A" and 20 involving "B", and no feedback was presented. The sequence of events in the trials of the transfer phase was identical to the sequence in the test trials.

Importantly, two different versions of the described probabilistic selection task were applied which only differed in the timing of the feedback during the learning phase. Half of the participants received "immediate feedback". In this condition feedback followed 500 ms after the subject's chosen symbol had been encircled in red for 300 ms to indicate the choice, that is, 800 ms after the button press. The other half of the participants learned from "delayed feedback", where information about choice accuracy was shown 6500 ms after the subject's choice had been indicated for 300 ms (6800 ms after the button press). The feedback stimulus was presented for 500 ms, then a fixation cross followed randomly for 500, 1000, or 2000 ms. As outlined above, no feedback was presented during the test (see Fig. 1A) and transfer phases.

2.3. Procedure

In order to reach the OFF medication state, the patients participated after 12 h of medication withdrawal over-night. Eight of the 24 patients also received long-lasting DA agonists as part of their anti-parkinsonian medication. These were usually administered once per day in the morning. As testing usually took place in the morning on the next day, they were OFF this medication for 24 h or more at the time of testing. Prior to the experiment, the participants gave informed written consent

and a structured interview was conducted to assess demographic data and potential health or drug problems. UPDRS scores were assessed for the first time before the experiment was conducted. After the patients had completed the probabilistic selection task, the MMST and the BDI were administered. Then the patients received their individual dose of PD specific medication by the medical staff in the Department of Neurology at the University Hospital Düsseldorf, and after a break of about 20 min, UPDRS scores were assessed again, now in the ON-state. Similarly, healthy controls signed an informed consent form, took part in the structured interview and completed then the probabilistic selection task.

2.4. Analysis

Statistical analysis was conducted using SPSS Statistics 23 software (IBM, Armonk, New York). As a general measure of performance in the probabilistic selection task, the number of correct choices from the first test phase was analysed with an analysis of variance (ANOVA) comprising the between-subjects factors *Group* (PD vs. control) and *Feedback Timing* (immediate vs. delayed) and the within-subjects factor *Pair* (A/B, C/D, and E/F; see also Kobza et al., 2012). As the participants could reach the learning criterion already after the first learning phase, the first test phase was the only one that all participants completed irrespective of their learning performance. Furthermore, an ANOVA with the factors *Group* and *Feedback timing* was used to analyse the number of learning phases needed to reach the learning criterion of 70% correct choices for the A/B pair in the test phase (see below) as a second measure of learning performance during acquisition.

Then scores for positive and negative feedback learning derived from choice accuracy in the transfer phase (the number of choices of symbol A and the number of avoidances of symbol B, see above) were analysed in an ANOVA with *Group* (PD vs. control) and *Feedback Timing* (immediate vs. delayed) as between-subjects factors and *Learning Type* (positive vs. negative) as within-subjects factor. For being included into this analysis, subjects had to reach an accuracy level of at least 70% for the A/B symbol pair during one test phase (see Kobza et al., 2012, for an application of the same criterion). This criterion was reached by all study participants. However, two control subjects learning from delayed feedback were excluded from the analysis, because they scored zero on learning from positive feedback and were outliers relative to the other control participants (*z*-scores < -4). None of the other control subjects or patients had z-scores lower than -3 or above 3 for learning from positive or negative feedback relative to the respective group of participants (controls or patients). An *alpha* level of $p < .05$ was used to determine statistical significance. For the ANOVAs, the *Greenhouse-Geisser* correction was performed, when the sphericity assumption was violated. To resolve interactions, post-hoc *t*-tests were conducted. Also, corrected *t*- and *p*-values were considered when non-homogenous variances were identified by applying the *Levene's test*.

3. Results

Fig. 2 depicts the mean percentage of correct responses for the different stimulus pairs during the first test phase, separately for both patient and control groups. Statistical analysis of performance during this phase revealed no main effects for *Group* and *Feedback Timing* (all $Fs < .673$; $ps > .414$; $\eta_p^2 < .011$), but a significant main effect for Pair ($F(2, 136) = 9.561$; $p < .001$; $\eta_p^2 = .123$), reflecting differences in difficulty depending on the feedback probabilities. All two-way and the three-way interaction involving all factors did not reach significance (all $Fs < 1.389$; $ps > .253$; $\eta_p^2 < .021$).

To further quantify general learning performance during acquisition, the number of learning phases needed to reach the learning criterion was analysed. On average, PD patients needed 1.7 ($SD = 1.2$) learning phases in the immediate and 1.5 ($SD = 1.1$) learning phases in the delayed feedback condition. Controls needed 1.3 ($SD = .6$) learning phases with immediate feedback and 1.6 ($SD = .9$) learning phases when feedback was delayed. The number of learning phases was neither different between groups ($F(1, 68) = .400$; $p = .529$; $\eta_p^2 = .006$) nor was it affected by the feedback delay conditions ($F(1, 68) = .204$; $p = .653$; $\eta_p^2 = 003$). The interaction between the factors did not reach significance, either ($F(1, 68) = 1.381$; $p = .244$; $\eta_p^2 = .020$).

Concerning the transfer phase, the mean numbers of correct choices of symbol A (learning from positive feedback) vs. the number of avoidances of symbol B (learning from negative feedback) for the control participants and the PD patients are depicted in Fig. 3A for the immediate feedback and Fig. 3B for the delayed feedback condition. Statistical analysis of choice behaviour during the transfer phase revealed no main effects for *Group*, *Feedback Timing* and *Learning Type* (all $Fs < 1.463$; $ps > .231$; $\eta_p^2 < .023$). However, a significant interaction between *Group* and *Feedback Timing* was found ($F(1, 66) = 4.198$; $p = .044$; $\eta_p^2 = .060$). Fig. 4 illustrates this interaction showing transfer phase performance averaged across negative and positive feedback learning for patients and controls who had learned from immediate and delayed feedback. Healthy controls who had learned from immediate feedback during the learning phase performed generally better during the transfer phase than the respective PD group ($t(34) = 2.252$; $p = .031$; $d = .796$), whereas no difference was seen between patients and controls who had learned from delayed feedback ($t(32) = .610$; $p = .703$; $d = .219$).

A significant interaction was found between *Learning Type* and *Group* ($F(1, 66) = 16.286$; $p < .001$; $\eta_p^2 = .198$) (see Fig. 5 for transfer phase performance averaged across subjects in the immediate and delayed feedback condition). The resolution of this interaction showed that PD patients avoided symbol B (indicating learning from negative feedback) more frequently than they chose symbol A (indicating

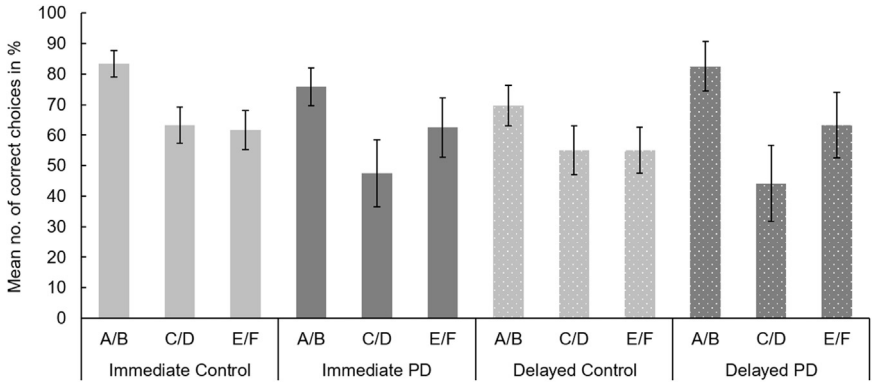

Fig. 2. Mean percentage of correct responses during the first test phase for controls and PD patients, both feedback timings, and the different stimulus pairs (error bars indicate the standard error of the mean).

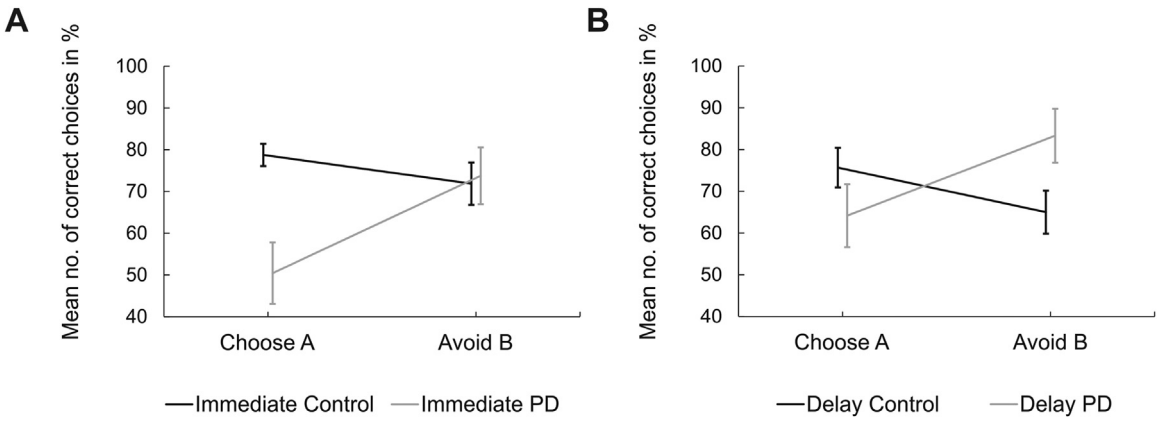

Fig. 3. Mean numbers of correct choices of symbol A (learning from positive feedback) and avoidances of symbol B (learning from negative feedback) during the transfer phase for controls and PD patients in the immediate feedback timing (A) and the delayed feedback timing condition (B) (error bars indicate the standard error of the mean).

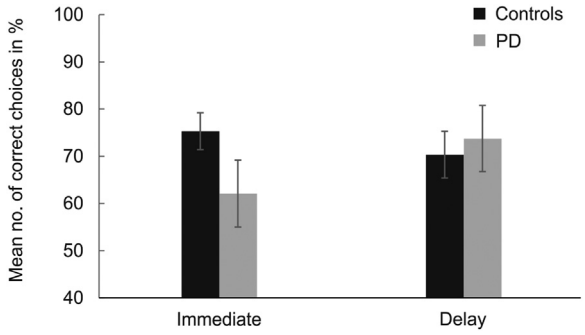

Fig. 4. Mean number of correct choices during the transfer phase for controls and PD patients in the immediate and in the delayed feedback timing condition (error bars indicate the standard error of the mean).

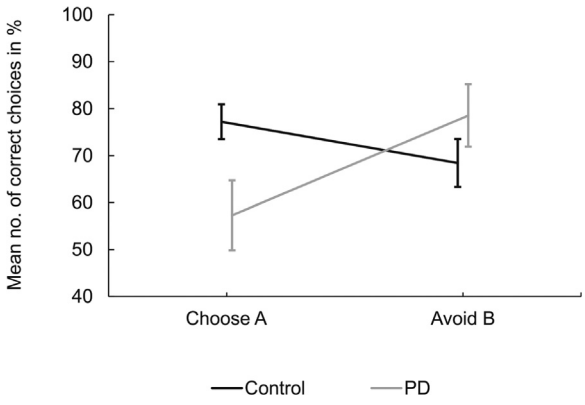

Fig. 5. Mean numbers of correct choices of symbol A (learning from positive feedback) and avoidances of symbol B (learning from negative feedback) during the transfer phase for controls and PD patients averaged across feedback timing conditions (error bars indicate the standard error of the mean).

learning from positive feedback, $t(23) = -2.877$; $p = .009$; $d = .587$), while the healthy controls preferred symbol A over symbol B ($t(46) = 2.661$; $p = .011$; $d = .397$). Both the two-way interaction between *Feedback Timing* and *Learning Type* ($F(1, 66) = .874$; $p = .353$; $\eta_p^2 = .013$) and the three-way interaction involving all factors ($F(1, 66) = 4.278$; $p = .662$; $\eta_p^2 = .003$) were not significant.

The absence of the three-way interaction and the pattern of results for immediate and delayed feedback learning appear to suggest that the difference between learning from positive and negative feedback is affected by a lack of DA irrespective of feedback timing. In order to support this conclusion, we first aimed to find out if our sample was large enough to detect a significant three-way interaction comparable to similar effects described in the literature. We thus conducted a systematic search of the literature considering all studies that cited the original study on the negative learning bias in PD patients OFF medication by Frank et al. (2004). The search revealed that only very few studies examined effects of two between subjects factors (one of them related to a DA depletion) on this bias (Jakob and Ehrentreich, 2018; Kobza et al., 2012; Lighthall and Gorlick, 2013). The study by Kobza et al. (2012) was the only one finding a significant three-way interaction of this kind, with an effect size of $\eta_p^2 = .0761$. Entering this value and our sample size of 70 participants into a compromise power analysis using the software G*Power (Faul and Erdfelder, 2007, 2009) revealed a relatively high power of .838 for detecting a three-way interaction effect of this size in our study.

We further reasoned that a similar result pattern for learning from immediate and delayed feedback processing would be reflected in interactions between *Group* and *Learning Type* for both feedback timings separately. Thus, exploratory separate analyses for subjects in the two feedback timing conditions were conducted. As expected, significant interactions between the factors *Group* (PD vs. control) and *Learning Type* (positive vs. negative) were found for both immediate ($F(1, 34) = 8.469$; $p = .006$; $\eta_p^2 = .199$) and delayed feedback ($F(1, 32) = 7.962$; $p = .008$; $\eta_p^2 = .199$).

We then further analysed to which extent the data support the conclusion that indeed only the DA level, and thus the factor *Group*, but not the factor *Feedback Timing* affects the relative bias to learn from positive or negative feedback by carrying out a Bayesian analysis using JASP (Version 0.8.3.1; JASP Team, 2017; Wagenmakers et al., 2017a, 2017b). For this purpose, we first computed the difference between the number of correct choices for A and avoidances of B (positive minus negative feedback learning) during the transfer phase for each participant, because this difference reflects the mentioned bias, which we were mostly interested in (see above). We then entered the differences in a Bayesian ANOVA with the between-subjects factors *Group* (PD vs. control) and *Feedback Timing* (immediate vs. delayed). In the Bayesian ANOVA, the null model was compared against four statistical models, each containing, respectively, the main effect of *Group*, the main effect of *Feedback Timing*, both main effects, and the interaction effect (Wagenmakers et al., 2017a, 2017b). The resulting Bayes factors (BF) for each model were computed as the ratio of the predictive adequacy (i.e. the change from prior to posterior odds brought about by the data) of each statistical model and that of the null model. Thus, the higher the BF, the more the evidence favours the specific statistical model (Wagenmakers et al., 2017a, 2017b). For the interpretation of the Bayes factors we adopted the classification suggested by Lee and

Wagenmakers (2013) (adapted from Jeffreys, 1961; see also Wagenmakers et al., 2017b), according to which a BF between 1 and 3, 3 and 10, 10 and 30, 30 and 100, and above 100 indicates, respectively, anecdotal, moderate, strong, very strong, and extreme evidence for a specific model against the null model. The priors were equally set to $p(m) = .200$ for each model, as, to our knowledge, no comparable previous data exists that would suggest any other prior for PD patients learning from immediate vs. delayed feedback.

The results revealed that the data provided extreme evidence (BF = 157.642) for the model including only the main effect of *Group* (which corresponds to the interaction between *Group* and *Feedback Type* in the conventional ANOVA, as the *Bayes* statistics was applied on the positive - negative feedback learning difference score, see above). In turn, the data did not provide evidence for the model including only the effect of *Feedback Timing* (BF = .362). However, very strong and strong evidence, was also found in favour of, respectively, the model including both the *Group* and the *Feedback Timing* main effects (BF = 66.065), and the model including the interaction of these factors (BF = 24.356). To quantify how adding the *Feedback Timing* effect and the interaction effect to the *Group*-only model weakens the evidence in favour of the model, we computed the BF by comparing the *Group*-only model against the other two models ($p(m) = .333$; Wagenmakers et al., 2017b). The data provided anecdotal evidence against including the *Feedback Timing* factor (BF = 2.407) and moderate evidence against including the interaction (BF = 6.358). Finally, to also quantify how much the data support the inclusion of each effect (main effect of *Group*, main effect of *Feedback Timing*, interaction effect), we applied the Bayesian model averaging, which computes the change from prior to posterior odds ($BF_{inclusion}$) for each effect, taking into account each candidate models' conclusions (Wagenmakers et al., 2017b). When averaged across all candidate models, the data strongly support the inclusion of the *Group* factor ($BF_{inclusion} = 120.770$), while the *Feedback Timing* and the interaction received very weak support ($BF_{inclusion} < 1$).

After we could verify that the relative bias to learn from positive or negative feedback is not affected by feedback timing, we were interested to explore the patterns underlying the interactions between *Group* and *Learning Type* that we found for both, participants learning from immediate and delayed feedback, separately in the two feedback timing conditions. For immediate feedback PD patients revealed a significant negative learning bias, learning better from negative than from positive feedback ($t(11) = -2.777$; $p = .018$; $d = .802$), while controls performed comparably for both types of feedback ($t(23) = 1.142$; $p = .265$; $d = .233$). For delayed feedback, a slightly different pattern emerged. PD patients descriptively also showed better learning from negative than positive feedback, but this effect was not significant ($t(11) = -1.531$; $p = .154$; $d = .442$). In contrast, healthy controls scored higher for positive than for negative feedback learning ($t(21) = 2.588$; $p = .017$; $d = .552$).

In an alternative resolution of the interaction focusing on between-group comparisons, however, we did find evidence of enhanced negative feedback learning in PD patients. They learned better from negative delayed feedback compared to controls ($t(32) = -2.284$; $p = .029$; $d = .820$), but did not differ from controls for learning from positive feedback ($t(32) = 1.778$; $p = .095$; $d = .734$). For immediate feedback patients scored lower than controls for positive feedback learning ($t(34) = 3.469$; $p = .004$; $d = 1.512$), but did not differ from controls for negative feedback learning ($t(34) = .211$; $p = .834$; $d = .074$). As can be seen in Fig. 3 (left) these comparisons are also affected by the generally lower performance of patients for immediate feedback learning.

4. Discussion

In this study, PD patients OFF DA replacement medication and healthy controls completed a probabilistic selection task that aimed to test the participant's preference for learning from positive or negative feedback (Frank et al., 2004). Importantly, half of the participants received immediate feedback (800 ms after stimulus choice), while the other half received feedback after a temporal delay (6800 ms after stimulus choice). The aim was to explore whether the effect of a reduced DA level in PD patients on choice preference is modulated by feedback delay.

While there were no differences between PD patients and controls or between feedback timing conditions for the acquisition of stimulus-response-outcome associations in the learning phase, participants' choices during the transfer phase revealed two different findings. First, PD patients performed generally worse than controls when feedback had previously been provided immediately, but not when it had been delayed. Second, concerning the relative bias for learning from negative or positive feedback and thus the main focus of the present study, PD patients generally scored higher for negative learning irrespective of feedback timing, while healthy controls showed the opposite bias. Separate analyses for the two *Feedback Timing* conditions revealed that the relative preference for learning from positive or negative feedback differed between patients and controls for both immediate and delayed feedback. This finding was further corroborated by a Bayesian analysis showing that the data strongly support a model in which only the *Group* factor (PD patients versus controls), and thus the DA level, affects the difference score between positive and negative feedback learning over a model including *Feedback Timing* and, importantly, over a model including the interaction between the factors.

The first finding of generally impaired performance in PD patients OFF medication for associations learned via immediate but not for those learned via delayed feedback is in line with and partially replicates findings from previous studies obtained in PD patients ON and OFF medication with similar feedback learning tasks (Foerde et al., 2013; Foerde and Shohamy, 2011). We found this general performance deficit in the transfer phase where no feedback was given. Interestingly, the deficits in the samples studied by Foerde and Shohamy (2011) and Foerde et al. (2013) were also seen in trials without feedback. However, the stimulus pairs that were presented in these trials were the same as during learning, whereas the transfer phase in the present study entailed new stimulus pairs and thus required the transfer of the learned associations.

The general difference between immediate and delayed feedback has been explained with differential roles of hippocampus and dorsal striatum in learning from delayed and immediate feedback, respectively. The importance of the hippocampus for delayed feedback learning has been shown by both imaging and patient evidence (Foerde et al., 2013; Foerde and Shohamy, 2011). It is thus conceivable that the PD patients of the present study rather used their intact MTL structures for learning from delayed feedback to reach a performance level as high as the one in control participants. The DA system with its projections to striatum and ACC might thus not be involved in delayed feedback processing.

On the other hand, the results of the present study raise doubts as to whether the mechanisms for learning from immediate and delayed feedback can be separated completely. If the postulated differential DA influence on learning from positive and negative feedback, and thus the influence on striatal information processing (Frank, 2005; Frank et al., 2004; Moustafa and Sherman, 2008; Samson and Frank, 2010), was only present for learning from immediate feedback, a different learning bias compared to healthy controls induced by a lack of DA should not be seen in PD patients OFF medication for delayed feedback. In our study, however, the relative preference for learning from positive or negative feedback differed between patients and controls for both learning from immediate and delayed feedback. This suggests that DA is involved also in learning from delayed feedback. The only difference between feedback timing conditions was that PD patients showed a significant negative bias for immediate feedback, which was absent in controls, whereas for delayed feedback no significant bias was found for the patients, but a positive bias was seen in controls. The latter finding is likely related to a general positivity effect during healthy aging

(Mather and Carstensen, 2005). Starting approximately at the age of 50 years, healthy older subjects pay more attention to and show better memory for information with a positive valence. In the context of feedback learning, the positivity effect may play a role especially when the DA influence is reduced (Mammarella et al., 2016), as for example in observational learning (Bellebaum et al., 2011; Kobza et al., 2012) or in learning from delayed feedback, as in the present study. This appears plausible because the positivity effect appears rather declarative in nature (Lind and Visentini, 2017; Mather and Carstensen, 2005), and for delayed feedback processing a stronger role of the hippocampus and thus of declarative memory processes has been proposed than for immediate feedback (Foerde and Shohamy, 2011). Thus, the positive learning bias in healthy controls for delayed feedback might be considered as a different baseline, against which the patients' performance has to be compared. The striatal influence on learning from delayed feedback then appears to be still strong enough to strengthen negative learning, which is reflected in the higher scores for learning from negative delayed feedback for patients than controls, corroborating the similarity to learning from immediate feedback.

Similarities between immediate and delayed feedback processing have also been described in other studies. In their fMRI study, Foerde and Shohamy (2011) found that feedback-locked activations in the ventral striatum did not differ between immediate and delayed feedback. In EEG studies it has repeatedly been observed that the FRN is reduced, but not absent during delayed feedback processing (Arbel et al., 2017; Peterburs et al., 2016; Weismüller and Bellebaum, 2016, but see Weinberg et al., 2012 for a different pattern). Moreover, in one study a modulation of the FRN by expectancy was found for both immediate and delayed feedback (Weismüller and Bellebaum, 2016), an effect, which has previously been linked to an influence of the DA system on feedback processing (Bellebaum and Daum, 2008; Hajcak, Moser, Holroyd, and Simons, 2007; Holroyd and Krigolson, 2007, 2009). One potential mechanism for a cooperation between the DA system/striatum and the hippocampus (Dickerson and Delgado, 2015; Dickerson et al., 2011) might relate to DA innervations of the hippocampus (Otmakhova and Lisman, 1998). DA release has been shown to facilitate hippocampal plasticity and episodic memory formation influencing adaptive behaviour based on past experiences (Shohamy and Adcock, 2010). The hippocampus might thus enable participants to link outcomes, processed by DA neurons, to actions across a temporal gap.

It has to be noted that some researchers have provided a different interpretation of the performance in the probabilistic selection task. In particular, it has been criticised that action selection rather than positive vs. negative feedback learning is tested, and that DA affects more the former than the latter (Shiner et al., 2012). Although this view does not necessarily exclude DA effects on learning, it postulates that the DA level during the transfer phase primarily determines the performance in the task: A lack of DA leads to stronger avoidance of actions that have been associated with negative feedback before, and a DA overdose leads to a stronger selection of actions that have been associated with positive feedback. Indeed, Shiner et al. (2012) reported a bias for choosing beneficial stimuli only in patients ON medication during the transfer phase, with no effect of the DA level during learning. In a related study, Smittenaar et al. (2012) assigned the valence to the outcome stimuli only when the transfer phase began and also found evidence for a role of DA on action selection. These findings are in line with a growing body of evidence stating that DA is involved in representing rewarded actions (e.g. Guitart-Masip et al., 2012; for a review see Guitart-Masip and Duzel, 2014).

On the other hand, it seems unlikely that the findings by Frank et al. (2004, 2007); and by our group (Kobza et al., 2012 and in the present study) on choice biases in PD patients can be fully explained in terms of a DA influence on action selection. First, the study by Smittenaar et al. (2012) found a bias for choosing beneficial stimuli in patients ON medication and not in controls, while the interaction between valence and medication (ON versus OFF) only approached significance. Second, convincing data on the effect of a lack of DA on action selection are still missing. While effects of a lack of DA on learning and action selection could not be disentangled due to the study design in the study by Shiner et al. (2012), they, as well as Smittenaar et al. (2012), did not find a significant effect on the choice bias in patients that were OFF medication during the transfer phase.

At the same time, the view of a critical role of DA in learning is supported by a great amount of animal studies (e.g. Andrzejewski et al., 2005; Clarke and Hill, 2011; Costa and Tran, 2015; Flagel et al., 2011; for review see Averbeck and Costa, 2017) and behavioural studies providing evidence for deficits in feedback learning in PD patients (Cools and Altamirano, 2006; Knowlton and Mangels, 1996; Myers et al., 2003). Finally, by applying the probabilistic selection task, our group found differences between groups of PD patients OFF medication concerning the positive/negative learning bias depending on how the patients learned during acquisition (actively or by observation, Kobza et al., 2012). This finding cannot be explained in terms of action selection. It is important to point out, however, that potential roles of DA in learning and action selection are not mutually exclusive and that potential interactions between the two effects need to be studied further.

Another potential problem for the interpretation of the present findings may be that patients and controls may have differed also in other factors as the lack of DA. It is known that PD also affects other neurotransmitter systems, as for example serotonin (Jellinger, 1991; Politis and Niccolini, 2015). Considering, for example, that PD patients frequently exhibit signs of depression (Marsh, 2013), a screening for depressive symptoms was performed. The screening revealed that the patients did not exhibit elevated depression scores. Nevertheless, an influence of a negative affective bias on the choice behaviour (Lemke et al., 2004; Thoma and Norra, 2015) can be widely, but not completely ruled out, as the patients were not systematically examined for psychiatric disorders.

Likewise, PD patients often exhibit different types of cognitive deficits. While a screening for signs of dementia did not reveal general cognitive impairments in our sample of patients, more specific problems such as executive dysfunctions were not tested and might have affected performance. Executive dysfunction might particularly affect their ability to transfer knowledge to a new task and thus the PD patients' performance during the transfer phase in this study. However, the current patient sample was not generally impaired in transfer phase performance, weakening this hypothesis.

Finally, it is conceivable that the sample size in the present study was too small to detect differences in the learning bias in patients and controls for immediate and delayed feedback. However, in previous studies samples of comparable size were tested, and in one study by our group we detected differential effects of DA depletion on the learning bias in two different types of learning (active vs. observational; Kobza et al., 2012). As was described in the Results section, the power for detecting a three-way interaction with the effect size of the study by Kobza et al. (2012) and the sample size of our study amounted to more than .8 and was thus quite high. Nevertheless, it cannot be excluded that with a larger sample the slight differences underlying the patterns of the two-way interactions between participants learning from immediate and delayed feedback would have resulted in a significant three-way interaction. Notwithstanding, the fact that significant two-way interactions between Learning Type (positive and negative) and Group (PD vs. controls) emerged for both feedback timing conditions separately shows that a bias towards stronger negative learning relative to control subjects is induced by a lack of DA for both immediate and delayed feedback.

In conclusion, the present study extends previous research on altered feedback learning in PD patients OFF medication when feedback is given immediately after a choice action (Frank et al., 2004, 2007; Kobza et al., 2012) by providing new evidence for a similar change when feedback is given after a short delay. For both feedback timings,

PD patients showed a different bias for positive/negative learning compared to healthy controls, with a stronger tendency towards negative learning. In those patients who learned from delayed feedback, this change in bias was seen despite overall spared learning performance, hinting at a nigro-striatal DA contribution to feedback learning, which affects how subjects learn from feedback, even if there is a temporal gap between action and outcome. Although it cannot be excluded that DA effects on action selection (Guitart-Masip et al., 2012, 2014; Shiner et al., 2012; Smittenaar et al., 2012) contributed to the present findings, our interpretation in terms of reduced but still measurable DA effects on delayed feedback learning is in line with many recent studies.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Andrzejewski, M.E., Spencer, R.C., Kelley, A.E., 2005. Instrumental learning, but not performance, requires dopamine D1-receptor activation in the amygdala. Neuroscience 135 (2), 335–345. http://dx.doi.org/10.1016/j.neuroscience.2005.06.038.

Arbel, Y., Hong, L., Baker, T.E., Holroyd, C.B., 2017. It's all about timing: an electrophysiological examination of feedback-based learning with immediate and delayed feedback. Neuropsychologia 99, 179–186. http://dx.doi.org/10.1016/j.neuropsychologia.2017.03.003.

Aubert, I., Ghorayeb, I., Normand, E., Bloch, B., 2000. Phenotypical characterization of the neurons expressing the D1 and D2 dopamine receptors in the monkey striatum. J. Comp. Neurol. 418 (1), 22–32. http://dx.doi.org/10.1002/(SICI)1096-9861(20000228)418:1<22::AID-CNE2>3.0.CO;2-Q.

Averbeck, B.B., Costa, V.D., 2017. Motivational neural circuits underlying reinforcement learning. Nat. Neurosci. 20 (4), 505–512. http://dx.doi.org/10.1038/nn.4506.

Bédard, P., Larochelle, L., Parent, A., Poirier, L.J., 1969. The nigrostriatal pathway: a correlative study based on neuroanatomical and neurochemical criteria in the cat and the monkey. Exp. Neurol. 25 (3), 365–377. http://dx.doi.org/10.1016/0014-4886(69)90131-9.

Bellebaum, C., Daum, I., 2008. Learning-related changes in reward expectancy are reflected in the feedback-related negativity. Eur. J. Neurosci. 27 (7), 1823–1835. http://dx.doi.org/10.1111/j.1460-9568.2008.06138.x.

Bellebaum, C., Rustemeier, M., Daum, I., 2011. Positivity effect in healthy aging in observational but not active feedback-learning. Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn. 19 (3), 402–420. http://dx.doi.org/10.1080/13825585.2011.629289.

Clarke, H.F., Hill, G.J., Robbins, T.W., Roberts, A.C., 2011. Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. J. Neurosci. 31 (11), 4290–4297. http://dx.doi.org/10.1523/JNEUROSCI.5066-10.2011.

Cools, R., Altamirano, L., D'Esposito, M., 2006. Reversal learning in Parkinson's disease depends on medication status and outcome valence. Neuropsychologia 44 (10), 1663–1673. http://dx.doi.org/10.1016/j.neuropsychologia.2006.03.030.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2001. Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. Cereb. Cortex 11 (12), 1136–1143. http://dx.doi.org/10.1093/cercor/11.12.1136.

Cools, R., Barker, R.A., Sahakian, B.J., Robbins, T.W., 2003. L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. Neuropsychologia 41 (11), 1431–1441. http://dx.doi.org/10.1016/S0028-3932(03)00117-9.

Costa, V.D., Tran, V.L., Turchi, J., Averbeck, B.B., 2015. Reversal learning and dopamine: a bayesian perspective. J. Neurosci. 35 (6), 2407–2416. http://dx.doi.org/10.1523/jneurosci.1989-14.2015.

Dickerson, K.C., Delgado, M.R., 2015. Contributions of the hippocampus to feedback learning. Cogn. Affect. Behav. Neurosci. 15 (4), 861–877. http://dx.doi.org/10.3758/s13415-015-0364-5.

Dickerson, K.C., Li, J., Delgado, M.R., 2011. Parallel contributions of distinct human memory systems during probabilistic learning. NeuroImage 55 (1), 266–276. http://dx.doi.org/10.1016/j.neuroimage.2010.10.080.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.G., 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41 (4), 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149.

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2), 175–191. 〈http://www.ncbi.nlm.nih.gov/pubmed/17695343〉.

Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akil, H., 2011. A selective role for dopamine in stimulus-reward learning. Nature 469 (7328), 53–57. http://dx.doi.org/10.1038/nature09588.

Foerde, K., Knowlton, B.J., Poldrack, R.A., 2006. Modulation of competing memory systems by distraction. Proc. Natl. Acad. Sci. USA 103 (31), 11778–11783. http://dx.doi.org/10.1073/pnas.0602659103.

Foerde, K., Race, E., Verfaellie, M., Shohamy, D., 2013. A role for the medial temporal lobe in feedback-driven learning: evidence from amnesia. J. Neurosci. 33 (13), 5698–5704. http://dx.doi.org/10.1523/JNEUROSCI.5217-12.2013.

Foerde, K., Shohamy, D., 2011. Feedback timing modulates brain systems for learning in humans. J. Neurosci. 31 (37), 13157–13167. http://dx.doi.org/10.1523/JNEUROSCI.2701-11.2011.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12 (3), 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6.

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. 17 (1), 51–72. http://dx.doi.org/10.1162/0898929052880093.

Frank, M.J., Samanta, J., Moustafa, A.A., Sherman, S.J., 2007. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science 318 (5854), 1309–1312. http://dx.doi.org/10.1126/science.1146157.

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. http://dx.doi.org/10.1126/science.1102941.

Gerfen, C.R., 1992. The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. Annu. Rev. Neurosci. 15, 285–320. http://dx.doi.org/10.1146/annurev.ne.15.030192.001441.

Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., … Movement Disorder Society, U. R. T. F, 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov. Disord. 23 (15), 2129–2170. http://dx.doi.org/10.1002/mds.22340.

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., Dolan, R.J., 2012. Action controls dopaminergic enhancement of reward representations. Proc. Natl. Acad. Sci. USA 109 (19), 7511–7516. http://dx.doi.org/10.1073/pnas.1202229109.

Guitart-Masip, M., Duzel, E., Dolan, R., Dayan, P., 2014. Action versus valence in decision making. Trends Cogn. Sci. 18 (4), 194–202. http://dx.doi.org/10.1016/j.tics.2014.01.003.

Haber, S.N., Fudge, J.L., 1997. The primate substantia nigra and VTA: integrative circuitry and function. Crit. Rev. Neurobiol. 11 (4), 323–342. http://dx.doi.org/10.1615/CritRevNeurobiol.v11.i4.40.

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 (6), 905–912. http://dx.doi.org/10.1111/j.1469-8986.2007.00567.x.

Hautzinger, M., Keller, F., Kühner, C., 2006. Das Beck Depressionsinventar II. Deutsche Bearbeitung und Handbuch zum BDI II. Frankfurt a. M.: Harcourt Test Services.

Hernandez-Lopez, S., Bargas, J., Surmeier, D.J., Reyes, A., Galarraga, E., 1997. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. J. Neurosci. 17 (9), 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., Surmeier, D.J., 2000. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC[beta]1-IP3-calcineurin-signaling cascade. J. Neurosci. 20 (24), 8987–8995.

Hoehn, M.M., Yahr, M.D., 1967. Parkinsonism: onset, progression, and mortality. Neurology 17 (5), 427–442.

Holroyd, C., 2004. A note on the oddball N200 and the feedback ERN. Neurophysiology 78, 447–455.

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

Holroyd, C.B., Coles, M.G., 2008. Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. Cortex 44 (5), 548–559. http://dx.doi.org/10.1016/j.cortex.2007.08.013.

Holroyd, C.B., Krigolson, O.E., 2007. Reward prediction error signals associated with a modified time estimation task. Psychophysiology 44 (6), 913–917. http://dx.doi.org/10.1111/j.1469-8986.2007.00561.x.

Holroyd, C.B., Krigolson, O.E., Baker, R., Lee, S., Gibson, J., 2009. When is an error not a prediction error? An electrophysiological investigation. Cogn. Affect. Behav. Neurosci. 9 (1), 59–70. http://dx.doi.org/10.3758/CABN.9.1.59.

Hughes, A.J., Daniel, S.E., Kilford, L., Lees, A.J., 1992. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J. Neurol. Neurosurg. Psychiatry 55 (3), 181–184. http://dx.doi.org/10.1136/jnnp.55.3.181.

Jakob, K., Ehrentreich, H., Holtfrerich, S.K.C., Reimers, L., Diekhof, E.K., 2018. DAT1-genotype and menstrual cycle, but not hormonal contraception, modulate reinforcement learning: preliminary evidence. Front. Endocrinol. 9, 60. http://dx.doi.org/10.3389/fendo.2018.00060.

Jeffreys, H., 1961. Theory of Probability, 3rd ed. Oxford University Press, Oxford, UK.

Jellinger, K.A., 1991. Pathology of Parkinsons-disease - changes other than the nigrostriatal pathway. Mol. Chem. Neuropathol. 14 (3), 153–197. http://dx.doi.org/10.1007/Bf03159935.

Knowlton, B.J., Mangels, J.A., Squire, L.R., 1996. A neostriatal habit learning system in humans. Science 273 (5280), 1399–1402.

Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., Sudmeyer, M., Bellebaum, C., 2012. Dissociation between active and observational learning from positive and negative feedback in Parkinsonism. PLos One 7 (11), e50250. http://dx.doi.org/10.1371/journal.pone.0050250.

Lavoie, B., Smith, Y., Parent, A., 1989. Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase immunohistochemistry. J. Comp. Neurol. 289 (1), 36–52. http://dx.doi.org/10.1002/cne.902890104.

Lee, M.D., Wagenmakers, E.J., 2013. Bayesian Cognitive Modeling: A Practical Course. Cambridge University Presshttp://dx.doi.org/10.1017/CBO9781139087759.

Lehéricy, S., Ducros, M., Van de Moortele, P.F., Francois, C., Thivard, L., Poupon, C., Kim, D.S., 2004. Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. Ann. Neurol. 55 (4), 522–529. http://dx.doi.org/10.1002/ana.20030.

Lemke, M.R., Fuchs, G., Gemende, I., Herting, B., Oehlwein, C., Reichmann, … J., Volkmann, J., 2004. Depression and Parkinson's disease. J. Neurol. 6, 24–27. http://dx.doi.org/10.1007/s00415-004-1606-6.

Lighthall, N.R., Gorlick, M.A., Schoeke, A., Frank, M.J., Mather, M., 2013. Stress modulates reinforcement learning in younger and older adults. Psychol. Aging 28 (1), 35–46. http://dx.doi.org/10.1037/a0029823.

Lind, M., Visentini, M., Mantyla, T., Del Missier, F., 2017. Choice-supportive misremembering: a new taxonomy and review. Front. Psychol. 8, 2062. http://dx.doi.org/10.3389/fpsyg.2017.02062.

Lynd-Balta, E., Haber, S.N., 1994. The organization of midbrain projections to the ventral striatum in the primate. Neuroscience 59 (3), 609–623. http://dx.doi.org/10.1016/0306-4522(94)90181-3.

Mammarella, N., Di Domenico, A., Fairfield, B., 2016. Aging and the genetic road towards the positivity effect in memory. Exp. Gerontol. 82, 120–124. http://dx.doi.org/10.1016/j.exger.2016.06.011.

Marsh, L., 2013. Depression and Parkinson's disease: current knowledge. Curr. Neurol. Neurosci. Rep. 13 (12), 409. http://dx.doi.org/10.1007/s11910-013-0409-5.

Mather, M., Carstensen, L.L., 2005. Aging and motivated cognition: the positivity effect in attention and memory. Trends Cogn. Sci. 9 (10), 496–502. http://dx.doi.org/10.1016/j.tics.2005.08.005.

Moustafa, A.A., Sherman, S.J., Frank, M.J., 2008. A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. Neuropsychologia 46 (13), 3144–3156. http://dx.doi.org/10.1016/j.neuropsychologia.2008.07.011.

Movement Disorder Society Task Force on Rating Scales for Parkinson's, D., 2003. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. Mov. Disord. 18 (7), 738–750. http://dx.doi.org/10.1002/mds.10473.

Myers, C.E., Shohamy, D., Gluck, M.A., Grossman, S., Kluger, A., Ferris, S., Schwartz, R., 2003. Dissociating hippocampal versus basal ganglia contributions to learning and transfer. J. Cogn. Neurosci. 15 (2), 185–193. http://dx.doi.org/10.1162/089892903321208123.

Otmakhova, N.A., Lisman, J.E., 1998. D1/D5 dopamine receptors inhibit depotentiation at CA1 synapses via cAMP-dependent mechanism. J. Neurosci. 18 (4), 1270–1279.

Peterburs, J., Kobza, S., Bellebaum, C., 2016. Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). Psychophysiology 53 (2), 209–215. http://dx.doi.org/10.1111/psyp.12560.

Politis, M., Niccolini, F., 2015. Serotonin in Parkinson's disease. Behav. Brain Res. 277, 136–145. http://dx.doi.org/10.1016/j.bbr.2014.07.037.

Samson, R.D., Frank, M.J., Fellous, J.M., 2010. Computational models of reinforcement learning: the role of dopamine as a reward signal. Cogn. Neurodyn. 4 (2), 91–105. http://dx.doi.org/10.1007/s11571-010-9109-x.

Schultz, W., 1997. Dopamine neurons and their role in reward mechanisms. Curr. Opin. Neurobiol. 7 (2), 191–197. http://dx.doi.org/10.1016/S0959-4388(97)80007-4.

Schultz, W., 2000. Multiple reward signals in the brain. Nat. Rev. Neurosci. 1 (3), 199–207. http://dx.doi.org/10.1038/35044563.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275 (5306), 1593–1599. http://dx.doi.org/10.1126/science.275.5306.1593.

Sheth, S.A., Mian, M.K., Patel, S.R., Asaad, W.F., Williams, Z.M., Dougherty, D.D., Eskandar, E.N., 2012. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. Nature 488 (7410), 218–221. http://dx.doi.org/10.1038/nature11239.

Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K.P., Dayan, P., Dolan, R.J., 2012. Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease. Brain 135 (Pt 6), 1871–1883. http://dx.doi.org/10.1093/brain/aws083.

Shohamy, D., Adcock, R.A., 2010. Dopamine and adaptive memory. Trends Cogn. Sci. 14 (10), 464–472. http://dx.doi.org/10.1016/j.tics.2010.08.002.

Smittenaar, P., Chase, H.W., Aarts, E., Nusselein, B., Bloem, B.R., Cools, R., 2012. Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection–learning or performance? Eur. J. Neurosci. 35 (7), 1144–1151. http://dx.doi.org/10.1111/j.1460-9568.2012.08043.x.

Thoma, P., Norra, C., Juckel, G., Suchan, B., Bellebaum, C., 2015. Performance monitoring and empathy during active and observational learning in patients with major depression. Biol. Psychol. 109, 222–231. http://dx.doi.org/10.1016/j.biopsycho.2015.06.002.

Wagenmakers, E.J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Morey, R.D., 2017a. Bayesian inference for psychology. Part II: example applications with JASP. Psychon. Bull. Rev. http://dx.doi.org/10.3758/s13423-017-1323-7.

Wagenmakers, E.J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Morey, R.D., 2017b. Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. Psychon. Bull. Rev. http://dx.doi.org/10.3758/s13423-017-1343-3.

Weinberg, A., Luhmann, C.C., Bress, J.N., Hajcak, G., 2012. Better late than never? The effect of feedback delay on ERP indices of reward processing. Cogn. Affect. Behav. Neurosci. 12 (4), 671–677. http://dx.doi.org/10.3758/s13415-012-0104-z.

Weismüller, B., Bellebaum, C., 2016. Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. Psychophysiology 53 (11), 1739–1750. http://dx.doi.org/10.1111/psyp.12738.

Psychophysiology, 53 (2016), 1739–1750. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/psyp.12738

Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning

BENJAMIN WEISMÜLLER AND CHRISTIAN BELLEBAUM

Institute for Experimental Psychology, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany

Abstract

Learning from feedback is a prerequisite for adapting to the environment. Prediction error signals coded by midbrain dopamine (DA) neurons are projected to the basal ganglia and anterior cingulate cortex (ACC). It has been suggested that neuronal activity shifts away from the DA system when feedback is delayed. The feedback-related negativity (FRN), an ERP that is generated in the ACC and has been shown to be sensitive to feedback valence and prediction error magnitude, was found to be reduced for delayed feedback. It has, however, not yet been investigated if the FRN for delayed feedback reflects a reward prediction error. In this study, effects of feedback delay (1 s vs. 7 s) on the processing of expected and unexpected positive and negative feedback were investigated in a between-subjects design in healthy human participants conducting a probabilistic feedback learning task. FRN and P300 amplitudes were decreased for subjects learning from delayed compared to immediate feedback. Importantly, the FRN, extracted from the negative-positive feedback difference wave, was significantly smaller for expected compared to unexpected feedback for both the immediate and delayed feedback conditions. Expectancy effects for the P300 were also seen, but did not interact with feedback valence. These results demonstrate an influence of feedback expectancy, and thus the prediction error, on early feedback processing even for delayed feedback, suggesting that neuronal structures underlying feedback processing are comparable for immediate and delayed feedback, at least to some extent. Modulations of the P300 by feedback delay may be linked to feedback salience.

Descriptors: EEG, Performance monitoring, Learning, Feedback delay, Expectancy, FRN

Humans face decisions on how to act or behave every day. These decisions can result in positive or negative consequences such as reward or punishment, which can serve as feedback to guide future action selection in order to receive maximal positive outcome.

Animal studies showed that the dopamine (DA) system is critically involved in feedback-based learning. Phasic bursts and dips of midbrain DA neuron firing are associated with unexpected positive and negative feedback, respectively (Bayer & Glimcher, 2005; Schultz, 1997, 1998, 2000; Schultz, Dayan, & Montague, 1997), and enable the basal ganglia, which receive DA projections (Bédard, Larochelle, Parent, & Poirier, 1969; Lavoie, Smith, & Parent, 1989; Lynd-Balta & Haber, 1994) to evaluate ongoing events (Barto, 1995; Montague, Dayan, & Sejnowski, 1996; for review, see Schultz, 2002). Furthermore, DA neurons project to the striatum and frontal cortex including the anterior cingulate cortex (ACC; Lehéricy et al., 2004; see Haber & Fudge, 1997), which is thus also involved in reward processing (e.g., Delgado, 2007; Delgado, Locke, Stenger, & Fiez, 2003; for review, see Knutson & Cooper, 2005; McClure, Berns, & Montague, 2003; O'Doherty et al., 2004).

ERP studies have associated the so-called feedback-related negativity (FRN) with processing in the ACC (Bellebaum & Daum, 2008; Gehring & Willoughby, 2002; Holroyd & Coles, 2002; for review, see Nieuwenhuis, Holroyd, Mol, & Coles, 2004; Walsh & Anderson, 2012). The FRN is an ERP component that peaks at about 250 ms postfeedback and is typically larger for negative compared to positive feedback (Gehring & Willoughby, 2002; Hajcak, Moser, Holroyd, & Simons, 2006; Hajcak, Moser, Yeung, & Simons, 2005; Holroyd & Coles, 2002; Holroyd, Hajcak, & Larsen, 2006; Holroyd & Krigolson, 2007; Miltner, Braun, & Coles, 1997; for review, see Simons, 2010; Walsh & Anderson, 2012). In accordance with the reinforcement learning (RL) theory (Holroyd & Coles, 2002) and with the mentioned DA inputs to the ACC, the FRN appears to reflect reward expectancy and thus prediction errors. While some researchers found evidence for generally higher amplitudes for unexpected compared to expected negative and positive feedback (e.g., Ferdinand, Mecklinger, Kray, & Gehring, 2012; Oliveira, McDonald, & Goodman, 2007), others focused on expectancy effects on the difference FRN, that is, the amplitude difference between negative and positive feedback (e.g., Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd & Coles, 2002; Holroyd & Krigolson, 2007; Yasuda, Sato, Miyawaki, Kumano, &

This research was financed by the Department of Biological Psychology, Heinrich-Heine University Düsseldorf. We would like to thank Denis Breuer and Jonathan Hagemeister for their help in EEG data acquisition and analysis.

Address correspondence to: Benjamin Weismüller, Institute for Experimental Psychology, Department of Biological Psychology, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany. E-mail: Benjamin.Weismueller@hhu.de

PSYCHOPHYSIOLOGY

Kuboki, 2004). Some earlier studies had yielded contradictory findings, without effects of reward expectation on the difference FRN (Hajcak, Holroyd, Moser, & Simons, 2005; see also Donkers, Nieuwenhuis, & van Boxtel, 2005; Larson, Kelly, Stigge-Kaufman, Schmalfuss, & Perlstein, 2007; for discussion, see Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; for review, see Walsh & Anderson, 2012). In this context, Hajcak et al. (2007) pointed out that the participants' subjective reward expectation needs to be taken into account and that the variation of objective reward probability does not necessarily suffice in inducing different reward expectations. Another important factor contributing to feedback processing is whether feedback can actually be used to optimize behavior (Holroyd et al., 2009). Effects of feedback expectancy are stronger in learning tasks than in gambling tasks in which the relative frequency of positive and negative feedback is predetermined.

Most ERP studies on feedback processing have used immediate feedback or at least very short response feedback intervals of 1 s or less (e.g., Bellebaum & Daum, 2008; Ferdinand et al., 2012; Hajcak, Holroyd, et al., 2005; Hajcak et al., 2007; Holroyd & Coles, 2002; Holroyd et al., 2009). Findings from an fMRI study (Foerde & Shohamy, 2011) suggested distinct neuronal mechanisms for the processing of immediate and delayed feedback. While prediction error processing for immediate feedback (delay of 1 s) recruited the dorsal striatum more strongly than for delayed feedback (7 s), the opposite pattern was found in the hippocampus (HC), where the prediction error for delayed feedback was more strongly represented. These findings were confirmed by evidence obtained in brain-damaged patients (Foerde, Race, Verfaellie, & Shohamy, 2013; Foerde & Shohamy, 2011). Parkinson patients suffering from striatal dysfunction were impaired in learning from immediate but not from delayed feedback, while patients with amnesia and suspected HC damage showed the opposite pattern. Foerde and Shohamy (2011) and Foerde et al. (2013) proposed that the HC may bind related elements across time and is therefore recruited when feedback is delayed. The medial temporal lobe (MTL), in particular the HC, has been associated with declarative learning, while the striatum is thought to underlie nondeclarative feedback-based learning (Knowlton, Mangels, & Squire, 1996; Sherry & Schacter, 1987; Squire, 1992; Squire & Zola, 1996). These distinct systems have been suggested to interact in a competitive manner (e.g., Poldrack et al., 2001; Poldrack & Packard, 2003), but recent studies provided evidence for parallel contributions of both systems during learning (Dickerson & Delgado, 2015; Dickerson, Li, & Delgado, 2011).

Only few studies examined the effect of different delays on ERP correlates of feedback processing. The difference FRN amplitude has been demonstrated to be diminished for delayed compared to immediate feedback (Opitz, Ferdinand, & Mecklinger, 2011; Peterburs, Kobza, & Bellebaum, 2015; Weinberg, Luhmann, Bress, & Hajcak, 2012; but see Wang, Chen, Lei, & Li, 2014, for a negative finding), whereby it is important to note that the experimental procedures and the delays used varied considerably between studies. This finding is compatible with the above described evidence on stronger involvement of the striatum for immediate than delayed feedback, as DA projections target both the striatum and the ACC. Furthermore, this is in line with recent studies linking the FRN to processing in the striatum (Becker, Nitsch, Miltner, & Straube, 2014; Foti, Weinberg, Dien, & Hajcak, 2011). However, as was demonstrated by our group, the difference FRN is not absent during delayed feedback processing (Peterburs et al., 2015). At the same time, a close look at the fMRI findings by Foerde and Shohamy (2011) reveals that both immediate and delayed feedback processing recruit the ventral striatum. Thus, the striatal/ACC system might also be involved in delayed feedback processing. At the same time, DA innervation can also be found in the HC where it influences hippocampal plasticity (Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010), suggesting that prediction error-related information in the HC, as in the striatum, originates in DA neurons. Together, these findings indicate that the processing mechanisms may be comparable for immediate and delayed feedback to some extent, so that the FRN for delayed feedback may also be modulated by the prediction error, as for immediate feedback. Effects of reward expectancy and thus the prediction error on the FRN following delayed feedback have, however, not been examined to date.

A recent study also found an influence of feedback delay on the P300 amplitude. The P300 is associated with attentional resource allocation, being larger for unexpected compared to expected target stimuli (Duncan-Johnson & Donchin, 1977; for review, see Polich, 2007), including performance feedback (Bellebaum & Daum, 2008). However, the influence of feedback delay is still under debate. Opitz et al. (2011) reported an increased P300 for delayed compared to immediate negative feedback, while Wang et al. (2014) reported the opposite pattern.

This study aimed to shed light onto the influence of feedback expectancy on the processing of delayed feedback by means of ERPs in healthy human subjects. Two groups of participants performed a probabilistic learning task, one with immediate and the other with delayed feedback. To induce different reward expectations, subjects could choose between stimuli with different objective reward probabilities. In addition, subjective reward expectations were assessed on a trial-to-trial basis in order to confirm that the different objective reward probabilities resulted in different subjective estimations of reward probabilities. In a first step, we wanted to confirm the previous finding that delayed feedback elicits smaller amplitudes of the difference FRN than immediate feedback. We then hypothesized that, despite the amplitude reduction, difference FRN amplitudes would be larger for unexpected than expected delayed feedback. Furthermore, effects of feedback delay and expectancy on the P300 were analyzed.

Method

Study Participants

Fifty students of the Heinrich-Heine University Düsseldorf were recruited and randomly assigned to one of two groups, the immediate and the delayed feedback group (see below for details). Mean age of the 25 participants in the immediate feedback group was 24.5 years (range 19–38, $SD = 4.3$; 10 males, 15 females). In the 25 participants of the delayed feedback group, mean age was 24.8 years (range 18–35, $SD = 5.1$; 10 males, 15 females). Exclusion criteria were self-reported history of psychiatric or neurological disorders, traumatic brain injury with sustained unconsciousness in the past, and consumption of alcohol or psychodynamic drugs within the last 24 h or on a regular basis. All participants had normal or corrected-to-normal vision. The study was approved by the Ethics Committee of the Faculty of Mathematics and Natural Sciences at Heinrich-Heine University Düsseldorf and was in accordance with the Declaration of Helsinki. Study participants were recruited by advertisement. They participated voluntarily and gave informed written consent before taking part in the experiment. Financial reimbursement or course credit was offered for participation in addition to monetary reward from the learning task (see below).

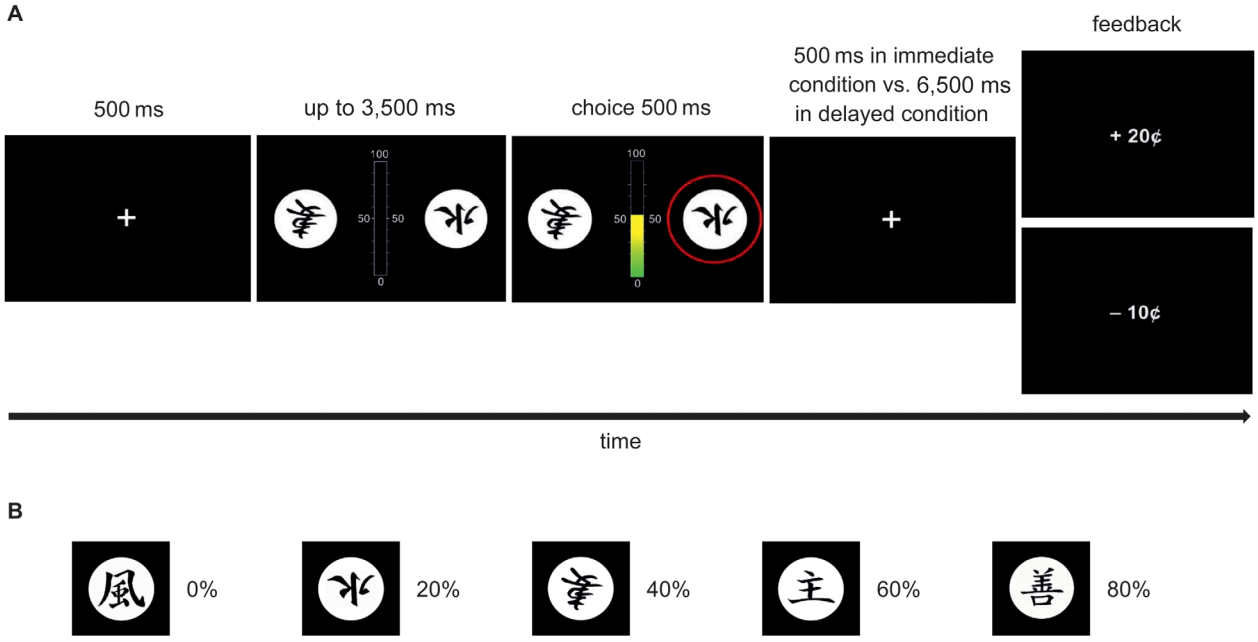

Figure 1. A: Time course of events in a single learning trial. When participants chose a symbol by pressing a button, they also indicated their subjective reward probability by holding the button. After that, the fixation cross was presented for either 500 ms in the immediate feedback condition or for 6,500 ms in the delayed feedback condition. Finally, monetary reward or punishment was presented. B: Symbols and objective reward probabilities used in the experiment.

Probabilistic Learning Task

Participants in the learning task were instructed (a) to try to receive as much money as possible in the form of monetary reward for choices between different symbols, and (b) to indicate their subjective reward probability for each choice. On each trial, participants were asked to choose one of two visual stimuli presented on the left and right side of the screen by pressing a button in order to receive monetary feedback, that is, reward (+20¢) or punishment (−10¢). By using a higher positive than negative feedback value, we ensured that participants always received a net gain at the end of the experiment. Furthermore, it has been shown that losses are subjectively weighed about twice as high as gains (Kahneman, Knetsch, & Thaler, 1991; Tversky & Kahneman, 1991). Five different Japanese symbols served as stimuli, presented in all 10 possible pair combinations, with the assignment of stimuli to the left and right side of the screen being counterbalanced for each pair. Unknown to the participant, symbol-locked reward probabilities were 0%, 20%, 40%, 60%, and 80% (referred to as objective reward probabilities in the following). These symbol-specific probabilities were kept constant throughout the experiment. Choices had to be made within 3,500 ms after symbol presentation, and a red circle around the chosen symbol indicated the participant's choice (see Figure 1 for details on the sequence of events in one trial and for the stimuli).

In addition, a scale ranging from 0 to 100% was presented between the symbols in the middle of the screen. With the help of this scale, participants were asked to indicate their subjective reward expectation when they made their choice (referred to as subjective probability, see above). The scale was manipulated by holding the button the participants pressed to choose a symbol. As the process of indicating the subjective reward probability took some time (maximally 1,600 ms), the scale either started at 0% and could be moved toward 100% by holding the button (scale up), or it started at 100% and holding moved it toward 0% (scale down) to exclude a confound of delay (between button press onset and feedback) and subjective probability. After releasing the button, the scale froze for 500 ms indicating the participants' choice and their final subjective reward expectation (i.e., the value at which the bar stopped when the participant stopped holding the button). Then, a fixation cross was presented until the positive or negative feedback stimulus came on. Importantly, this feedback was presented after an additional 500 ms in the immediate feedback condition and after 6,500 ms in the delayed feedback condition. This sums up to 1-s feedback delay (from button press offset) in the immediate and 7-s delay in the delayed feedback condition. Each participant completed only one version of the experiment (with immediate or delayed feedback).

In total, participants completed five blocks. The number of blocks with a rising or decreasing subjective probability scale (two or three of each type) was varied across participants. Each block included 100 trials, 10 for each of the 10 symbol pairs. Between the blocks, participants were allowed to have short breaks. In the immediate condition, one block took approx. 8–9 min. In the delayed condition, a block took approximately 18–20 min. After the experiment, participants received the sum of money they earned during their most successful block, which amounted to maximally 8€. To increase motivation throughout the experiment, participants were instructed at the beginning that their performance would result in real monetary reward and that they would be paid out the sum of their most successful block.

Procedure

Before EEG recording, a structured interview was conducted with each participant in which demographic data, medication, and previous illnesses were assessed. For the EEG session, participants were comfortably seated about 60 cm away from a computer screen of 24-inch diagonal dimension (ASUS VG248QE Full HD). They were asked to not move, to relax, and to avoid blinks as far as

Table 1. *Mean Number of Trials Averaged and Rejected Trials for Each Feedback Condition*

		High expectancy		Low expectancy	
Learning condition		Positive	Negative	Positive	Negative
Immediate feedback	Trials in averages	165.32 (44.17)	82.67 (21.38)	40.11 (12.35)	60.26 (19.76)
	Rejected trials	17.05 (28.87)	9.58 (15.09)	4.05 (6.51)	7.32 (10.45)
Delayed feedback	Trials in averages	159.00 (48.19)	78.53 (25.82)	38.06 (10.49)	63.06 (24.02)
	Rejected trials	6.29 (8.14)	2.18 (3.30)	0.82 (1.42)	2.64 (4.12)

Note. Standard deviations appear in parentheses.

possible. After the task was explained, the participants conducted 10 test trials. Testing started when the task was completely understood and participants were sufficiently trained to handle the subjective probability bar (both the rising and the decreasing variant). EEG recordings lasted between about 40–45 min in the immediate condition and 95–100 min in the delayed feedback condition.

EEG Recording

Continuous EEG data were recorded from electrode positions Fz, F3, F7, F4, F8, FCz, FC3, FC7, FC4, FC8, Cz, C3, C4, CPz, CP3, CP4, Pz, P3, P7, P4, P8, POz, PO3, PO7, PO4, PO8, T7, T8, and AFz against nose reference (according to the International 10-20 system) using BrainVision Recorder software, version 2.0 (Brain Products, Munich, Germany). Vertical eye movements and blinks were recorded with electrodes above and below the right eye; horizontal eye movements were recorded at electrode sites F9 and F10. Impedances were below 5 kΩ. Data were continuously sampled at 1000 Hz.

Data Analysis

Behavioral data. Choice accuracy was derived from the mean number of correct trials in each block in order to investigate learning performance. In accordance with the procedure in previous studies (Bellebaum et al., 2016; Foerde et al., 2013; Frank, Seeberger, O'Reilly, 2004; Holroyd et al., 2009; Knowlton et al., 1996), a choice was considered correct when participants chose the symbol with the higher reward probability (irrespective of the feedback they actually received on that trial). The number of correct responses was pooled across all trial types and averaged for each block.

ERPs. EEG data were analyzed offline by means of BrainVision Analyzer software, version 2.0 (Brain Products) and MATLAB *R2013*a (MathWorks, Natick, MA). Data were 40 Hz low-pass and 0.5 Hz high-pass filtered. In a first step, data were segmented from 250 ms before to 2,250 ms after feedback stimulus onset and baseline corrected using the mean amplitude in the 250 ms preceding feedback stimulus onset. In order to prepare the data for ocular correction (see below), automatic artifact rejection excluded any epochs with data points exceeding an absolute amplitude value of 150 μV or exceeding an amplitude difference of 200 μV between the lowest and the highest data point, ignoring frontal and frontocentral electrode sites. The data of all electrodes in the remaining segments were corrected for eye movement and blink artifacts using the algorithm described by Gratton, Coles, and Donchin (1983). Data then were segmented from 200 ms before to 800 ms after feedback stimulus onset for each experimental condition, which were defined based on the combination of the chosen symbols' objective reward probability and feedback valence. Trials in which participants chose stimuli with a relatively high objective reward probability of 60% or 80% were pooled, as well as trials in which participants chose low reward probability stimuli (20% or 40%). The stimuli with 0% reward probability (see above) were important, as they made participants also choose the stimulus with the very low reward probability of 20%, when the two stimuli were paired. Rewarded choices of stimuli with a high reward probability were considered as expected positive feedback, whereas unrewarded choices of such stimuli were considered to yield unexpected negative feedback. Accordingly, rewarded choices of stimuli with low reward probability yielded unexpected positive feedback, whereas unrewarded choices led to expected negative feedback. Note that, due to the combination of all potential reward probabilities, choices of stimuli with 20% or 40% reward probability could reflect correct responses, for example, when these stimuli were presented together with a 0% reward probability stimulus. After another baseline correction relative to the 200-ms baseline interval, segments containing data points exceeding an absolute amplitude value of 100 μV or in which the amplitude difference between the highest and lowest data point exceeded 100 μV were excluded (Table 1 lists the mean absolute numbers of included and rejected trials). Finally, average ERPs were computed for each of the four conditions (expected and unexpected positive and negative feedback).

According to the definition of the conditions, participants expected certain types of outcomes more or less strongly. This notion is only true for those subjects who have gained insight into reward probabilities of the different stimuli. We therefore expected to find effects of reward probability (and thus feedback expectancy) only in those participants who reached a certain learning criterion. Participants were considered as learners if they reached a choice accuracy of at least 65% per block (pooled across all trial types) and were able to keep up this level on average in at least two consecutive blocks. In these subjects, only those trials were considered for the analysis after the criterion was reached (see online supporting information for a qualitative comparison with data of the nonlearners).

In line with the previous literature on delay effects (Peterburs et al., 2015; Wang et al., 2014; Weinberg et al., 2012), we focused on the punishment– reward difference waves in our FRN analysis, which were computed separately for expected and unexpected outcomes. For each participant, the FRN amplitude was defined as the maximum negative peak in the difference wave in the 180–350 ms time window after feedback onset at electrode site FCz. This electrode was chosen because topographic maps of difference wave amplitudes showed that the difference FRN was very pronounced over frontocentral cortex for both immediate and delayed feedback (see Results).

For the P300 amplitude, visual inspection suggested that the latency of the component differed between conditions. At the same time, individual subjects' ERPs did not always show clear peak amplitudes. We thus decided to use a mean amplitude approach based on the peaks of the grand averages of the different

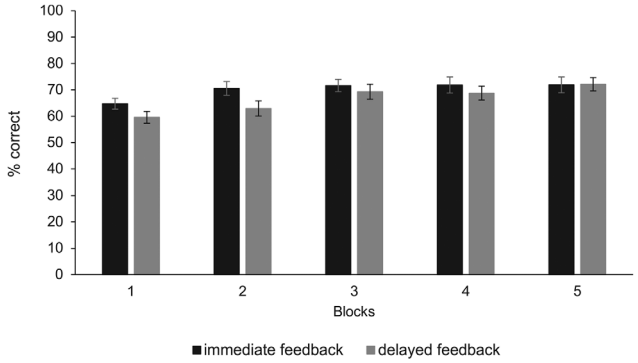

Figure 2. Choice accuracy (i.e., number of correct responses), pooled across trial types, in percent across blocks (error bars indicate standard errors).

conditions. For each timing and feedback valence condition, we extracted the maximum positive peak of the grand average in the time window between 325 and 525 ms after stimulus onset in a first step. Then, each participant's mean amplitude in a 100-ms time window centered around the respective grand average peak was extracted for every experimental condition. The P300 was analyzed at FCz and Pz electrode sites.

Statistical analyses. Statistical analyses were performed with SPSS Statistics 23 (IBM, Armonk, NY). Choice accuracy was analyzed with an analysis of variance (ANOVA) with block (1 to 5) as within-subject factor and feedback timing (immediate vs. delay) as between-subjects factor.

Delay effects on FRN peak amplitudes and latencies were analyzed by means of an ANOVA with the within-subject factor expectancy (low vs. high) and the between-subjects factor feedback timing (immediate vs. delayed). P300 amplitudes were analyzed by an ANOVA with the within-subject factors expectancy (low vs. high), valence (positive vs. negative), and electrode (FCz vs. Pz) and the between-subjects factor feedback timing (immediate vs. delayed). For all analyses, Greenhouse-Geisser correction was performed when sphericity was violated. An alpha level of $p < .05$ (two-sided) was accepted as statistically significant.

Results

Behavioral Data

Figure 2 shows the choice accuracy, pooled across trial types, in percent across blocks. Statistical analysis of choice accuracy revealed a significant main effect of block, $F(4,192) = 12.444$; $p < .001$; $\eta_p^2 = .206$. The ANOVA revealed neither a significant effect of feedback timing, $F(1,48) = 1.227$; $p = .273$; $\eta_p^2 = .025$, nor a significant interaction between block and feedback timing, $F(4,192) = 1.763$; $p = .160$; $\eta_p^2 = .035$, demonstrating that participants of both groups learned to choose the stimuli with higher reward probability similarly well. Controlling for multiple comparisons, post hoc paired t tests for the block main effect revealed that accuracy scores were significantly higher for Blocks 3, 4, and 5 than Block 1 (all $t(49) < -2.670$; all $p < .001$). Furthermore, accuracy was higher in Block 5 than Block 2, $t(49) = -3.201$; $p = .002$.

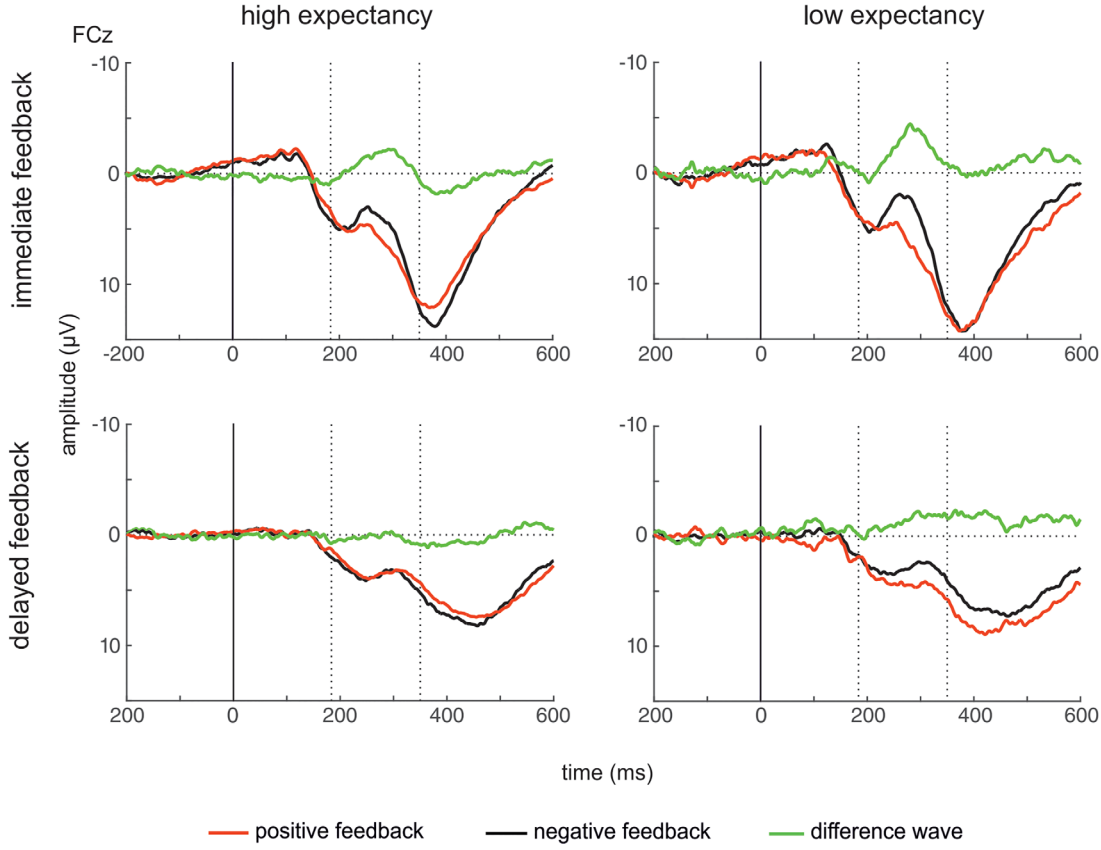

Figure 3. Feedback-locked grand-averaged ERPs for positive feedback, negative feedback, and the negative-positive difference wave at FCz. Dotted vertical lines indicate the time window used to extract FRN peaks.

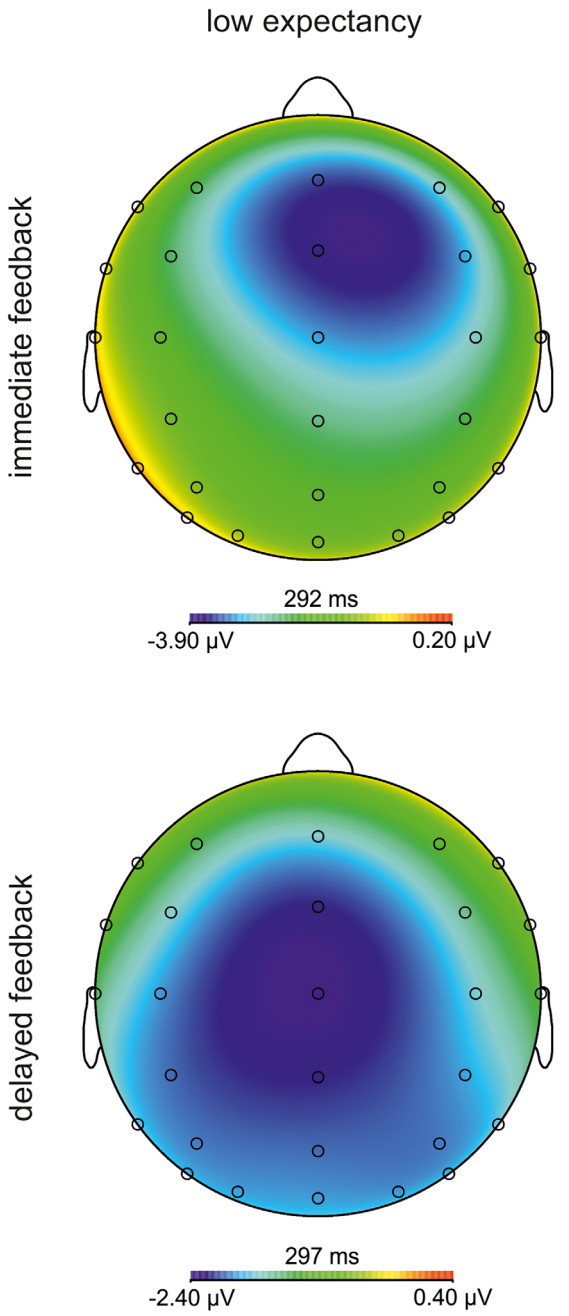

Figure 4. Topographic maps of the scalp distribution of the difference FRN peak for unexpected outcomes in the immediate and delayed feedback conditions.

ERP Data

Fourteen participants (six in the immediate and eight in the delayed feedback timing condition) were excluded because they did not reach the learning criterion. The final sample for EEG analysis thus consisted of 19 participants in the immediate (mean age 24.3, range 19–38, $SD = 4.6$; 8 males, 9 females) and 17 participants in the delayed condition (mean age 25.2, range 18–35, $SD = 4.7$; 8 males, 11 females).

FRN. Figure 3 shows feedback-locked ERPs and difference waves in both experimental groups. Figure 4 shows topographic maps for the difference FRN amplitudes in the unexpected feedback condition for both immediate and delayed feedback. As can be seen, the region of the largest relative negativity for negative feedback includes (fronto)central scalp sites for both delay conditions. Table 2 lists the mean amplitudes of the difference FRN. These were significantly higher in the unexpected than expected feedback condition, $F(1,34) = 28.425$; $p < .001$; $\eta_p^2 = .455$. Furthermore, difference FRN amplitudes were generally higher for the immediate feedback compared to the delayed feedback condition, $F(1,34) = 6.322$; $p = .017$; $\eta_p^2 = .157$. No significant interaction of Expectancy × Feedback Timing was found, $F(1,34) = 0.121$; $p = .731$; $\eta_p^2 = .004$. To examine if the effect of expectancy was significant in the two groups (and thus feedback timing conditions) separately, exploratory paired t tests were conducted. Difference FRN amplitudes were significantly higher for unexpected compared to expected feedback in both the immediate feedback condition, $t(18) = -3.385$; $p = .003$, and delayed feedback condition, $t(16) = -4.280$; $p = .001$.

Table 2. *FRN Difference Wave Amplitudes in µV*

Learning condition	High expectancy	Low expectancy
Immediate feedback	−4.08 (3.23)	−6.28 (3.61)
Delayed feedback	−1.84 (1.27)	−4.34 (2.37)

Note. Standard deviations appear in parentheses.

Table 3 lists mean difference FRN latencies. Statistical analysis revealed neither a main effect of expectancy, $F(1,34) = 2.114$; $p = .155$; $\eta_p^2 = .059$, nor of feedback timing, $F(1,34) = 0.045$; $p = .834$; $\eta_p^2 = .001$. No interaction was found between expectancy and feedback timing, $F(1,34) = 0.063$; $p = .804$; $\eta_p^2 = .002$.

To investigate whether the expectancy effect on the difference FRN amplitude was caused by the processing of positive or negative feedback or both, we extracted mean amplitudes of the original waveforms from 20-ms time windows centered around each individual subject's difference FRN peak latency in the expected and unexpected feedback conditions. If, for example, one particular subject had a difference FRN peak latency of 250 ms in the unexpected feedback condition, mean amplitudes between 240 and 260 ms were extracted for unexpected positive and negative feedback, respectively. The resulting mean amplitudes were analyzed in a 2 × 2 × 2 repeated measures ANOVA, again with feedback timing (immediate vs. delayed) as a between-subjects factor and feedback expectancy (expected vs. unexpected) as within-subject factor and the additional within-subject factor feedback valence (positive vs. negative). In accordance with the results of the difference wave analysis, we found a significant interaction of feedback valence and expectancy, $F(1,34) = 20.704$; $p < .001$; $\eta_p^2 = .378$, which was also significant in each feedback timing condition separately (for immediate feedback, $F(1,18) = 7.879$; $p = .012$; $\eta_p^2 = .304$; for delayed feedback, $F(1,16) = 15.387$; $p = .001$; $\eta_p^2 = .490$). The resolution of these interactions showed that the amplitudes were higher for unexpected positive compared to expected positive feedback in the whole group of participants, $t(35) = 3.538$; $p = .001$, as well as in each group separately (immediate feedback,

Table 3. *FRN Peak Latencies in ms*

Learning condition	High expectancy	Low expectancy
Immediate feedback	281 (42)	292 (35)
Delayed feedback	281 (48)	297 (42)

Note. Standard deviations appear in parentheses.

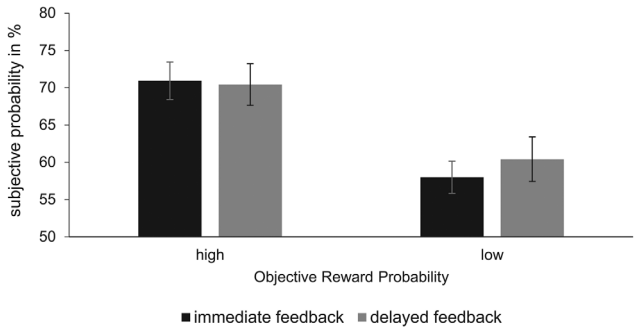

Figure 5. Subjective reward probability in percent, separately for trials in which subjects chose stimuli with higher or lower objective reward probability (error bars indicate standard errors).

$t(18) = 2.344$; $p = .031$; delayed feedback, $t(16) = 2.596$; $p = .019$). For negative feedback, however, no significant differences were found between the amplitudes for unexpected and expected feedback for the total group of participants, $t(35) = -0.615$; $p = .543$, or the feedback timing conditions separately (immediate feedback, $t(18) = -0.555$, $p = .586$; delayed feedback, $t(16) = -0.275$; $p = .787$).

As outlined above, the different expectancy conditions were based on participant's choices of stimuli with different objective reward probabilities. As task difficulty was unequally distributed across the symbol combinations, a potential confound between expectancy and difficulty could not be excluded. We thus reanalyzed the data, excluding easy symbol pairs that combine high and low reward probabilities (i.e., 80% vs. 0%, 80% vs. 20%, 60% vs. 0%, and 60% vs. 20%), so that all outcomes entering the analysis of feedback-locked ERPs now stemmed from choices in difficult trials (for details, see supporting information). The difference FRN still showed a clearly higher amplitude for unexpected compared to expected feedback in both feedback timing conditions. In this analysis, we could not replicate the effect of feedback delay on the difference FRN. Generally, this analysis has to be interpreted with caution, however, as the number of trials entering the analysis dropped below 15 for several subjects in one or two conditions.

Finally, in order to make sure that the processing differences between the groups receiving immediate or delayed feedback were not caused by differences in subjective reward probabilities, mean subjective probabilities were analyzed for the low and high objective reward probability conditions, on which the definition of the ERP conditions was based (i.e., 20–40% vs. 60–80% reward probability). Figure 5 shows the mean subjective reward probability ratings for the different objective reward probability conditions and in the two feedback timing groups. Although subjects overestimated the lower reward probabilities, ANOVA with probability (low vs. high) as within-subject factor and feedback timing as between-

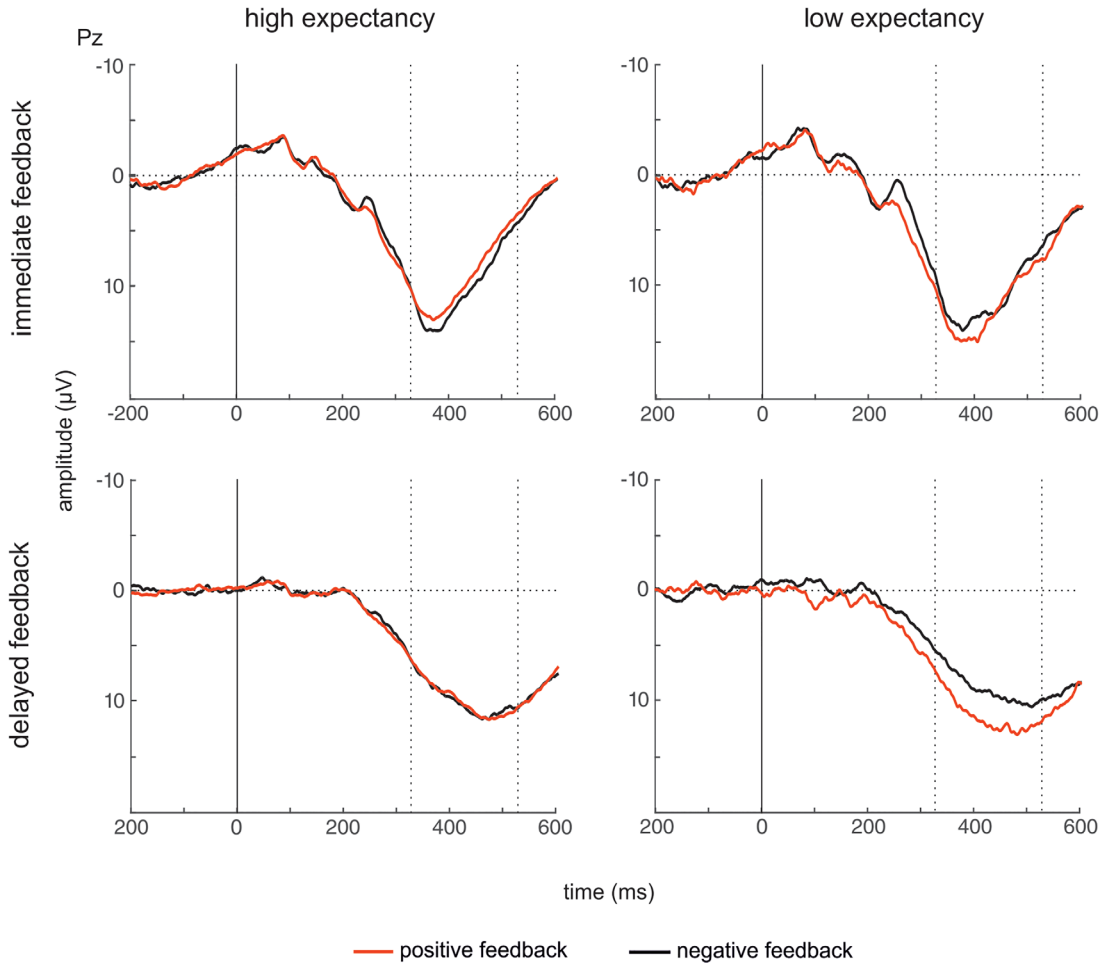

Figure 6. Feedback-locked grand-averaged ERPs for positive and negative feedback at Pz. Dotted vertical lines indicate the time window used to extract P300 peaks.

Table 4. *P300 Peak Amplitudes in μV*

Learning condition		High expectancy		Low expectancy	
		Positive	Negative	Positive	Negative
FCz	Immediate feedback	10.96 (4.77)	11.99 (3.92)	12.79 (4.62)	12.47 (4.39)
	Delayed feedback	7.14 (3.34)	7.60 (4.35)	8.28 (3.73)	6.77 (4.03)
Pz	Immediate feedback	11.83 (5.49)	12.74 (4.14)	13.90 (4.27)	12.76 (5.39)
	Delayed feedback	11.15 (4.50)	11.02 (5.62)	12.38 (5.32)	9.95 (4.77)

Note. Standard deviations appear in parentheses.

subjects factor yielded a significant main effect of probability, $F(1,32) = 59.848$; $p < .001$; $\eta_p^2 = .652$. Subjectively rated reward probability was indeed higher for stimuli with high compared to low objective reward probability. Ratings did, however, not significantly differ between feedback timing conditions, $F(1,32) = 0.077$; $p = .783$; $\eta_p^2 = .002$. Also, no significant interaction of Probability × Feedback Timing was found, $F(1,32) = 0.974$; $p = .331$; $\eta_p^2 = .030$, suggesting that subjective reward expectations were not modulated by feedback timing.

P300. Figure 6 shows feedback-locked ERPs at electrode Pz. Table 4 lists the mean values of P300 amplitudes for all experimental

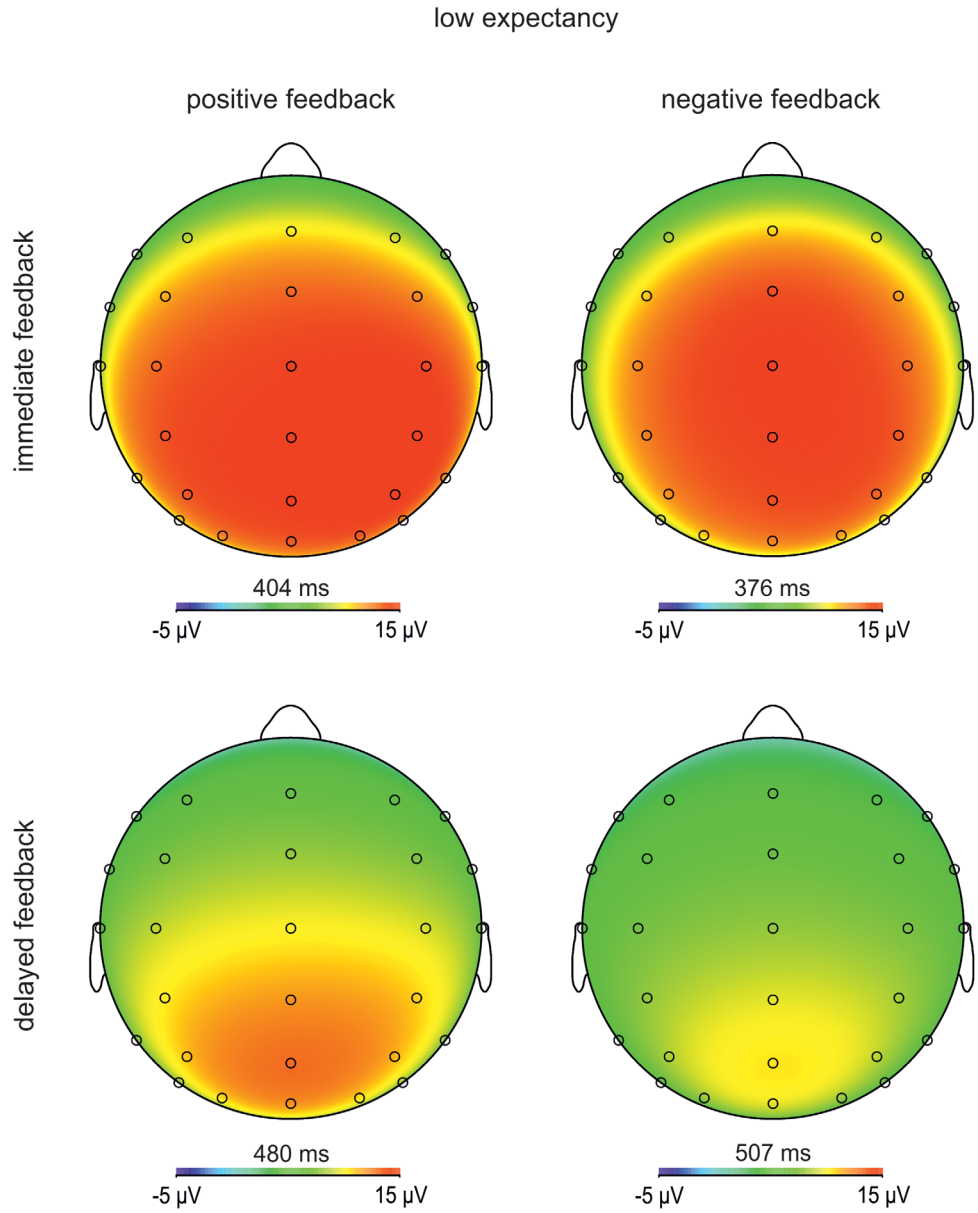

Figure 7. Topographic maps of the scalp distribution of the P300 peak for unexpected outcomes in the immediate and delayed feedback conditions.

conditions at electrode sites Pz and FCz. Statistical analysis revealed main effects of the three factors expectancy, electrode, and feedback timing: P300 amplitudes were generally higher for unexpected than for expected feedback, $F(1,34) = 5.576$; $p = .024$; $\eta_p^2 = .141$, at electrode site Pz compared to FCz, $F(1,34) = 27.826$; $p < .001$; $\eta_p^2 = .450$, and for immediate compared to delayed feedback, $F(1,34) = 5.421$; $p = .026$; $\eta_p^2 = .138$. The main effect of valence was not significant, $F(1,34) = 10.972$; $p = .376$; $\eta_p^2 = .023$. Significant interactions were found in Feedback Timing × Electrode, $F(1,34) = 12.055$; $p = .001$; $\eta_p^2 = .262$; Electrode × Valence, $F(1,34) = 6.155$; $p = .018$; $\eta_p^2 = .153$; and Expectancy × Valence, $F(1,34) = 21.309$; $p < .001$; $\eta_p^2 = .385$. No other significant two- or three-way interaction was found (all ps $> .066$). The resolution of the first interaction revealed that P300 amplitudes were larger for the immediate than the delayed feedback condition at FCz, $t(34) = -3.590$; $p = .001$, but not at Pz, $t(34) = -1.097$; $p = .280$. The interaction of Valence × Electrode resulted from a larger amplitude difference between electrode sites for positive, $t(35) = -4.603$; $p < .001$, than negative feedback, $t(35) = -3.972$; $p = .001$. Finally, concerning the interaction of Expectancy × Valence, we found that P300 amplitudes were significantly larger for unexpected positive compared to unexpected negative feedback, $t(35) = 2.682$; $p = .044$, whereas no significant difference emerged for expected positive and negative feedback, $t(35) = 1.244$; $p = .888$. Figure 7 shows the topographic distribution of the P300 following unexpected positive and negative feedback for both feedback timing conditions. These maps illustrate both the feedback timing main effect and the interaction of Feedback Timing × Electrode, with larger amplitude differences between groups at frontal than at parietal sites.

Discussion

In the present study, ERPs in healthy human participants performing a probabilistic learning task were analyzed. The aim was to investigate the influence of feedback expectancy on the processing of delayed compared to immediate feedback. Participants received feedback for choices between stimuli with different objective reward probabilities. Different groups of subjects received the feedback either 1 s (immediate condition) or 7 s (delayed condition) after the offset of the button press that indicated their choice. First, we could replicate previous findings that difference FRN amplitudes for delayed feedback are smaller compared to immediate feedback. Furthermore, and in accordance with our hypothesis, unexpected feedback elicited larger difference FRN amplitudes than expected feedback in both timing conditions, which was caused by a modulation of the positive feedback ERPs. Control analyses conducted on difficult trials only and in nonlearners support that the FRN modulation was related to expectancy (see supporting information). An expectancy effect could also be seen for P300 amplitudes. Moreover, P300 amplitudes were found to be smaller for delayed compared to immediate feedback, especially over the frontal cortex. Analyses of the behavioral data showed that the participants learned similarly well from immediate and delayed feedback. Here, it is important to note that the choice accuracy displayed in Figure 2 also comprises data of the nonlearners. When only the learners are considered, the average values are higher. Assessment of subjective reward probabilities also did not yield differences between the delay conditions. The subjective probabilities were higher than the objective values, especially for stimuli associated with low reward probabilities, which has been ascribed to an overoptimistic bias of reward expectation (see Miller & Ross, 1975; Oliveira et al., 2007; Radhakrishnan, Arrow, & Sniezek, 1996). Importantly, there was still a significant difference between high and low objective reward probability conditions, with reward expectation being higher for the former than the latter.

The present study's result of higher difference FRN amplitudes for immediate than delayed feedback corroborated the findings of a recent ERP study by our group (Peterburs et al., 2015). However, the overall pattern of results in the literature is mixed. Wang et al. (2014), for example, reported no differences between the difference FRN amplitudes following immediate and delayed feedback. In other studies (Opitz et al, 2011; Weinberg et al., 2012), effects of feedback delay did occur, but in the study by Weinberg et al. (2012) only for immediate feedback was an amplitude difference between negative and positive feedback found, whereas in the study by Peterburs et al. (2015), as in the present study, amplitudes differed between negative and positive feedback for different delay conditions. The type of task used might account for some differences in findings. Some previous studies (Wang et al., 2014; Weinberg et al., 2012) used gambling tasks, and the participants could not use the feedback to increase the likelihood of receiving positive outcomes by learning stimulus-outcome mappings. As prediction error representations in the FRN are particularly strong when stimulus-outcome associations can be learned (Holroyd et al., 2009), it is conceivable that the FRN is generally more pronounced in studies in which learning is possible. Opitz et al. (2011) observed an FRN decrease for negative feedback processing when comparing delays of 0 and 1 s, showing that short delays already affect feedback processing.

Applying a conceptually similar learning task as Peterburs et al. (2015), the present study thus adds to the evidence that delayed feedback elicits a difference FRN that is reduced in amplitude. As the ACC has also been linked to motivational processes, one interpretation could be that the reduction for delayed feedback reflects reduced motivation in the subjects due to the increased waiting time (Holroyd & Yeung, 2012; Kouneiher, Charron, & Koechin, 2009). Given the comparable learning performance for immediate and delayed feedback, this appears unlikely, however.

Instead, Foerde and Shohamy (2011) and Foerde et al. (2013) proposed a distinction of neuronal circuits being involved in immediate and delayed feedback processing. They found prediction error representations for immediate feedback in the striatum and for delayed feedback in the HC, which may bind relevant elements across time in learning from delayed feedback, in line with a growing body of literature (Cohen, Poldrack, & Eichenbaum, 1997; Shohamy & Wagner, 2008; Staresina & Davachi, 2009). Together with evidence from source-localization studies (Foti et al., 2011), fMRI and EEG separately (Carlson, Foti, Mujica-Parodi, Harmon-Jones, & Hajcak, 2011) or combined (Becker et al., 2014), which all point to a striatal contribution to the FRN, the reduced difference FRN might thus indicate a less strong role of the striatum and the DA system in delayed feedback processing (Peterburs et al., 2015; Weinberg et al., 2012). Importantly, however, the fact that a difference FRN can also be observed following delayed feedback appears to indicate that the striatal/DA system also contributes to delayed feedback processing.

This notion is supported by the main new finding of the present study: For both timing conditions, larger difference FRN amplitudes for unexpected compared to expected feedback were revealed. As outlined above, the difference FRN is thought to reflect prediction error signals conveyed from the midbrain to the striatal/ACC system with larger amplitudes for unexpected feedback compared to expected feedback (Gehring & Willoughby,

2002; Holroyd & Coles, 2002; Holroyd et al., 2009; Miltner et al., 1997; Oliveira et al., 2007; for review, see Walsh & Anderson, 2012). Therefore, an expectancy effect in the delayed feedback timing condition suggests a comparable involvement of nigrostriatal and/or mesocortical DA circuits, and consequently of the striatum and/or ACC, in the processing of immediate and delayed feedback. According to the original version of the RL theory (Holroyd & Coles, 2002), the expectancy effect is mainly driven by larger signals for unexpected negative feedback. Later studies found that the expectancy effect in the difference wave was mainly driven by a positivity for (unexpected) positive feedback (Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Weinberg, Riesel, & Proudfit, 2014). Holroyd (2004) and Holroyd et al. (2008) proposed that unexpected positive feedback causes phasic increases in DA activity, which in turn inhibits conflict-related dorsal ACC activity and thus drives the signal in the positive direction, increasing the difference FRN. Several other studies (Becker et al., 2014; Carlson et al., 2011; Foti et al., 2011; Weinberg et al., 2014) suggest that the difference FRN might reflect two independent overlapping processes that are linked to the striatum/ACC system and are influenced by positive and negative feedback, respectively.

In the present study, expectancy effects for both immediate and delayed feedback processing were driven by modulations of the positive feedback amplitude, adding to the overall impression that the mechanisms of feedback processing were comparable for the two delay conditions. Converging evidence for DA innervations influencing long-term potentiation in the HC (Otmakhova & Lisman, 1998; Shohamy & Adcock, 2010) suggests that prediction error signals are projected to both the striatum/ACC system and the HC via dopaminergic neurons. These results may indicate a cooperation of a more nondeclarative learning system associated with the striatum and a more declarative learning system associated with the HC (Sherry & Schacter, 1987; Squire, 1992; Squire & Zola, 1996) during feedback learning (Dickerson & Delgado, 2015; Dickerson et al., 2011; Knowlton et al., 1996). A related dissociation refers to a habitual system and a goal-directed system recruiting the basal ganglia and the prefrontal cortex and MTL, respectively (Corbit & Balleine, 2000; Cosman & Vecera, 2013; Daw & Shohamy, 2008; Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995). The decreased, but not absent, difference FRN amplitudes for delayed feedback reported above and in our previous study (Peterburs et al., 2015), as well as our finding of similar expectancy effects for both delay conditions, thus appear to indicate that feedback delay modulates the relative involvement of the striatal (habitual) and the medial temporal/prefrontal (goal-directed) systems in a more-or-less rather than in an all-or-nothing fashion (Peterburs et al., 2015).

Instead of a reduced involvement of the striatal system for the processing of delayed feedback, the current results could also indicate that prediction errors are computed when feedback is delayed, but that these signals are not used for learning. Instead, the goal-directed system based on the MTL could direct learning and behavior alone (see Walsh & Anderson, 2012). This interpretation seems to be supported by behavioral findings from patients suffering from Parkinson's disease (Foerde et al., 2013; Foerde & Shohamy, 2011), in whom learning from immediate but not delayed feedback is affected, suggesting that the striatum is not necessary for learning from delayed feedback. This, however, does not mean that learning from delayed feedback exclusively relies on the hippocampus and MTL in the healthy brain.

Finally, Opitz et al. (2011) linked reduced FRN amplitudes for delayed feedback to reduced prediction error processing due to stronger working memory demands. At first glance, this interpretation appears to contradict the present finding of comparable prediction error processing in immediate and delayed feedback. It has to be noted, however, that the tasks in the study by Opitz et al. (2011) and the present study differed. Opitz et al. (2011) had their participants learn an artificial grammar so that they had to keep alternative rules in mind across the delay. Working memory demands in the present study were likely smaller, as only two stimuli (out of five) and the choice had to be kept in mind. The high number of nonlearners was not specific for the delayed feedback condition and was probably more related to the probabilistic stimulus–outcome associations than to working memory demands. The different working memory demands could also explain the different result pattern concerning the P300 component. While Opitz et al. (2011) reported increased P300 amplitudes, our results indicated reduced P300 amplitudes for delayed feedback, strongly resembling the findings by Wang et al. (2014), who used comparable delays as the present study (600–1,000 ms vs. 4,000–5,000 ms). As in our study, they reported a significant difference between feedback delay conditions only at frontal electrodes. As one potential cause for the lower P300 amplitudes, they suggested that the feeling of relevance for oneself might decrease for longer time intervals between action and outcome.

Conclusion

The current data show that the FRN codes prediction errors in immediate as well as delayed feedback processing. This supports the idea that neuronal mechanisms eliciting the FRN component are comparable for immediate and delayed feedback. At the same time, overall reduced FRN amplitudes for delayed feedback processing suggest that the neural structures driving the FRN, the striatum, and the ACC are less strongly involved for delayed than immediate feedback processing, possibly due to reduced feedback relevance or salience. A reduction of the frontal P300 with increasing feedback delay supports this interpretation.

References

Barto, A. G. (1995). Adaptive critics and the basal ganglia. In J. C. Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 215–232). Cambridge, MA: MIT Press.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141. doi: 10.1016/j.neuron.2005.05.020

Becker, M. P., Nitsch, A. M., Miltner, W. H., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience*, *34*(8), 3005–3012. doi: 10.1523/JNEUROSCI.3684-13.2014

Bédard, P., Larochelle, L., Parent, A., & Poirier, L. J. (1969). The nigrostriatal pathway: A correlative study based on neuroanatomical and neurochemical criteria in the cat and the monkey. *Experimental Neurology*, *25*(3), 365–377. doi: 10.1016/0014-4886(69)90131-9

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*(7), 1823–1835. doi: 10.1111/j.1460-9568.2008.06138.x

Bellebaum, C., Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., & Sudmeyer, M. (2016). Strategies in probabilistic feedback learning in Parkinson patients OFF medication. *Neuroscience*, *320*, 8–18. doi: 10.1016/j.neuroscience.2016.01.060

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD

activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. doi: 10.1016/j.neuroimage.2011.05.037

Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory*, *5*(1–2), 131–178. doi: 10.1080/741941149

Corbit, L. H., & Balleine, B. W. (2000). The role of the hippocampus in instrumental conditioning. *Journal of Neuroscience*, *20*(11), 4233–4239.

Cosman, J. D., & Vecera, S. P. (2013). Learned control over distraction is disrupted in amnesia. *Psychological Science*, *24*(8), 1585–1590. doi: 10.1177/0956797613475632

Daw, N. D., & Shohamy, D. (2008). The cognitive neuroscience of motivation and learning. *Social Cognition*, *26*(5), 593–620. doi: 10.1521/soco.2008.26.5.593

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*, 70–88. doi: 10.1196/annals.1390.002

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: Effects of valence and magnitude manipulations. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(1), 27–38. doi: 10.3758/CABN.3.1.27

Dickerson, K. C., & Delgado, M. R. (2015). Contributions of the hippocampus to feedback learning. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 861–877. doi: 10.3758/s13415-015-0364-5

Dickerson, K. C., Li, J., & Delgado, M. R. (2011). Parallel contributions of distinct human memory systems during probabilistic learning. *NeuroImage*, *55*(1), 266–276. doi: 10.1016/j.neuroimage.2010.10.080

Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior*, *23*(2), 197–206. doi: 10.3758/BF03199935

Donkers, F. C., Nieuwenhuis, S., & van Boxtel, G. J. (2005). Mediofrontal negativities in the absence of responding. *Brain Research. Cognitive Brain Research*, *25*(3), 777–787. doi: 10.1016/j.cogbrainres.2005.09.007

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467. doi: 10.1111/j.1469-8986.1977.tb01312.x

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *Journal of Neuroscience*, *32*(35), 12087–12092. doi: 10.1523/JNEUROSCI.1410-12.2012

Foerde, K., Race, E., Verfaellie, M., & Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: Evidence from amnesia. *Journal of Neuroscience*, *33*(13), 5698–5704. doi: 10.1523/JNEUROSCI.5217-12.2013

Foerde, K., & Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience*, *31*(37), 13157–13167. doi: 10.1523/JNEUROSCI.2701-11.2011

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping*, *32*(12), 2207–2216. doi: 10.1002/hbm.21182

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. doi: 10.1126/science.1102941

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282. doi: 10.1126/science.1066893

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. doi: 10.1016/0013-4694(83)90135-9

Haber, S. N., & Fudge, J. L. (1997). The primate substantia nigra and VTA: Integrative circuitry and function. *Critical Reviews in Neurobiology*, *11*(4), 323–342 doi: 10.1615/CritRevNeurobiol.v11.i4.40

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*(2), 161–170. doi: 10.1111/j.1469-8986.2005.00278.x

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154. doi: 10.1016/j.biopsycho.2005.04.001

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*(6), 905–912. doi: 10.1111/j.1469-8986.2007.00567.x

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

Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. In M. Ullsperger, & M. Falkenstein (Eds.), *Errors, conflicts, and the brain: Current opinions on performance monitoring* (pp. 211–218). Leipzig, Germany: MPI of Cognitive Neuroscience.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. doi: 10.1037/0033-295X.109.4.679

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*(1), 93–101. doi: 10.1016/j.brainres.2005.12.015

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. doi: 10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective, & Behavior Neuroscience*, *9*(1), 59–70. doi: 10.3758/CABN.9.1.59

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*(5), 688–697. doi: 10.1111/j.1469-8986.2008.00668.x

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Science*, *16*(2), 122–128. doi: 10.1016/j.tics.2011.12.008

Kahneman, D., Knetsch, J. K., & Thaler, R. H., (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, *5*(1), 193–206. doi: 10.1257/jep.5.1.193

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399–1402. doi: 10.1126/science.273.5280.1399

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology*, *18*(4), 411–417. doi: 10.1097/01.wco.0000173463.24758.f6

Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, *12*(7), 939–945. doi: 10.1038/nn.2321

Larson, M. J., Kelly, K. G., Stigge-Kaufman, D. A., Schmalfuss, I. M., & Perlstein, W. M. (2007). Reward context sensitivity impairment following severe TBI: An event-related potential investigation. *Journal of International Neuropsychological Society*, *13*(4), 615–625. doi: 10.1017/S1355617707070762

Lavoie, B., Smith, Y., & Parent, A. (1989). Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase immunohistochemistry. *Journal of Comparative Neurology*, *289*(1), 36–52. doi: 10.1002/cne.902890104

Lehéricy, S., Ducros, M., Van de Moortele, P. F., Francois, C., Thivard, L., Poupon, C., … Kim, D. S. (2004). Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Annals of Neurology*, *55*(4), 522–529. doi: 10.1002/ana.20030

Lynd-Balta, E., & Haber, S. N. (1994). The organization of midbrain projections to the ventral striatum in the primate. *Neuroscience*, *59*(3), 609–623. doi: 10.1016/0306-4522(94)90181-3

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*(2), 339–346. doi: 10.1016/S0896-6273(03)00154-5

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*(2), 213–225. doi: 10.1037/h0076486

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. doi: 10.1162/jocn.1997.9.6.788

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*(5), 1936–1947.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins

and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*(4), 441–448. doi: 10.1016/j.neubiorev.2004.05.003

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454. doi: 10.1126/science.1094285

Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, *19*(12), 1994–2004. doi: 10.1162/jocn.2007.19.12.1994

Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: The impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience*, *5*, 8. doi: 10.3389/fnhum.2011.00008

Otmakhova, N. A., & Lisman, J. E. (1998). D1/D5 dopamine receptors inhibit depotentiation at CA1 synapses via cAMP-dependent mechanism. *Journal of Neuroscience*, *18*(4), 1270–1279.

Peterburs, J., Kobza, S., & Bellebaum, C. (2015). Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). *Psychophysiology*, *53*(2), 209–215. doi: 10.1111/psyp.12560

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. doi: 10.1038/35107080

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. doi: 10.1016/S0028-3932(02)00157-4

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Radhakrishnan, P., Arrow, H., & Sniezek, J. A. (1996). Hoping, performing, learning, and predicting: Changes in the accuracy of self-evaluations of peformance. *Human Performance*, *9*(1), 23–49. doi: 10.1207/s15327043hup0901_2

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Current Opinion in Neurobiology*, *7*(2), 191–197. doi: 10.1016/S0959-4388(97)80007-4

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*(3), 199–207. doi: 10.1038/35044563

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241–263. doi: 10.1016/S0896-6273(02)00967-4

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. doi: 10.1126/science.275.5306.1593

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review*, *94*(4), 439–454. doi: 10.1037/0033-295X.94.4.439

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Science*, *14*(10), 464–472. doi: 10.1016/j.tics.2010.08.002

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378–389. doi: 10.1016/j.neuron.2008.09.023

Simons, R. F. (2010). The way of our errors: Theme and variations. *Psychophysiology*, *47*(1), 1–14. doi: 10.1111/j.1469-8986.2009.00929.x

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243. doi: 10.1162/jocn.1992.4.3.232

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences U S A*, *93*(24), 13515–13522. doi: 10.1073/pnas.93.24.13515

Staresina, B. P., & Davachi, L. (2009). Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron*, *63*(2), 267–276. doi: 10.1016/j.neuron.2009.06.024

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, *103*(1), 1039–1061. doi: 10.2307/2937956

Walsh, M. M., & Anderson, J. R. (2012). Learning from delayed feedback: Neural responses in temporal credit assignment. *Cognitive, Affective, & Behavioral Neuroscience*, *11*(2), 131–143. doi: 10.3758/s13415-011-0027-0

Wang, J., Chen, J., Lei, Y., & Li, P. (2014). P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *NeuroReport*, *25*(13), 1044–1048. doi: 10.1097/WNR.0000000000000226

Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (2012). Better late than never? The effect of feedback delay on ERP indices of reward processing. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(4), 671–677. doi: 10.3758/s13415-012-0104-z

Weinberg, A., Riesel, A., Proudfit, G. H. (2014). Show me the money: The impact of actual rewards and losses on the feedback negativity. *Brain and Cognition*, *87*, 134–139. doi: 10.1016/j.bandc.2014.03.015

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004). Error-related negativity reflects detection of negative reward prediction error. *NeuroReport*, *15*(16), 2561–2565. doi: 10.1097/00001756-200411150-00027

(Received January 19, 2016; Accepted July 26, 2016)

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1: Analysis of difference FRN amplitude with balanced trial difficulty in learners.

Appendix S2: Analysis of difference FRN amplitudes in nonlearners.

Table S1: FRN difference wave amplitudes for ERPs derived from difficult trials only.

Table S2: FRN difference wave amplitudes for nonlearners.

Figure S1: Feedback-locked grand-averaged ERPs derived from difficult trials only.

Figure S2: Feedback-locked grand-averaged ERPs for nonlearners.

Effects of feedback delay and agency on feedback-locked beta and theta power during reinforcement learning

Benjamin Weismüller, Janna Kullmann, Matthias Hoenen, and Christian Bellebaum

Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf, Germany

Author Note

Benjamin Weismüller, Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf; Janna Kullmann, Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf; Matthias Hoenen, Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf; Christian Bellebaum, Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf

Correspondence concerning this article should be addressed to Benjamin Weismüller, Institute of Experimental Psychology, Department of Biological Psychology, Heinrich-Heine University Düsseldorf, Germany, Phone: +49 211 81-11596, E-Mail: Benjamin.Weismueller@hhu.de

Abstract

Feedback-based learning initiated by dopamine (DA) cell firing is crucial for adaptive behaviour. The nature and context of feedback can vary, however, affecting how feedback is processed. For example, the feedback-related negativity (FRN) in the event-related potential in humans, which has been linked to the DA system, is reduced for delayed feedback and for observational compared to active learning. Recent research suggested that oscillations in the theta and beta band over the medio-frontal cortex reflect distinct feedback processing mechanisms. In the present study we hypothesized that the power in both frequency bands is also affected by feedback delay and agency. We thus investigated effects of feedback delay (1 s vs. 7 s) on induced theta and beta band power and the FRN in a probabilistic feedback learning task in two groups of participants, one learning actively and one by observation. For theta and beta a larger power difference between negative and positive feedback for immediate than delayed feedback was found, driven by positive feedback for beta and by negative feedback for theta. Only for theta band power the difference between negative and positive feedback was stronger for active than observational learning. These results indicate that feedback-locked beta and theta power indeed reflect distinct neuro-cognitive mechanisms during feedback processing, with theta being linked to behavioural adaptation after negative feedback and beta to memory consolidation after positive feedback. With respect to the FRN amplitude, we could replicate previous findings of both delay and agency modulations, without an interaction of the two factors.

Keywords: feedback-learning, feedback delay, observational learning, Theta band, Beta band, FRN

1 Introduction

Across the animal kingdom, individuals are able to adapt their behaviour according to the consequences of earlier actions. Behaviour that previously resulted in positive feedback (e.g. a reward) is shown more frequently, while behaviour that previously led to negative feedback (e.g. a punishment or the omission of reward) is shown less frequently.

In primates feedback processing by the midbrain dopamine (DA) system is considered as the neuronal substrate for feedback learning. As shown in monkeys and humans, feedback that is better than expected (i.e. a positive prediction error) is associated with phasic bursts of activity, while feedback that is worse than expected (i.e. a negative prediction error) causes phasic dips in DA neuron firing rates (Schultz, 1997, 2000; Schultz, Dayan, & Montague, 1997; Zaghoul et al., 2009). These signals are projected to the striatum and frontal cortex including the anterior cingulate cortex (ACC; Bédard, Larochelle, Parent, & Poirier, 1969; Haber & Fudge, 1997; Lavoie, Smith, & Parent, 1989; Lehéricy et al., 2004; Lynd-Balta & Haber, 1994), where prediction error related activity has been located by means of functional neuroimaging in humans (Delgado, 2007; Delgado, Locke, Stenger, & Fiez, 2003; for a review, see Knutson & Cooper, 2005; O'Doherty et al., 2004).

Important insights into the temporal dynamics of reward processing have been gained by applying electroencephalography (EEG), in particular by analysing event-related potentials (ERPs). The feedback-related negativity (FRN), peaking at about 250 ms after feedback onset (Holroyd, 2004; Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997), is not only larger for negative compared to positive feedback (e.g. Ernst & Steinhauser, 2012, 2017; Gehring & Willoughby, 2002; Hajcak, Moser, Holroyd, & Simons, 2006; Hajcak, Moser, Yeung, & Simons, 2005; Holroyd, Hajcak, & Larsen, 2006; Holroyd & Krigolson, 2007; for a review, see Simons, 2010; Walsh & Anderson, 2012), but also for unexpected compared to expected feedback (e.g. Ferdinand, Mecklinger, Kray, & Gehring, 2012; Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Oliveira, McDonald, &

Goodman, 2007; Weismuller & Bellebaum, 2016). While some studies using dipole-modelling or low-resolution-brain-electromagnetic-tomography (LORETA) suggested that the ACC is the neural generator of the FRN (e.g., Bellebaum & Daum; Zhou, Yu, Zhou, 2010; for review, see Walsh & Anderson, 2012), a more recent dipole-modelling study provided evidence for a striatal contribution (Foti, Weinberg, Bernat, & Proudfit, 2015). Becker et al. (2014) reported that the blood-oxygen-level-dependent (BOLD) signal in the ventral striatum, the midcingulate and midfrontal cortex was significantly predicted by the ERP amplitude in the FRN time window. However, this result was obtained for ERPs following positive feedback, suggesting that the mentioned regions which are all involved in reward processing (Becker et al., 2014; Bush et al., 2002; Rogers et al., 2004; for a review, see Ullsperger, Danielmeier, & Jocham, 2014), primarily cause a shift of the ERP signal in the positive direction when rewards are processed. This finding is in line with one current view on ERP correlates of feedback processing, according to which the difference between ERP responses for positive and negative feedback mostly results from a positive deflection in response to unexpected positive feedback (Holroyd, Pakzad-Vaezi, & Krigolson, 2008), which has been labelled reward positivity (Holroyd, Krigolson, & Lee, 2011; Mulligan & Hajcak, 2017).

In recent years, analyses of oscillatory signals have frequently been used to study the neural mechanisms involved in feedback processing. Oscillations in certain frequency bands have been linked to processes of communication between distant brain regions and thus to the integration of different types of information related to attention, learning and memory (Cohen, Wilmes, & van de Vijver, 2011). In particular, signals in two frequency bands which are pronounced in the time window of the FRN have been discussed as being implicated in feedback processing. First, power in the medial frontal theta band (4-8 Hz) is increased for negative compared to positive feedback (Janssen, Poljac, & Bekkering, 2016). It has, at least by some researchers, been found to reflect a negative reward prediction error (Cavanagh & Frank, 2014; Cavanagh, Frank, Klein, & Allen, 2010; Cohen, 2011a; Li, Baker, Warren, & Li,

2016; but see Jannsen et al. (2016) for a negative finding on theta and outcome expectancy) and has been localised in the medial frontal/prefrontal lobe (van der Molen, Dekkers, Westenberg, van der Veen, & van der Molen, 2016; see also Cohen, 2014). Feedback-locked theta thus shares many features with the FRN, but has also been suggested to underlie other ERP components such as the N200 and the response-locked ERN. Accordingly, the signal in this frequency range is thought to reflect a general signal communicating the need for cognitive control and behavioural adaptation (Cavanagh & Frank, 2014; Cohen, 2014).

Oscillations in higher frequency bands in the beta and low gamma range (20-35 Hz), on the other hand, have been associated with reward-related activity in the ventral striatum (Marco-Pallares et al., 2008; Mas-Herrero, Ripolles, HajiHosseini, Rodriguez-Fornells, & Marco-Pallares, 2015), being stronger for positive compared to negative feedback (Cohen et al., 2011) and for larger compared to smaller reward magnitudes (Marco-Pallares et al., 2008). In a recent review, Marco-Pallares, Munte, and Rodriguez-Fornells (2015) proposed a model in which beta-gamma oscillatory activity reflects a motivational signal involved in reward processing and underlying memory formation by signalling which events are better than expected.

The neural underpinnings of feedback processing and learning appear to vary, however, depending on the context and the experimental settings. For example, the timing of feedback relative to the preceding response affects, which neural structures mediate feedback learning. When feedback is presented 7 s after stimulus choice, striatal activity is reduced and hippocampal activity enhanced compared to immediate feedback (1 s after stimulus choice; Foerde & Shohamy, 2011). Furthermore, patients with striatal dysfunction suffering from Parkinson's Disease (PD) can learn from delayed, but not immediate feedback, while amnestic patients with suspected hippocampal damage show the opposite pattern (Foerde, Race, Verfaellie, & Shohamy, 2013).

In addition, (perceived) agency or lack of agency seems to alter feedback processing. For example, functional imaging studies showed that the regions involved in active and vicarious reward processing overlap to some extent, but that typical reward processing regions in the striatum are more strongly involved when rewards refer to a person him-herself compared to an observed person (Morelli, Sacchet, & Zaki, 2015). The dorsal striatum, which has previously been implicated in the coding of action-outcome associations (O'Doherty et al., 2004), seems to play an important role in linking own actions to outcomes (Bellebaum, Jokisch, Gizewski, Forsting, & Daum, 2012; Kobza & Bellebaum, 2015).

Concerning the electrophysiological correlates of feedback processing, effects on feedback-locked ERPs were found for both delay and agency manipulations, with higher FRN amplitudes for immediate compared to delayed feedback (Arbel, Hong, Baker, & Holroyd, 2017; Opitz, Ferdinand, & Mecklinger, 2011; Peterburs, Kobza, & Bellebaum, 2016; Weinberg, Luhmann, Bress, & Hajcak, 2012; Weismuller & Bellebaum, 2016; but see Wang, Chen, Lei, & Li, 2014 for a negative finding) and for feedback given to active performers versus for observed persons (Bellebaum & Colosio, 2014; Bellebaum, Kobza, Thiele, & Daum, 2010; Koban, Pourtois, Bediou, & Vuilleumier, 2012). These findings suggest a reduced involvement of the striatum/ACC in feedback processing when feedback does not follow an own action immediately.

As outlined above, signals in the beta-gamma and theta range have been suggested to reflect cognitive processes that can be segregated from the processes underlying the FRN. For both frequency bands, effects of agency and/or feedback delay are conceivable and would yield new insights into their functional role in the context of feedback processing. In the present study we thus recorded electroencephalographic activity while participants engaged in active or observational feedback learning tasks with either short or long response-feedback delays. For feedback-locked theta we hypothesised that observation relative to active responding would lead to reduced power, especially for negative feedback, due to a reduced

need for cognitive control, while for feedback delay no clear predictions could be made. For the beta-gamma band, we expected a power modulation for positive feedback, which was expected to yield high power values particularly when it followed recent behaviour, as beta-gamma power has been associated with the reinforcement of preceding rewarded actions (Feingold, 2011; Marco-Pallares et al., 2015; Mas-Herrero et al., 2015). Moreover, the proposed link to striato-frontal information processing (Marco-Pallares et al., 2015) suggests reduced beta-gamma power for feedback given to observed persons and/or after a delay, as striatal involvement is reduced in both situations (see above). Finally, we also analysed the FRN for which a combined influence of delay and agency modulations has not been investigated to date.

2 Method

2.1 Study Participants

Forty healthy students of the Heinrich-Heine-University Düsseldorf were recruited as participants for the present study via advertisement, and each participant was assigned randomly to one of two groups. One group took part in active feedback-learning tasks while the other group engaged in observational versions of the task. The mean age of the participants of the active group (n = 20 participants) was 24.8 years (SD = 2.7, 9 male, 11 female), while the 20 participants of the observational group were on average 24.7 years old (SD = 3.1, 4 male, 16 female). All Participants had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders, traumatic brain injury with sustained unconsciousness or regular consumption of alcohol or psychodynamic drugs. The participants gave informed written consent before testing and were reimbursed with 15 €. The study conformed to the guidelines outlined in the Declaration of Helsinki and was approved by the Ethics Committee of the Faculty of Mathematics and Natural Sciences at Heinrich-Heine University Düsseldorf.

2.2 Probabilistic Learning Task

2.2.1 Active task. For this experiment, we used a modified version of a paradigm that we applied recently in another study (Weismuller & Bellebaum, 2016). In each trial two stimuli (Asian symbols) were presented to the participants, one on the left and one on the right side of a computer screen. The participants then were asked to choose one of the two stimuli. By means of monetary feedback (i.e. monetary gain of +20¢ or loss of -10¢") for their choice participants could learn to predict the outcomes associated with each stimulus and maximize reward. In total, five different stimuli were used, each linked to a particular reward probability (0%, 20%, 40%, 60%, and 80%) that was unknown to the participants. Figure 1A shows a set of stimuli and their corresponding reward probabilities. All 10 possible combinations of stimuli were presented equally often, with the assignment of stimulus to the side of the computer screen being counterbalanced. As soon as the participants responded, the chosen stimulus was encircled in red for 500 ms to indicate the choice. Afterwards, a fixation cross was shown for 500 ms, followed by the presentation of the feedback stimulus for another 500 ms. In case the participant did not reply within 3000 ms he/she was asked to respond faster. After feedback presentation a fixation cross appeared again until the next trial followed with an intertrial interval between 1200 and 1600 ms. Importantly, each participant completed two versions of the task, which differed in the delay between response and feedback. Immediate feedback followed 1000 ms after the choice, whereas delayed feedback followed 7000 ms after the choice in line with several previous studies using similar temporal delays (Arbel et al., 2017; Foerde et al., 2013; Foerde & Shohamy, 2011; Peterburs et al., 2016; Weismüller & Bellebaum, 2016). A different set of stimuli was used for each feedback timing. Figure 1B illustrates the sequence of events during the active learning task.

2.2.2 Observational task. Those participants who completed the observational learning tasks were not asked to choose one of two stimuli themselves, but rather watched

another participants' performance. As in the active learning task, two stimuli were shown on each trial. Then a picture of a left or right hand appeared on one side of the screen indicating the observed choice of one stimulus. This stimulus was then encircled in red and the observer was asked to confirm the choice within 3000 ms by pressing the corresponding (left or right) button in order to ensure that the observers were attending to the observed person's choices. If the participants did not respond in this time window, a reminder was presented to react faster. After the choice was confirmed, a fixation cross was shown (duration 500 ms), followed by the feedback (500 ms) before the next trial came on. Figure 1C illustrates the exact timing of events during one trial of the observational learning task. As the participants of the active learning task, the observers completed two versions, which differed in the delay between (observed) button press and feedback presentation (1000 ms vs. 7000 ms). Importantly, the performance that was observed by the participants of the observational learning task was taken from one participant who had completed the active learning task. Thereby, each participant in the observation group observed the performance of one participant from the active group for both feedback timings.

2.2.3 Test phase. As in the observational task versions participants only confirmed others' choices, learning could not be assessed in the learning trials. Thus, an additional test phase was included after each block of learning trials, which required active responding but did not entail feedback. This was also done in the active learning task so that performance could be directly compared between active and observational learning. Similar procedures for observational learning paradigms have been reported in previous studies of our group (e.g. Bellebaum & Colosio, 2014; Bellebaum et al., 2012). Instead of trial-by-trial feedback, the overall amount of money lost/gained was reported after the whole block in test phases. The symbols and symbol combinations presented in the test phase were identical to those in the learning phase so that participants could use the knowledge gained in the learning trials to complete the test phase. Figure 1D shows details of the timing of the test trials.

---------- Insert Figure 1 about here ----------

2.3 Procedure

The participants' demographic data, medication, and medical history were assessed in a structured interview before EEG recording started. Then, participants were prepared for the EEG session and comfortably seated about 70 cm away from a computer screen of 27" diagonal dimension. They were asked to relax and not to move. Breaks in which the participants could rest were included after every block. After the task was explained, participants completed six practice trials with feedback and six practice trials without feedback for each of the two feedback delay versions. When the experimenter was sure that the task was understood, the actual experiment and the EEG recording started.

All participants completed three (active or observational) learning blocks with immediate feedback and three blocks with delayed feedback, each consisting of 100 trials (10 trials per symbol combination). Whether participants started with the immediate feedback or the delayed feedback version of the learning task was counterbalanced across participants. A test phase consisting of 60 trials (6 per symbol combination) followed after every learning block, regardless of feedback timing or agency (i.e. active or observational). Learning blocks took 8-10 min for immediate feedback and 18-20 min for delayed feedback. Test phases took approx. 3 min. The total duration of the EEG session was approx. 90 min. After the experiment, participants were reimbursed with 15 €.

2.4 Electroencephalographic Recording

EEG data were continuously sampled at 1000 Hz via BrainVision Recorder software, version 1.20 (Brainproducts, Munich, Germany). Active silver/silver-chloride electrodes were attached in accordance with the international 10-10-system (Chatrian, Lettich & Nelson,

1988) at 29 electrode sites (F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, and PO10; FCz served as online reference) on a 32-channel-ActiCAP-electrode cap (ActiCAP; Brain Products GmbH, Germany). Horizontal eye movements were recorded with an electrode placed lateral to the left outer canthus. Blinks and vertical eye-movements were recorded by an electrode above the left eye (Fp1). Impedances were kept below 10 kΩ.

2.5 Data Analysis

2.5.1 Behavioural data. In accordance with earlier studies using probabilistic learning tasks (e.g. Bellebaum et al., 2016; Foerde et al., 2013; Frank, Seeberger, & O'Reilly R, 2004; Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Knowlton, Mangels, & Squire, 1996; Weismuller & Bellebaum, 2016), a participant's choice was considered correct if the symbol with the higher objective reward probability was chosen (regardless of the actual feedback the participant received). The percentage of correct choices per block, pooled across all stimulus combinations, was then the dependent variable for behavioural data analysis.

2.5.2 EEG-data. BrainVision Analyzer software, version 2.1 (Brainproducts, Munich, Germany) and MATLAB R2013a (MathWorks, Natrick, Massachusetts) were used to analyse EEG data offline.

2.5.2.1 Time-Frequency data. For the time-frequency analysis the data were first down-sampled to 500 Hz and re-referenced to linked mastoids and corrected for direct current trends. Then, the data were 40 Hz low-pass and 0.5 Hz high-pass filtered. Blink artifacts were identified and corrected in single participant EEG data by independent component analysis (ICA). ICA decomposes the spatially summed multivariate EEG signal into independent virtual components in order to reconstruct possible independent sources of the signal. For each participant 29 components were created from the 29 active electrode channels in the EEG signal. Components were considered to represent blink artifacts, if they explained a large

part of the variance and had a symmetrical, frontally positive topography. The identified component was then excluded before back transformation. Then, the data were segmented into intervals from 700 ms before the feedback stimulus until 1500 ms post-stimulus, separately for the four different feedback conditions (positive and negative immediate and delayed feedback). For each condition, a time-frequency analysis was performed by continuous complex Morlet wavelet transformation using wavelets that were optimised for the two frequency bands that were examined. The Morlet wavelets' spectral bandwidth was $2\sigma_f$ = 1.61 Hz at the central frequency of 5.65 Hz for theta frequency and $2\sigma_f$ = 7.50 Hz at the central frequency of 26.22 Hz for beta band. For both frequency bands the Morlet parameters were set to $2\pi\sigma_t = 7f^{-1}$ and baseline corrections were performed using a reference interval from 500 ms to 200 ms preceding the feedback stimulus. The spectral power was extracted and averaged across trials. Then, a continuous complex Morlet wavelet transformation was also performed for the average positive and negative feedback ERPs for both feedback timings in active and observational learners, using the same parameters as before. The ERP spectral power was then subtracted from the spectral power averaged across trials for each condition separately, so that the induced power remained. For the statistical analysis of theta frequency, mean induced power values were extracted for the time window between 200 and 500 ms after feedback onset for frequencies in the range between 4.12 and 7.75 Hz at FCz, separately for each individual participant and for each condition. Beta-gamma frequency power values were extracted between 19.63 and 35 Hz in the time window between 200 and 500 ms after feedback onset for each participant and condition, again at the FCz electrode. This frequency range will be referred to as beta band in the following. The time windows were chosen in accordance with values described in the literature and because the power modulation in the respective windows was most pronounced.

2.5.2.2 Event-related potentials. The first steps of analysis, (re-referencing, direct current trend correction, filtering, ICA, and artifact correction) were identical for the analysis

of ERPs and of the frequency bands (see above). Following those steps, the data were segmented, again separately for the feedback valence (positive or negative) and timing conditions (immediate vs. delayed), creating epochs from 200 ms preceding the feedback stimulus until 800 ms after the feedback stimulus. After segmentation, data were baseline corrected using the 200 ms interval before feedback onset, followed by artifact rejection. Segments were excluded if they contained data points exceeding an absolute amplitude value of 100 μV or if the amplitude difference between the highest and lowest data point exceeded 100 μV. In a final step, data from each of the different feedback conditions were averaged across trials.

In agreement with earlier studies on effects of feedback timing, the focus was on the negative – positive feedback difference wave (Peterburs et al., 2016; Wang et al., 2014; Weinberg et al., 2012; Weismuller & Bellebaum, 2016). The maximum negative peak in this difference wave was identified between 200 and 370 ms after feedback onset at electrode site FCz and considered as FRN amplitude.

2.6 Statistical Analyses

Data were statistically analysed using SPSS Statistics 23 software (IBM, Armonk, New York). Choice accuracy, extracted from test phases without feedback, was analysed by means of an analysis of variance (ANOVA) with Block (1st to 3rd) and Feedback Timing (immediate vs. delayed) as within-subjects factors and Agency (active vs. observational) as between-subjects factor. Time-frequency data were analysed by means of an ANOVA with the within-subjects factors Feedback Timing and Valence (positive vs. negative) and the between-subjects factor Agency, separately for theta and beta power. Finally, FRN amplitudes were analysed using an ANOVA with Feedback Timing as within-subjects factor and Agency as between-subjects factor. For all analyses, results were considered as statistically significant when an *alpha* level of $p < 0.05$ was reached. Post-hoc t-tests were performed to resolve

interactions. Whenever sphericity was violated, *Greenhouse-Geisser* correction was applied. As measures of effect size η_p^2 is reported for ANOVAS and Cohen's *d* for t-tests.

3 Results

In the active group one participant had 72 misses (more than 10 % of all trials, the maximum value for the other participants was 21). This participant as well as the corresponding observer were excluded from data analysis. The resulting sample for the analysis of behavioural and EEG data thus consisted of 19 participants in the active group (age: M = 24.9 years, SD = 2.64, 9 males, 10 females) and 19 in the observer group (age: M = 24.6 years, SD = 3.13, 4 males, 15 females).

3.1 Behavioural Data

Figure 2 shows the accuracy of the participants of both experimental groups in the different timings across the three blocks. The mixed ANOVA revealed a significant main effect of Block ($F(2, 72) = 4.254$; $p = .018$; $\eta_p^2 = .106$) with a linear increase in the percentage of correct choices ($F(1, 36) = 7.777$; $p = .008$; $\eta_p^2 = .178$). Learning from observed and active choices did not differ significantly, as no main effect of Agency was found ($F(1, 36) = 0.017$; $p = .897$; $\eta_p^2 < .001$). Neither did Feedback Timing influence performance ($F(1, 36) = 0.416$; $p = .523$; $\eta_p^2 = .011$), nor was any significant interaction found (all $p > .200$).

---------- Insert Figure 2 about here ----------

3.2 EEG-data

3.2.1 Time-frequency data. The mean induced power for negative and positive feedback and the negative-positive difference at electrode site FCz for each Feedback Timing condition, in active and observational learners, is depicted in figure 3.

3.2.1.1 Theta band. Figure 4 shows the mean induced theta band power values for all conditions. Statistical analysis revealed a significant main effect of Valence ($F(1, 36) = 17.201$; $p < .001$; $\eta_p^2 = .323$) indicating generally higher theta band power for negative compared to positive feedback. Significant interactions were found between the factors Valence and Agency ($F(1, 36) = 8.669$; $p = .006$; $\eta_p^2 = .194$), and between Valence and Feedback Timing ($F(1, 36) = 15.997$; $p < .001$; $\eta_p^2 = .308$). Post-hoc t-tests revealed that negative feedback caused higher theta frequency power than positive feedback when participants learned actively ($t(18) = 4.560$; $p < .001$; $d = 1.075$), while theta power was comparable between feedback valences during observational learning ($t(18) = 0.957$; $p = .351$; $d = 0.225$). A resolution of the second interaction showed that theta frequency power was higher for negative compared to positive immediate feedback ($t(37) = 4.522$; $p < .001$; $p = 0.743$), while no difference was found for delayed feedback ($t(37) = 0.952$; $p = .347$; $d = 0.157$). All remaining main effects and interactions including the three-way interaction involving all factors did not reach significance (all $p > .05$).

3.2.1.2 Beta band. One active learner was excluded from the analysis of beta power, as his power values differed by more than 5 standard deviations from the group mean in two of the four conditions. Figure 5 depicts the mean induced beta band power values for the remaining participants. Statistically, a trend for a main effect was revealed for Valence ($F(1, 35) = 3.038$; $p = .090$; $\eta_p^2 = .080$) with higher (i.e. less negative) power values for positive than negative feedback. Furthermore, Valence significantly interacted with Feedback Timing ($F(1, 35) = 5.022$; $p = .031$; $\eta_p^2 = .125$). This interaction arose from significantly higher power values following positive compared to negative immediate feedback ($t(36) = -2.442$; $p = .020$; $d = 0.407$), while the power values were comparable for positive and negative delayed feedback ($t(36) = 0.461$; $p = .648$; $d = 0.077$). No other main effect or interaction reached or approached significance (all $p > .10$).

---------- Insert Figure 3 about here ----------

---------- Insert Figure 4 about here ----------

---------- Insert Figure 5 about here ----------

3.2.2 ERP data. Feedback-locked grand average ERPs are illustrated in figure 6 for both learning conditions and feedback timings, table 1 lists mean FRN amplitudes. Statistical analysis of the FRN difference wave indicated a significant main effect of Feedback Timing ($F(1, 36) = 6.712$; $p = .014$; $\eta_p^2 = .157$), with larger amplitudes for immediate compared to delayed feedback. Furthermore, amplitudes differed significantly between Agencies ($F(1, 36) = 5.817$; $p = .021$; $\eta_p^2 = .139$), being larger for active compared to observational learning. Both main effects were mainly driven by the results for immediate feedback during active learning which is expressed in a significant interaction between Feedback Timing and Agency ($F(1, 36) = 5.126$; $p = .030$; $\eta_p^2 = .125$). Further investigating this interaction with post-hoc t-tests revealed larger FRN difference wave amplitudes for immediate compared to delayed feedback only for active ($t(18) = -2.761$; $p = .013$; $d = 0.651$), but not for observational learning ($t(18) = -0.343$; $p = .736$; $d = 0.081$).

---------- Insert Table 1 about here ----------

---------- Insert Figure 6 about here ----------

4 Discussion

The present study investigated the influence of feedback delay and agency (i.e. ascribing feedback to a self-generated vs. an observed action) on the neural correlates of feedback processing. EEG data were assessed during a probabilistic feedback learning task in which one group of participants received feedback for actively choosing between stimuli, while the other group learned by observing the choices of another person. For participants of

both groups feedback was provided immediately (1 s after stimulus choice) or after a temporal delay (7 s after stimulus choice) in different task versions.

Behavioural data showed that neither feedback timing, nor agency (active vs. observational learning) affected the participants' performance. As hypothesised, theta power was generally increased for negative compared to positive feedback. Moreover, interactions between Valence and Feedback Timing, as well as Valence and Agency revealed that both feedback delay and agency, modulated this valence effect on theta power independently. The difference between negative and positive feedback was larger for immediate than delayed feedback and for active than observational learning. For beta power, as for the theta band, an interaction was found for Valence and Feedback Timing, with higher power values for positive than negative feedback during immediate, but not delayed feedback learning. Effects of agency on beta power were not found. Finally, concerning the ERP result pattern, Feedback Timing and Agency showed a significant interaction: the largest difference wave amplitude (negative minus positive feedback processing) was found for immediate feedback during active learning.

4.1 Theta power and behavioural adaptation

Theta oscillations over the medial frontal cortex have been suggested to underlie several ERP components, including the FRN and the feedback locked N2, but also the ERN and CRN as response-locked signals (Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012). In accordance with the neural processing indicated by the FRN (Holroyd & Coles, 2002; Holroyd et al., 2004), some previous studies linked increased feedback-locked theta power to negative feedback processing (Janssen et al., 2016) or generally to "bad" events (Cavanagh & Shackman, 2015). Given that the N2 is sensitive to conflict and novelty (Folstein & Van Petten, 2008) and that the ERN and CRN indicate error and/or conflict processing (Cavanagh & Frank, 2014; Cavanagh et al., 2012), a more general

account thus proposed that fronto-medial theta oscillations reflect neuronal top-down computations in the sense of a general alarm signal, possibly also indicating the need for cognitive control (Cavanagh et al., 2010; Cohen, 2014), by synchronising neuronal oscillatory activity across various cortical areas (Cavanagh & Frank, 2014; Cohen, 2011b). Summarising a variety of different findings, Cavanagh and Shackman (2015) postulated that the increased need for cognitive control signalled by medial frontal theta facilitates behavioural adaption and results from a general uncertainty about action outcomes, including prediction errors and conflict.

In the present study theta power was specifically enhanced for negative compared to positive feedback when it was given immediately for own choices. An interpretation of theta in terms of an increased need for cognitive control appears to be plausible for the effect of agency, as in observational learning the own behavioural strategy does not have to be updated following negative feedback. For delayed negative feedback for active choices, however, the response strategy needs to be changed, so that diminished theta power in this condition appears to be surprising at first sight. As we and others already argued, altered electrophysiological responses in learning from delayed feedback in the absence of behavioral changes probably indicate that different neural mechanisms underlie learning in this condition (Arbel et al., 2017; Peterburs et al., 2016; Weismuller & Bellebaum, 2016). Theta activity associated with conflict and reward processing has been found in the dorsal (Cohen, Ridderinkhof, Haupt, Elger, & Fell, 2008; Wang, Ulbert, Schomer, Marinkovic, & Halgren, 2005) and rostral ACC in humans. But also in other regions signals in the theta band have been described. In rats, medial prefrontal neurons have been found to be phase locked to hippocampal theta oscillations in spatial working memory tasks (Siapas, Lubenov, & Wilson, 2005). In humans, errors have been found to cause enhanced theta power synchrony between FCz and lateral frontal (Cavanagh, Cohen, & Allen, 2009) and occipital electrode sites (Cohen et al., 2009). It is thus conceivable that the processing of conflict induced by delayed

feedback is also reflected in theta signals, but more in the medial temporal lobe and possibly the hippocampus, given that these structures underlie delayed feedback processing (see Foerde et al., 2013; Foerde & Shohamy, 2011). In the data of the present study modulations of theta band power at temporal or occipital electrode sites were not seen, but it may nevertheless be possible that signals in the theta band which could not be measured by surface electrodes contributed to feedback processing for delayed feedback.

4.2 Beta frequency and memory consolidation

Beta power has been associated with reward-related activity in the ventral tegmental area (VTA) and its projection sites in the ventral striatum (Mas-Herrero, et al., 2015) and the prefrontal cortex (HajiHosseini & Holroyd, 2015). Similar to theta band oscillations, synchronisation in the beta band may reflect communication across different systems and distant cortical and subcortical structures (Marco-Pallares et al., 2015). In contrast to theta, however, beta band oscillations have been linked to more specific processes in the context of feedback processing, that is, the integration of recent behaviours and their (positive) consequences, with the aim of memory consolidation for the established link (Feingold, 2011; Marco-Pallares et al., 2015; Mas-Herrero et al., 2015).

To our knowledge, the present study is the first to investigate beta oscillations for feedback that is temporally delayed and/or referring to another person's action. The results for immediate feedback are in line with previous findings (Cohen et al., 2011; Marco-Pallares et al. 2008; Mas-Herrero, et al., 2015; HajiHosseini & Holroyd, 2015) showing that beta power was increased for positive compared to negative feedback. However, this was not the case when feedback was given with a short delay, possibly indicating less fronto-striatal communication in response to feedback that is delayed by a few seconds, which is again in line with the described findings on reduced striatal involvement in delayed feedback processing (Foerde & Shohamy, 2011; Foerde et al., 2013). In contrast to theta, differences in

beta power for positive versus negative feedback were not affected by agency. This finding might suggest that the described integration process takes place to a similar extent in observational and active learning. Indeed, the descriptive pattern for immediate feedback in observational learning resembles the pattern for immediate feedback in active learning. Previous studies by our group and others (Morelli et al., 2015; Kobza & Bellebaum, 2015; Bellebaum et al., 2012) have shown at least partially reduced striatal activity for feedback processing in observational learning. The process indicated by beta, however, might reflect an integration process that is similar for active and observational learning. Despite the differences, the neural structures involved in active and observational learning also overlap to a large extent and it has been suggested that the dorsal striatum, similar to its role in active instrumental learning, underlies the integration of actions and outcomes also during observation (Cooper, Dunne, Furey, & O'Doherty, 2012).

Considering the proposal that beta reflects memory for rewarded as opposed to not-rewarded actions, the difference in power for positive and negative feedback might either arise from memory enhancement for rewarded behaviours (Mas-Herrero et al., 2015; Murty & Adcock, 2013) or from memory suppression for not-rewarded actions (Feingold, 2011). Beta power values in this study appear to primarily reflect the latter mechanism, as beta power values were negative (relative to baseline) for both positive and negative feedback, with the smallest decrease for immediate positive feedback for self-performed actions. In this context it is important to note that performance accuracy was quite high in the present study, so that positive feedback, on average, elicited less strong prediction errors. Reward expectation has been proposed to affect beta power negatively, so that it should be relatively strengthened for unexpected positive feedback (Cohen, Elger, & Raganath, 2007; Marco-Pallares et al., 2015). Thus, the process of memory consolidation was probably more suppressed after (unexpected) negative feedback than strengthened after (expected) positive feedback.

Together, the present results corroborate the idea that theta and beta band activation supposably reflect valence specific top-down learning signals generated in different medial frontal neuronal systems (Cohen et al., 2011). Theta frequency conveys the need of behavioural adaption (Cavanagh et al., 2010; Cohen, 2014), which is only necessary for own actions, while signals in the beta range possibly convey a motivational value signal for memory storage of relevant information including consequences of others' actions (Mas-Herrero et al., 2015).

4.3 Combined agency and delay effects on the FRN

Concerning the ERPs, the present study is generally in line with previous findings of both feedback delay (Arbel et al., 2017; Opitz et al., 2011; Peterburs et al., 2016; Weinberg et al., 2012; Weismuller & Bellebaum, 2016) and agency effects (Bellebaum & Colosio, 2014; Bellebaum et al., 2010; Koban et al., 2012; Yeung, Botvinick, & Cohen, 2004) on FRN amplitude. However, it has to be noted that the FRN in the present study was mostly affected by the combined influence of both manipulations. As ERPs in the FRN time window have been shown to be driven by the ACC and striatum (Becker et al., 2014; Foti et al., 2015) and a theoretical link to the DA system has been proposed (Holroyd & Coles, 2002; Walsh & Anderson, 2012), diminished amplitudes might indicate a reduced involvement of these feedback processing structures. Indeed, reduced striatal involvement has been suggested for both delayed (Foerde et al., 2013; Foerde & Shohamy, 2011) and non-personal feedback or reward processing (Morelli et al., 2015; Bellebaum et al., 2012; Kobza & Bellebaum, 2015; Bellebaum et al., 2016; Kobza et al., 2012). The present data show that the FRN is not additionally affected by feedback delay during observational feedback learning. Accordingly, this study is the first to show that the FRN, and thus the underlying neural mechanisms are particularly involved in feedback processing if feedback is given a) for own actions and b) shortly after the action.

4.4 Conclusion

In the context of feedback processing, signals in the beta and theta frequency bands appear to represent activity in two distinct, valence-specific neuronal systems. Comparing actual outcomes to the predictions, negative prediction errors may elicit a signal for behavioural adaptation reflected by theta oscillations (Cohen, 2014, Cavanagh et al., 2010) and positive prediction errors activate processes for memory consolidation reflected by beta oscillations (Feingold, 2011; Marco-Pallares et al., 2015; Mas-Herrero et al., 2015). The presented data provide corroborating evidence for this distinction and suggest that the processes indicated by theta are weakened when feedback is delayed or when the feedback does not refer to a self-generated action, while the mechanism indicated by beta is only affected by feedback delay and not by agency. Feedback-locked ERPs were modulated by a combination of feedback delay and agency showing that the FRN reflects yet another mechanism in feedback processing.

Acknowledgements

The project was conducted at and financed by the Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf. The authors declare that they have no conflict of interest.

References

Arbel, Y., Hong, L., Baker, T. E., & Holroyd, C. B. (2017). It's all about timing: An electrophysiological examination of feedback-based learning with immediate and delayed feedback. *Neuropsychologia, 99*, 179-186. https://www.doi.org/10.1016/j.neuropsychologia.2017.03.003

Becker, M. P., Nitsch, A. M., Miltner, W. H., & Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *Journal of Neuroscience, 34*(8), 3005-3012. doi: https://www.doi.org/10.1523/JNEUROSCI.3684-13.2014

Bédard, P., Larochelle, L., Parent, A., & Poirier, L. J. (1969). The nigrostriatal pathway: a correlative study based on neuroanatomical and neurochemical criteria in the cat and the monkey. *Exp Neurol, 25*(3), 365-377. https://www.doi.org/10.1016/0014-4886(69)90131-9

Bellebaum, C., & Colosio, M. (2014). From feedback- to response-based performance monitoring in active and observational learning. *J Cogn Neurosci, 26*(9), 2111-2127. https://www.doi.org/10.1162/jocn_a_00612

Bellebaum, C., Jokisch, D., Gizewski, E. R., Forsting, M., & Daum, I. (2012). The neural coding of expected and unexpected monetary performance outcomes: dissociations between active and observational learning. *Behav Brain Res, 227*(1), 241-251. https://www.doi.org/10.1016/j.bbr.2011.10.042

Bellebaum, C., Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., & Sudmeyer, M. (2016). Strategies in probabilistic feedback learning in Parkinson patients OFF medication. *Neuroscience, 320*, 8-18. https://www.doi.org/10.1016/j.neuroscience.2016.01.060

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cereb Cortex, 20*(12), 2874-2883. https://www.doi.org/10.1093/cercor/bhq038

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: A role in reward-based decision making. Proceedings of the National Academy of Sciences, 99(1), 523-528. https://www.doi.org/10.1073/pnas.012470999

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience, 29*(1), 98-105. https://www.doi.org/10.1523/JNEUROSCI.4137-08.2009

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci, 18*(8), 414-421. https://www.doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage, 49*(4), 3198-3209. https://www.doi.org/10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *J Physiol Paris, 109*(1-3), 3-15. https://www.doi.org/10.1016/j.jphysparis.2014.04.003

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49*(2), 220-238. https://www.doi.org/10.1111/j.1469-8986.2011.01293.x

Chatrian, G. E., Lettich E., & Nelson, P. L. (1988). Modified nomenclature for the "10%" electrode system. *Journal of Clinical Neurophysiology, 5*(2), 183-186. https://www.doi.org/10.1097/00004691-198804000-00005

Cohen, M. X. (2011a). It's about Time. *Frontiers in Human Neuroscience, 5*, 2. https://www.doi.org/10.3389/fnhum.2011.00002

Cohen, M. X. (2011b). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage, 55*(3), 1373-1383. https://www.doi.org/10.1016/j.neuroimage.2010.12.072

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends Neurosci, 37*(9), 480-490. https://www.doi.org/10.1016/j.tins.2014.06.004

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., & Schlaepfer, T. E. (2009). Nuclei accumbens phase synchrony predicts decision-making reversals following negative feedback. *Journal of Neuroscience, 29*(23), 7591-7598. https://www.doi.org/10.1523/JNEUROSCI.5335-08.2009

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. Neuroimage, 35(2), 968-978. https://www.doi.org/10.1016/j.neuroimage.2006.11.056

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E., & Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain Res, 1238*, 127-142. https://www.doi.org/10.1016/j.brainres.2008.07.114

Cohen, M. X., Wilmes, K., & Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends Cogn Sci, 15*(12), 558-566. https://www.doi.org/10.1016/j.tics.2011.10.004

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. J Cogn Neurosci, 24(1), 106-118. https://www.doi.org/10.1162/jocn_a_00114

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann N Y Acad Sci, 1104*, 70-88. https://www.doi.org/10.1196/annals.1390.002

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci, 3*(1), 27-38. https://www.doi.org/10.3758/CABN.3.1.27

Ernst, B., & Steinhauser, M. (2012). Feedback-related brain activity predicts learning from feedback in multiple-choice testing. Cognitive, Affective, & Behavioral Neuroscience, 12(2), 323-336. https://www.doi.org/10.3758/s13415-012-0087-9

Ernst, B., & Steinhauser, M. (2017). Top-down control over feedback processing: The probability of valid feedback affects feedback-related brain activity. Brain Cogn, 115, 33-40. https://doi.org/10.1016/j.bandc.2017.03.008

Feingold, J. (2011). Beta oscillations in frontal cortex and striatum represent post-processing of successful behavior (doctoral dissertation). Massachusetts Institute of Technology, Cambridge.

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *Journal of Neuroscience, 32*(35), 12087-12092. https://www.doi.org/10.1523/JNEUROSCI.1410-12.2012

Foerde, K., Race, E., Verfaellie, M., & Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: evidence from amnesia. *Journal of Neuroscience, 33*(13), 5698-5704. https://www.doi.org/10.1523/JNEUROSCI.5217-12.2013

Foerde, K., & Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience, 31*(37), 13157-13167. https://www.doi.org/10.1523/JNEUROSCI.2701-11.2011

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170. https://www.doi.org/10.1111/j.1469-8986.2007.00602.x

Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clin Neurophysiol, 126*(7), 1338-1347. https://www.doi.org/10.1016/j.clinph.2014.08.025

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://www.doi.org/10.1126/science.1102941

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279-2282. https://www.doi.org/10.1126/science.1066893

Haber, S. N., & Fudge, J. L. (1997). The primate substantia nigra and VTA: integrative circuitry and function. *Crit Rev Neurobiol, 11*(4), 323-342. https://www.doi.org/10.1615/CritRevNeurobiol.v11.i4.40

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biol Psychol, 71*(2), 148-154. https://www.doi.org/10.1016/j.biopsycho.2005.04.001

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*(6), 905-912. https://www.doi.org/10.1111/j.1469-8986.2007.00567.x

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

HajiHosseini, A., & Holroyd, C. B. (2015). Reward feedback stimuli elicit high-beta EEG oscillations in human dorsolateral prefrontal cortex. Sci Rep, 5, 13021. https://www.doi.org/10.1038/srep13021

Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. *Neurophysiology, 78*, 447-455.

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

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Res, 1105*(1), 93-101. https://www.doi.org/10.1016/j.brainres.2005.12.015

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology, 44*(6), 913-917. https://www.doi.org/10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cogn Affect Behav Neurosci, 9*(1), 59-70. https://www.doi.org/10.3758/CABN.9.1.59

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport, 22*(5), 249-252. https://www.doi.org/10.1097/WNR.0b013e328345441d

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport, 14*(18), 2481-2484. https://www.doi.org/10.1097/01.wnr.0000099601.41403.a5

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience, 7*(5), 497-498. https://www.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(5), 688-697. https://www.doi.org/10.1111/j.1469-8986.2008.00668.x

Janssen, D. J., Poljac, E., & Bekkering, H. (2016). Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. *Soc Cogn Affect Neurosci, 11*(8), 1280-1289. https://www.doi.org/10.1093/scan/nsw033

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*(5280), 1399-1402. https://www.doi.org/10.1126/science.273.5280.1399

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr Opin Neurol, 18*(4), 411-417. https://www.doi.org/10.1097/01.wco.0000173463.24758.f6

Koban, L., Pourtois, G., Bediou, B., & Vuilleumier, P. (2012). Effects of social context and predictive relevance on action outcome monitoring. *Cogn Affect Behav Neurosci, 12*(3), 460-478. https://www.doi.org/10.3758/s13415-012-0091-0

Kobza, S., & Bellebaum, C. (2015). Processing of action- but not stimulus-related prediction errors differs between active and observational feedback learning. *Neuropsychologia, 66*, 75-87. https://www.doi.org/10.1016/j.neuropsychologia.2014.10.036

Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., Sudmeyer, M., & Bellebaum, C. (2012). Dissociation between active and observational learning from positive and negative feedback in Parkinsonism. *Plos One, 7*(11), e50250. https://www.doi.org/10.1371/journal.pone.0050250

Murty, V. P., & Adcock, R. A. (2014). Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. Cereb Cortex, 24(8), 2160-2168. https://www.doi.org/10.1093/cercor/bht063

Lavoie, B., Smith, Y., & Parent, A. (1989). Dopaminergic innervation of the basal ganglia in the squirrel monkey as revealed by tyrosine hydroxylase immunohistochemistry. *J Comp Neurol, 289*(1), 36-52. https://www.doi.org/10.1002/cne.902890104

Lehéricy, S., Ducros, M., Van de Moortele, P. F., Francois, C., Thivard, L., Poupon, C., . . . Kim, D. S. (2004). Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann Neurol, 55*(4), 522-529. https://www.doi.org/10.1002/ana.20030

Li, P., Baker, T. E., Warren, C., & Li, H. (2016). Oscillatory profiles of positive, negative and neutral feedback stimuli during adaptive decision making. *Int J Psychophysiol, 107*, 37-43. https://www.doi.org/10.1016/j.ijpsycho.2016.06.018

Lynd-Balta, E., & Haber, S. N. (1994). The organization of midbrain projections to the ventral striatum in the primate. *Neuroscience, 59*(3), 609-623. https://www.doi.org/10.1016/0306-4522(94)90181-3

Marco-Pallares, J., Cucurell, D., Cunillera, T., Garcia, R., Andres-Pueyo, A., Munte, T. F., & Rodriguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. Neuropsychologia, 46(1), 241-248. https://www.doi.org/10.1016/j.neuropsychologia.2007.07.016

Marco-Pallares, J., Munte, T. F., & Rodriguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. Neurosci Biobehav Rev, 49, 1-7. https://www.doi.org/10.1016/j.neubiorev.2014.11.014

Mas-Herrero, E., Ripolles, P., HajiHosseini, A., Rodriguez-Fornells, A., & Marco-Pallares, J. (2015). Beta oscillations and reward processing: Coupling oscillatory activity and hemodynamic responses. Neuroimage, 119, 13-19. https://www.doi.org/10.1016/j.neuroimage.2015.05.095

Miltner, W. H., Braun, C. H., Coles, M. G. (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), 788-98. https://www.doi.org/10.1162/jocn.1997.9.6.788

Morelli, S. A., Sacchet, M. D., & Zaki, J. (2015). Common and distinct neural correlates of personal and vicarious reward: A quantitative meta-analysis. *Neuroimage, 112*, 244-253. https://www.doi.org/10.1016/j.neuroimage.2014.12.056

Mulligan, E. M., & Hajcak, G. (2017). The electrocortical response to rewarding and aversive feedback: The reward positivity does not reflect salience in simple gambling tasks. *Int J Psychophysiol*. https://www.doi.org/10.1016/j.ijpsycho.2017.11.015

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454. https://www.doi.org/10.1126/science.1094285

Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J Cogn Neurosci, 19*(12), 1994-2004. https://www.doi.org/10.1162/jocn.2007.19.12.1994

Opitz, B., Ferdinand, N. K., & Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Frontiers in Human Neuroscience, 5*, 8. https://www.doi.org/10.3389/fnhum.2011.00008

Peterburs, J., Kobza, S., & Bellebaum, C. (2016). Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). *Psychophysiology, 53*(2), 209-215. https://www.doi.org/10.1111/psyp.12560

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., & Smith, S. M. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. Biological Psychiatry, 55(6), 594-602. https://www.doi.org/10.1016/j.biopsych.2003.11.012

Schultz, W. (1997). Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol, 7*(2), 191-197. https://www.doi.org/10.1016/S0959-4388(97)80007-4

Schultz, W. (2000). Multiple reward signals in the brain. *Nat Rev Neurosci, 1*(3), 199-207. https://www.doi.org/10.1038/35044563

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599. https://www.doi.org/10.1126/science.275.5306.1593

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal phase locking to hippocampal theta oscillations. *Neuron, 46*(1), 141-151. https://www.doi.org/10.1016/j.neuron.2005.02.028

Simons, R. F. (2010). The way of our errors: theme and variations. *Psychophysiology, 47*(1), 1-14. https://www.doi.org/10.1111/j.1469-8986.2009.00929.x

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. Physiol Rev, 94(1), 35-79. https://www.doi.org/10.1152/physrev.00041.2012

van der Molen, M. J., Dekkers, L. M., Westenberg, P. M., van der Veen, F. M., & van der Molen, M. W. (2016). Why don't you like me? Midfrontal theta power in response to unexpected peer rejection feedback. *Neuroimage*. https://www.doi.org/10.1016/j.neuroimage.2016.08.045

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*(8), 1870-1884. https://www.doi.org/10.1016/j.neubiorev.2012.05.008

Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K., & Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring,

stimulus-response mapping, familiarity, and orienting. *Journal of Neuroscience, 25*(3), 604-613. https://www.doi.org/10.1523/JNEUROSCI.4151-04.2005

Wang, J., Chen, J., Lei, Y., & Li, P. (2014). P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *Neuroreport, 25*(13), 1044-1048. https://www.doi.org/10.1097/WNR.0000000000000226

Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (2012). Better late than never? The effect of feedback delay on ERP indices of reward processing. *Cogn Affect Behav Neurosci, 12*(4), 671-677. https://www.doi.org/10.3758/s13415-012-0104-z

Weismuller, B., & Bellebaum, C. (2016). Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. *Psychophysiology, 53*(11), 1739-1750. https://www.doi.org/10.1111/psyp.12738

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev, 111*(4), 931-959. https://www.doi.org/10.1037/0033-295X.111.4.939

Zaghoul, K. A., 2009, Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science,* 323(5920), 1496-9. https://www.doi.org/10.1126/science.1167342

Tables

Table 1

Mean FRN amplitudes (extracted from negative – positive feedback difference waves, in μV), standard deviation in parentheses

Learning Condition	Immediate Feedback	Delayed Feedback
Active Learning	-2.41 (1.87)	-1.28 (1.03)
Observational Learning	-1.08 (0.60)	-1.01 (1.13)

Figures

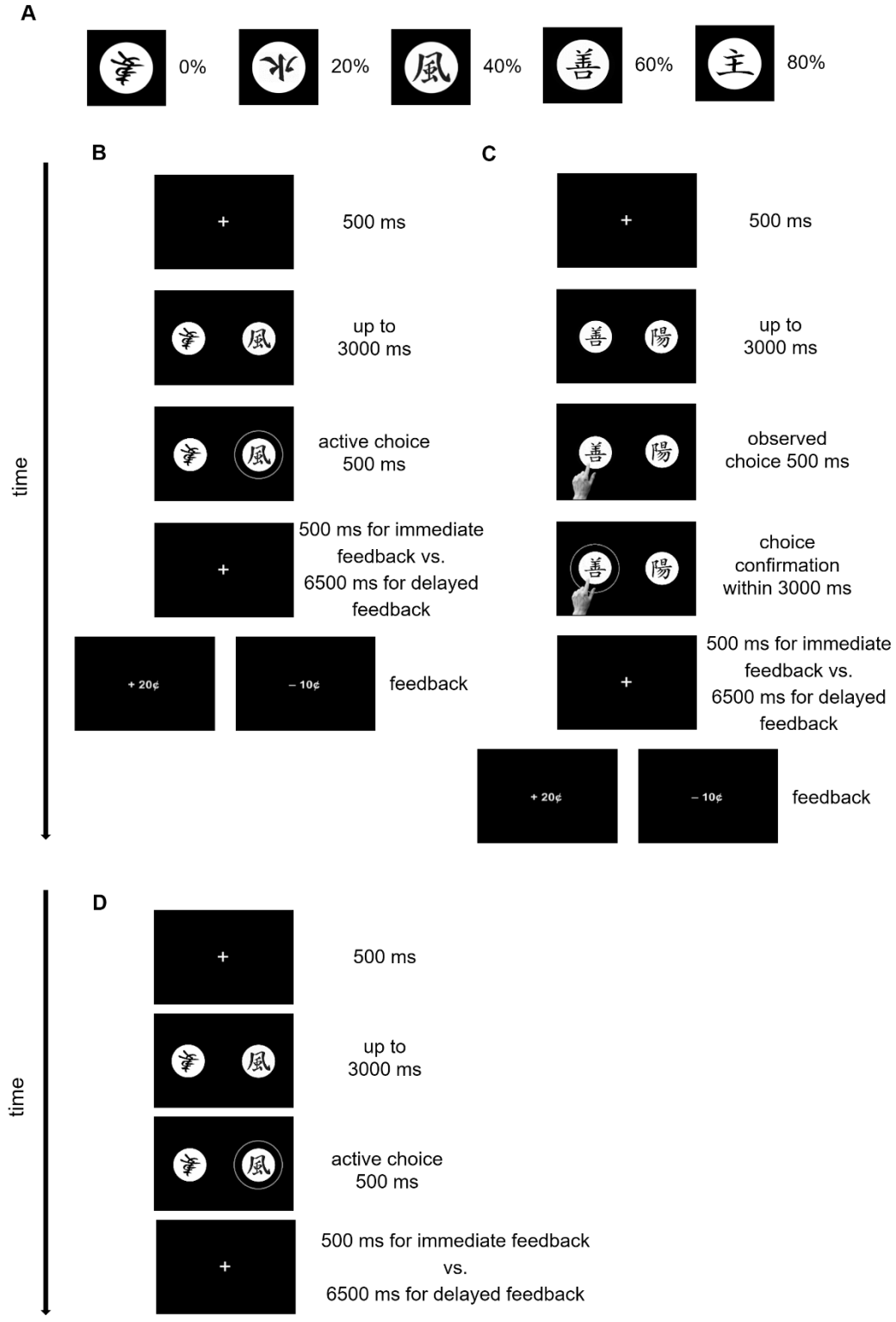

Figure 1. A) Symbols and reward probabilities used in the experiment. B) Time course of events in a single active learning trial. C) Time course of events in a single observational learning trial. D) Time course of events in a single test trial without feedback.

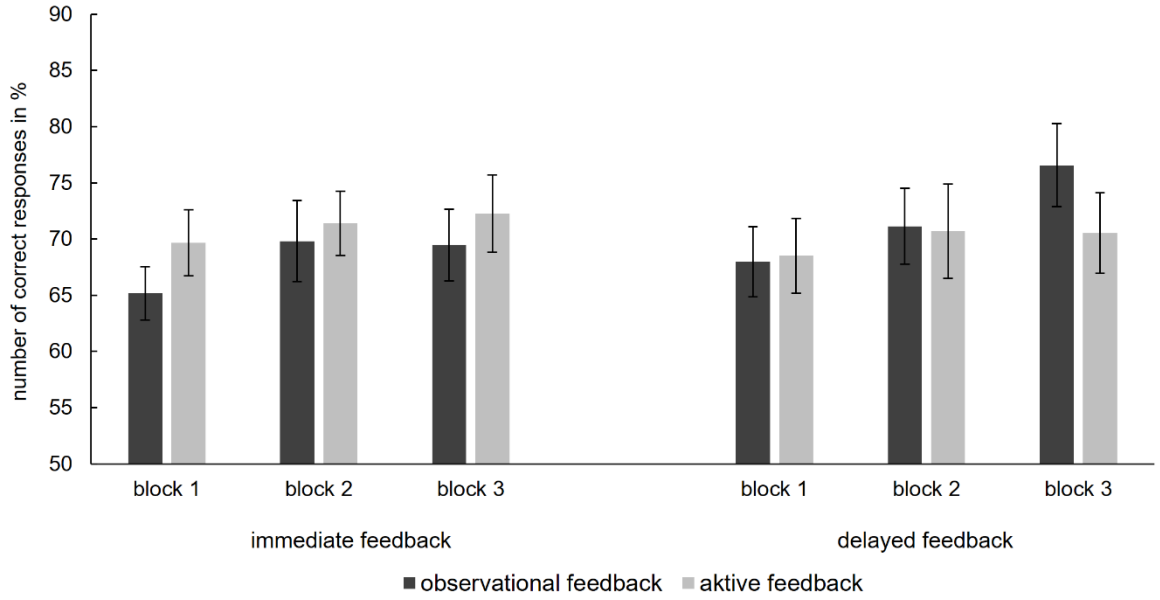

Figure 2. Choice accuracy across all reward probabilities in percent for each block (error bars indicate standard errors).

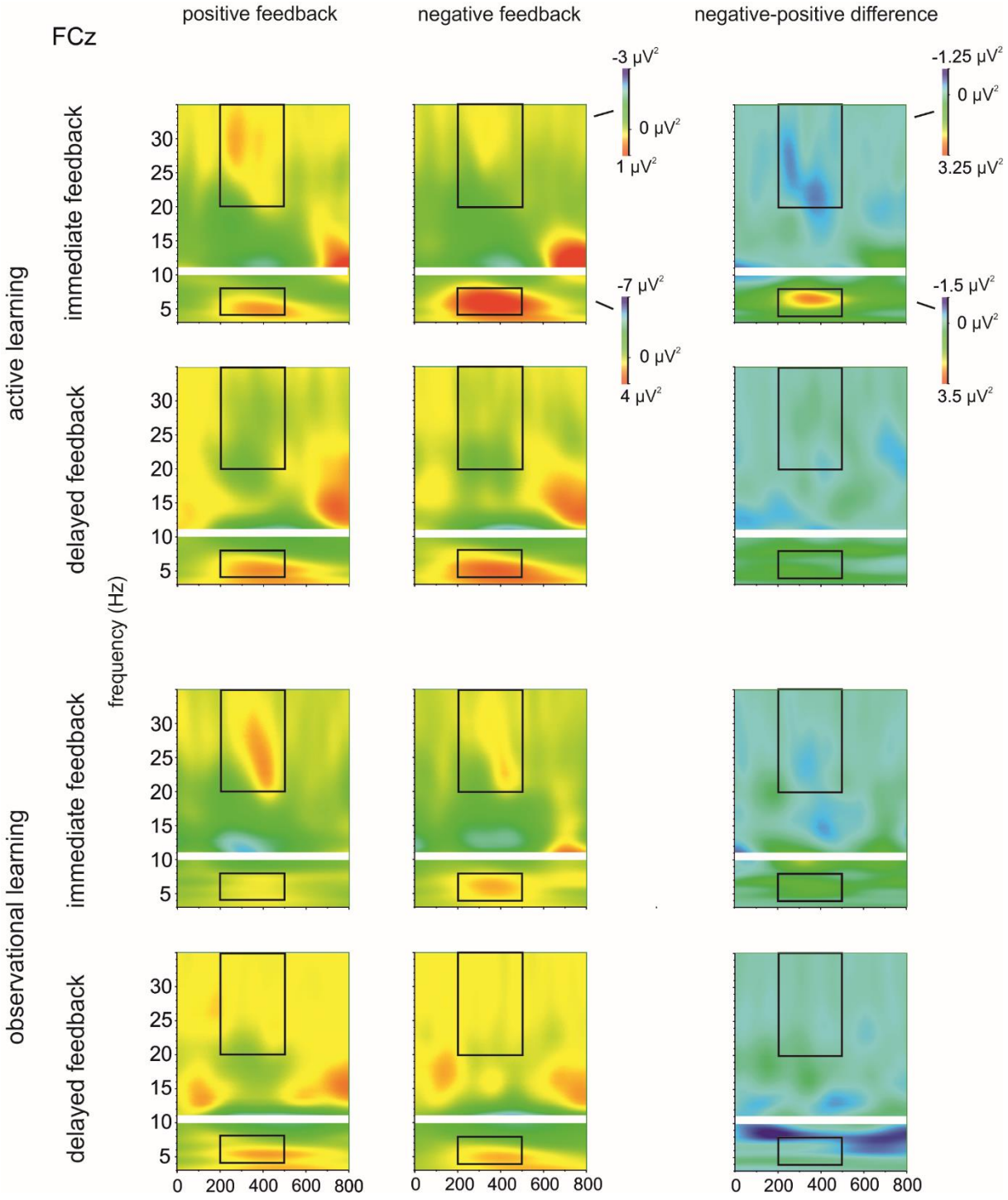

Figure 3. Grand-average time-frequency plots for positive and negative feedback and the difference between negative and positive feedback for both feedback timings and agency conditions at electrode site FCz. Plots represent the baseline-corrected induced spectral power (μV^2). The rectangles indicate the extracted time and frequency windows for theta (200-500 ms, 4.12-7.75 Hz) and beta power (200-500 ms, 19.63-35 Hz). Note that the colour scaling differs between lower frequencies (3-10 Hz) and higher frequencies (11-35 Hz) in each plot.

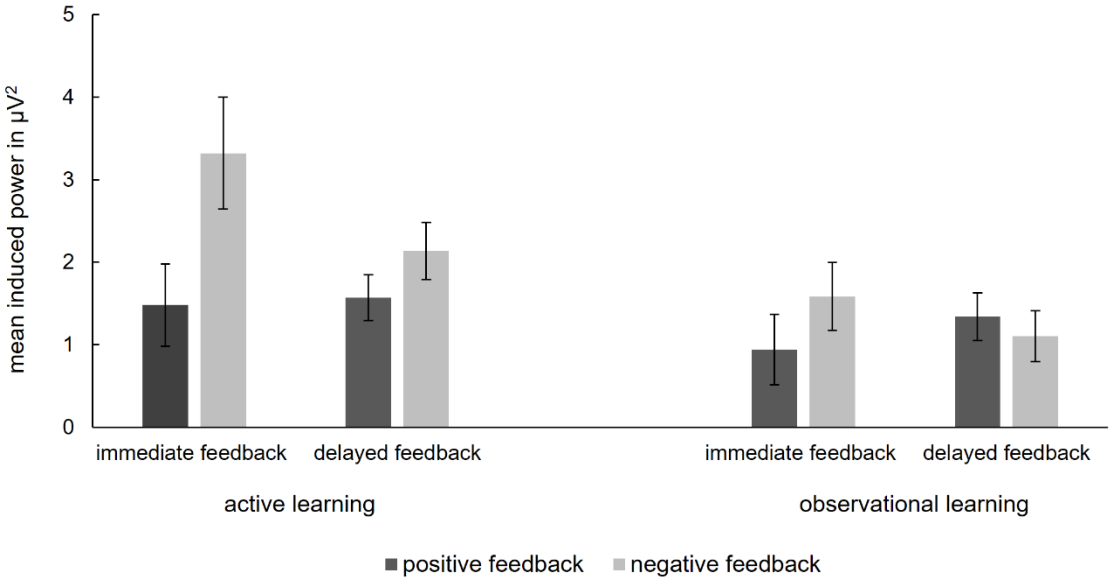

Figure 4. Mean induced theta power (μV^2) for each Feedback Timing, Agency, and Valence condition (error bars indicate standard errors).

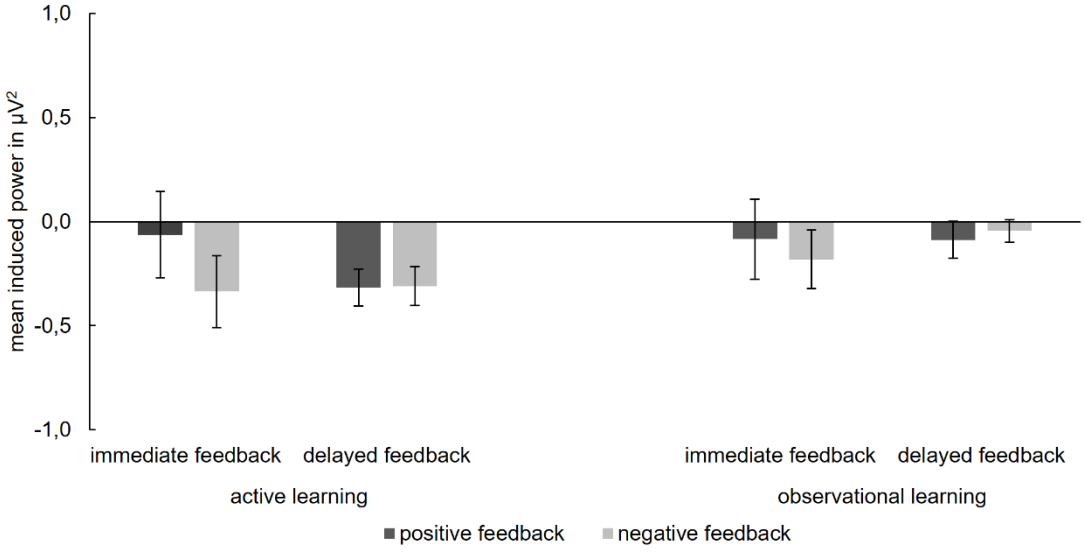

Figure 5. Mean induced beta power (μV^2) for each Feedback Timing, Agency, and Valence condition (error bars indicate standard errors).

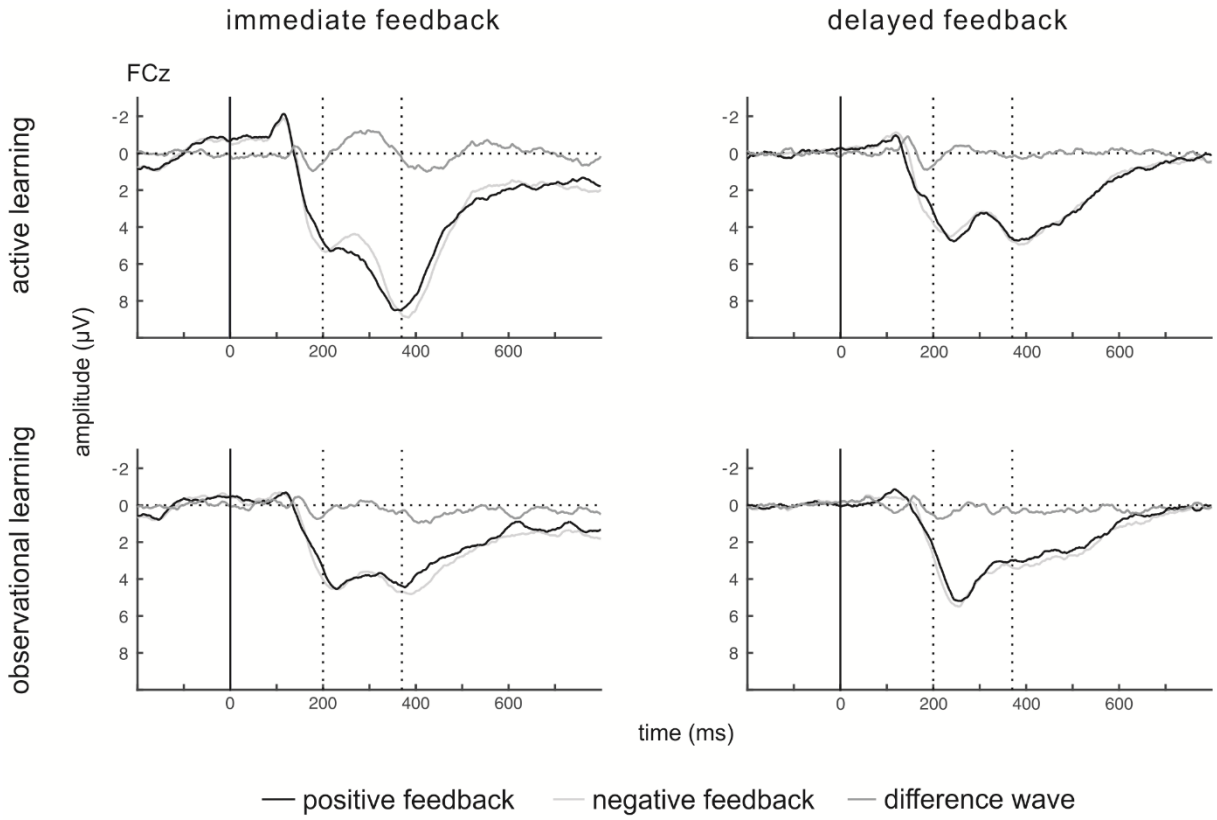

Figure 6. Feedback-locked grand average ERPs for positive feedback, negative feedback, and the negative-positive difference wave at FCz. Dotted vertical lines indicate the time window used to extract FRN peaks from the difference waves.