Dissertation

Titel der Arbeit

„Decision Making and Feedback Processing investigated with Event-Related Potentials“

Verfasserin

Mag. Daniela Pfabigan

Angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer. nat.)

Wien, im Dezember 2010

Studienkennzahl lt. Studienblatt: A 091 290

Dissertationsgebiet lt. Studienblatt: Psychologie

Betreuer: Univ.-Prof. Dr. Herbert Bauer

The Brain – is wider than the Sky -
For – put them side by side -
The one the other will contain
With ease – and You beside -

The Brain is deeper than the sea -
For – hold them Blue to Blue -
The one the other will absorb -
As Sponges – Buckets – do -

The Brain is just the weight of God -
For – Heft them – Pound for Pound -
And they will differ – if they do -
As Syllable from Sound -

Emily Dickinson, 1862

Acknowledgements

To keep a long list short, I would like to thank my family, all my friends, my colleagues at the lab and my participants for supporting me during the endeavour of planning, conducting and writing my PhD-thesis at the Faculty of Psychology in Vienna.

Completing my PhD-thesis wouldn’t have been possible without the support and encouragement of two people - Univ.-Prof. Dr. Herbert Bauer and PD. Dr. Uta Sailer - special thanks to you both!

Table of Contents

1. General Introduction

Our environment is constantly changing, thus behavioural flexibility is highly important in our daily lives. This flexibility requires continuous performance monitoring and decision making. Thus, monitoring of our own performance in everyday life is essential in order to adjust behavioural tendencies if necessary. External feedback, e. g. the observed behaviour of fellow human beings, traffic lights, noise, etc. influences our decisions. The ability to differentiate between favourable and unfavourable events or decision outcomes is a prerequisite to learn from these external cues to guide our behaviour.

Research on decision making and feedback processing has been of interest for several years now. The present thesis focuses on external feedback processing investigated with electrophysiological measures.

Event-related potentials (ERPs) were chosen to investigate feedback processing in healthy volunteers. ERPs can be considered as neuronal responses to specific internal and external stimuli and have been proven valuable to illustrate the strong relationship between electric brain activity and overt human behaviour (Andreassi, 2007).

2. Event-Related Potentials and Feedback Processing

2.1 Error-Related Negativity (ERN) and Feedback-Related Negativity (FRN)

2.1.1 Description of the two Components

In cognitive neuroscience research related to decision making and performance monitoring mainly focused on the investigation of negative consequences. To recap the paragraph on electrophysiology, event-related potentials (ERPs) of the on-going electroencephalogram (EEG) are a suitable neuroscientific approach to investigate unfavourable events and their consequences. The following section describes two ERP components which are assumed to be highly related to each other since both are reflecting different aspects of error processing, i.e. mechanisms of performance monitoring.

Internal error processing is thought to be reflected by the ERP-component called 'Error-Related Negativity (ERN; Gehring et al., 1993) or Error Negativity (Ne; Falkenstein et al., 1991)'. Typically, the ERN is elicited by errors peaking about 80-100 ms after error commission in simple reaction time tasks, thus reflecting subjective judgements of response accuracy (Scheffers & Coles, 2000). The negative deflection reaches maximum amplitudes over fronto-central electrode sites (Gehring et al., 1993). The size of the ERN amplitude depends on contextual factors. For example, studies emphasizing speed over accuracy in reaction time tasks, thus devaluating the subjective significance of errors, have shown a decrease in ERN amplitude (Falkenstein et al., 1995; Gehring et al., 1993). Furthermore, the ERN is reduced after erroneous responses to stimuli which occur rather infrequently (Holroyd & Coles, 2002).

External error processing, i.e. the processing of external feedback, is associated with another ERP component. Miltner and colleagues (1997) observed a negative ERP deflection about 200-350 ms after the presentation of external negative feedback cues at fronto-central electrode sites. These cues indicated an incorrect response in their time estimation paradigm, where subjects had to estimate the duration of one second after the visual presentation of an imperative cue. Since subjects had no internal representation as to whether they performed the estimation well or not, they had to rely on the external feedback to know whether they had to improve their performance or to maintain their response pattern. Miltner et al. (1997) labelled this ERP component Feedback-Related Negativity (FRN). The authors reported that an FRN could be evoked by negative feedback comprising of visual, acoustic, or somato-sensory feedback, thus they assumed that the FRN was independent of the physical appearance of the negative feedback. In 2002, Gehring and Willoughby presented a study where subjects were involved in a gambling task where they gained or lost real money. The stimuli comprising the visual information about a monetary loss also evoked a negative-going ERP component with a latency and scalp distribution comparable to the FRN. Gehring and Willoughby (2002) named this component, which was sensitive to loss feedback, Medial Frontal Negativity (MFN). Most researchers are in accordance now that the FRN and the MFN are neuronal signs of the same underlying mechanism related to external feedback processing. That's because Nieuwenhuis and colleagues (2004) were able to show that a fronto-central negative ERP deflection in the latency range of 200-350 ms could be elicited by either monetary or performance feedback. The authors provided subjects with an experimental paradigm where feedback simultaneously contained information about monetary or performance feedback. An FRN component was elicited by either one, only depending on which dimension was emphasized during the task. Thus, the FRN is thought to reflect an early outcome evaluation either based on a binary classification of good vs. bad (Hajcak et al., 2006), or whether a goal has been achieved or not (Holroyd et al., 2006).

Since there is no general agreement to call the response-locked ERP component ERN and the feedback-locked component FRN or MFN, the terms ERN, FRN, MFN are used synonymous in the whole manuscript.

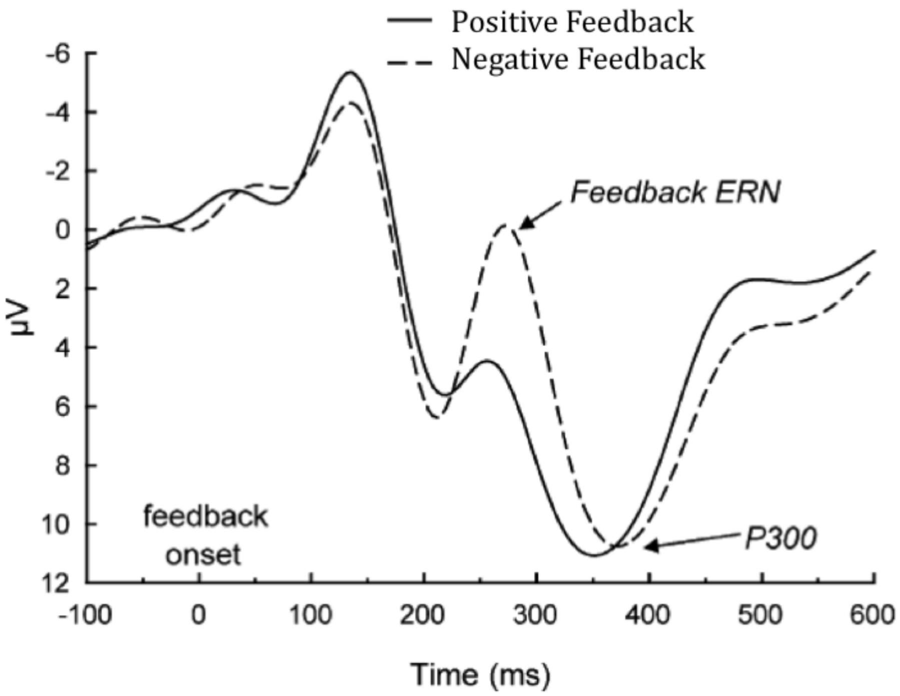


Figure 1.

Graphic adapted from Nieuwenhius and colleagues (2004). Typical amplitude courses after negative and positive feedback presentation at frontal electrode sites, the FRN component is observable between 200 and 300 ms after feedback onset, as depicted by the arrow and the text line.

2.1.2 Neuronal Generator of the ERN and FRN

Both ERN and FRN are assumed to be generated in the anterior cingulate cortex (ACC) – in particular in its caudal/dorsal portions - and in adjacent frontal regions, as reported by source localization methods (Gehring and Willoughby, 2002; Herrmann et al., 2004; Holroyd & Coles, 2002; Miltner et al., 1997) as well as functional brain imaging studies (Holroyd et al., 2004; Ullsperger & von Cramon, 2003). However, some studies failed to demonstrate ACC activation at all (Van Veen et al., 2004), or reported enhanced rostral ACC activation after positive feedback (Niewenhuis, Slagter et al., 2005).

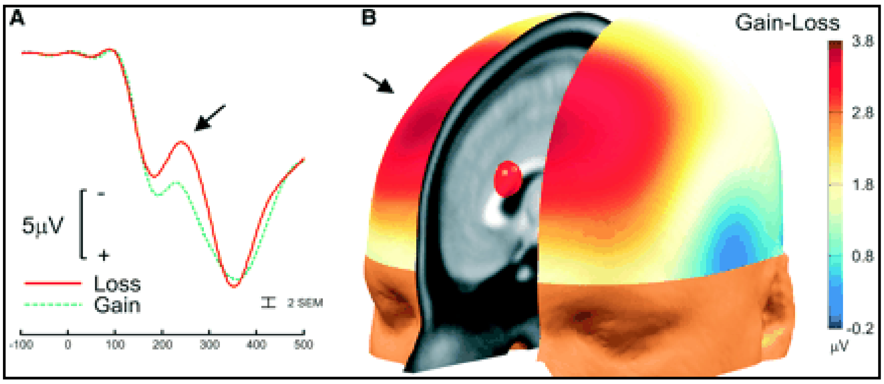


Figure 2.

Adapted from Gehring and Willoughby (2002); ERP waveforms at electrode location Fz, scalp topography, and probable neural generator of the FRN/MFN. (A) The red line indicates the grand-mean waveform for all trials where participants lost money; the green line indicates the grand-mean waveform for all trials where participants won money. The FRN/MFN component is indicated by an arrow. (B) Scalp topography computed at 265 ms after feedback onset, depicting voltage values derived from subtracting the loss-waveform from the gain-waveform. The color red indicates a greater FRN/MFN effect. The red sphere indicates the best-fitting dipole model of the FRN/MFN component which is centred in the ACC.

The ACC (BA[1] 24, BA 32) is a cortex region located on the medial surface of the frontal lobes. It is of importance for the integration of cognitive, affective, and visceral information (Allman et al., 2001; Critchley, 2005; Thayer & Lane, 2000).

The nomenclature of ACC sub-divisions has been refined recently. Vogt (2005) subsumes recent structural and functional observations under a four-region neurobiological model of cingulate cortex. Firstly, the term ACC refers to the most anterior parts of the cingulate cortex; the subgenual anterior cingulate cortex (sACC) and the pregenual anterior cingulate cortex (pACC) respectively. The ACC is reported to be involved in autonomic control and emotion processing via extensive connections to different nuclei of the amygdale (Vogt, 2009). Secondly, medial parts of the cingulate cortex are labelled MCC. The MCC is divided into anterior medial cingulate cortex (aMCC) and posterior medial cingulate cortex

---

[1] BA: short for Brodmann Area.

(pMCC). In general, the MCC is reported to be involved in response selection. It incorporates two separate motor areas which project to the spinal cord and to motor cortices. Furthermore, the MCC is involved in coding the reward value of behavioral outcomes. Anterior parts of the MCC share connections with the amygdale, posterior parts project to posterior parietal cortices. Thirdly, posterior parts of the cingulate cortex are labelled PCC. The PCC is involved in visuospatial orientation (dorsal posterior cingulate cortex – dPCC) and the assessment of self-relevance (ventral posterior cingulate cortex – vPCC) via parietal lobe connections. Fourthly, the retrosplenial cortex (RSC) is associated with memory formation and memory access.

Prior to the four-region model of cingulate cortex (Vogt, 2005), Bush and colleagues (2000) reported a dichotomy of affective and cognitive sub-divisions of the cingulate cortex. Dorsal parts were assumed to be involved in cognitive processes such as the assessment of motivational significance of external stimuli, action monitoring (Devinsky et al., 1995), and error processing (Carter et al., 1998). Rostral-ventral parts were assumed to be involved in affective processing (Bush et al., 2000). This dichotomy of cognitive and affective sub-regions of cingulate cortex can also be found in Vogt's four-region model for the cyto-architectural border between ACC and MCC.

Therefore, future studies investigating the neuronal generators of ERN and FRN should refer to the term MCC instead of the term ACC.

2.1.3 Theoretical Background of the ERN and FRN

There are several theoretical frameworks trying to give a reasonable account of both the ERN and the FRN.

Error Monitoring System

At first, ERN and FRN were simply interpreted in terms of operations of an error-processing system (Gehring et al., 1993; Miltner et al., 1997). Gehring (1992) suggested the ERN to be sensitive to the degree of an error. Miltner and colleagues (1997) were the first to propose that the ERN and the FRN were functionally similar processes. They interpreted these two components as indicators of a generic error detection system. This system should be involved in the comparison of an actual response and knowledge about an intended one. If the system detects a mismatch between these two, an ERN or FRN is elicited (Coles et al., 2001; Falkenstein et al., 1991; Gehring et al., 1993; Scheffers & Coles, 2000). Furthermore, the ERN was thought to index the need for error correction or error compensation as proposed by Gehring and colleagues (1993). However, this assumption was adjusted by Scheffers and co-workers (1996). These authors found that an ERN is observable even in cases without the possibility of response correction, thus disproving a mere error correction function of the ERN.

Conflict Monitoring Theory

Building on the aforementioned theories, Botvinick and colleagues (2001) related ERN and FRN activity to the aspect of conflict monitoring. The authors postulated that the ACC is the instance that monitors decision processes and response outcomes. In cases the ACC detects conflict, e. g. conflicting response tendencies, ACC activity is reflected by these electrophysiological conflict signals, i.e. the ERN or FRN. Based on such ACC activities behavioural errors could be easily detected and hence avoided, since the ACC is conveying this information to brain regions directly responsible for the control of cognitive processing, e. g. to the lateral prefrontal cortex (Botvinick et al., 2001; Carter et al., 1998). Botvinick and co-workers (2001) described the variable ERN or FRN size in a way that it might reflect the magnitude of the perceived response conflict. Thus, a rather marginal response conflict should be reflected in rather small ERN and FRN amplitudes, whereas a large response conflict should be reflected in larger ERN and FRN amplitudes. To be more specific, response conflicts may be exemplified by means of situations where multiple responses compete for the control of action (Yeung et al., 2004). Yeung and colleagues (2004) reported three possible situations where an ERN could be observed. The component is evoked either after an overt response error in choice reaction-time tasks, or after external feedback about response accuracy, or after late responses in choice reaction-time tasks when speed is emphasized over accuracy (Johnson et al., 1997). Furthermore, the converse argument is applicable because the presence of enhanced response conflict causes higher probability of error commission since more attention resources are necessary for solving the conflicting situation.

The conflict monitoring account of the ERN further suggests that a negative deflection should by observable after a correct response when the experimental trial contained highly conflicting elements. Based on this suggestion, Botvinick and colleagues (2001) developed a connectionist model for conflict monitoring. This model suggested a main distinction between error and correct trails. Error trials might be characterized by response conflicts in the period after the response, whereas correct trails might be characterized by response conflicts which occur mostly prior to the response. Therefore, Botvinick and co-workers (2001) argued that the N2 component (a negative deflection of the ERP about 200 ms after stimulus presentation) is the possible physiological correlate of the cognitive conflict prior to a correct response.

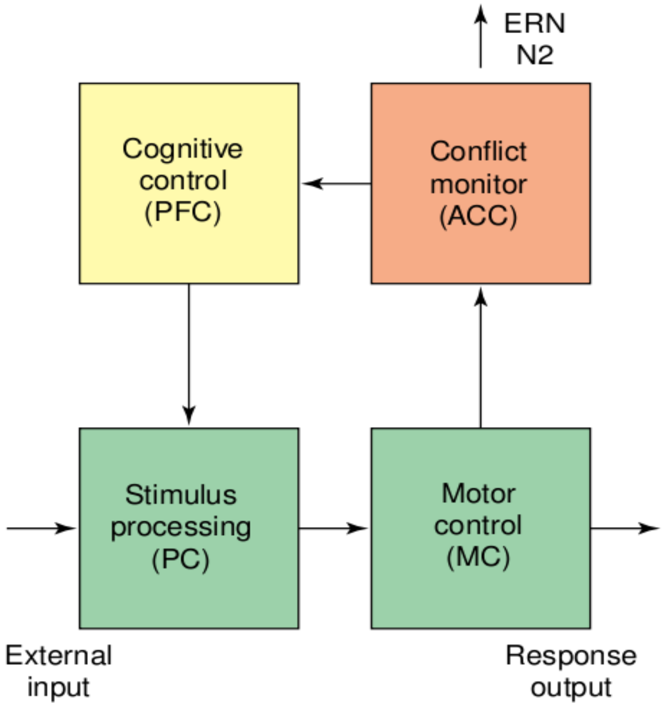


Figure 3.

Model of conflict monitoring (adapted from Holroyd and Yeung, 2003). The red box indicates mechanisms that monitor performance; the green boxes indicate mechanisms that map external input into response outputs. The yellow box indicates a separate mechanism for cognitive control. Abbreviations: ACC, anterior cingulate cortex; MC, motor cortex; PC, posterior cortex; PFC, prefrontal cortex.

Reinforcement Learning Theory

Another influencing theoretical account for ERN and FRN generation is the reinforcement learning theory of the ERN (RL-theory) proposed by Holroyd and Coles (2002). The authors suggested a response-monitoring system within the basal ganglia producing a specific error signal (reflected by the ERN or FRN) whenever detecting that the outcome of an event is worse than expected. This error signal would be conveyed to different cortical regions, amongst them the ACC, via a phasic decrease in tonic mesencephalic/dopaminergic activity. Thus, ERN and FRN appear to reflect the activation of this reinforcement learning system, which rapidly evaluates decision outcomes to guide reward seeking behaviour (Holroyd & Coles, 2002). Nieuwenhuis and colleagues (2004) elaborated the theoretical framework of the RL-theory. Firstly, they proposed that the ERN/FRN indicates the evaluation of decision outcomes along an abstract dimension 'good versus bad' (c. f. Hajcak et al., 2006). Secondly, they stated that the amplitude of the ERN/FRN depends on the relation of the expected versus the actual outcome. Thirdly, Nieuwenhuis and co-workers (2004) summarized that the FRN amplitude varies inversely to the ERN amplitude (Holroyd & Coles, 2002). This has been observed in gambling paradigms, where subjects used external feedback to learn specific cue-response contingencies. In these cases, the negative ERP deflection virtually 'propagated back' in time, namely from the negative feedback stimuli at the begin of learning, to the motor response, when individuals had learnt the contingency but made nevertheless an error. And finally, they agree with the notion that the ERN/FRN is generated in the ACC. Furthermore, Holroyd and Coles (2002) suggested that the ERN/FRN could be seen as a sign for a reward prediction error indicating the difference between a reward received and a predicted one. To be more specific, this so-called reward prediction error may also incorporate information about the next prediction mode by the reward system (Bertsekas & Tsitsiklis, 1996). To convey these theoretical assumptions in a physiological framework, the work of Schultz (1998) and Montague and colleagues (2004) was used. Holroyd and colleagues (2003) summarized that the RL-theory account of a negative reward prediction error signal (i.e., worse than expected; which is thought to induce a disinhibition of the ACC, thus leading to a negative deflection in the ongoing EEG) is caused by a phasic decrease of mesencephalic dopamine activity. These thoughts led to a major prediction of the RL-theory: the FRN amplitude should be positively related the size of the prediction error (i.e., numerically larger FRN amplitudes can be expected after highly unexpected events), thus implying that the relation between the expected and the actual outcome is the most important variable influencing the actual ERP amplitude (Hajcak, Holroyd et al., 2005).

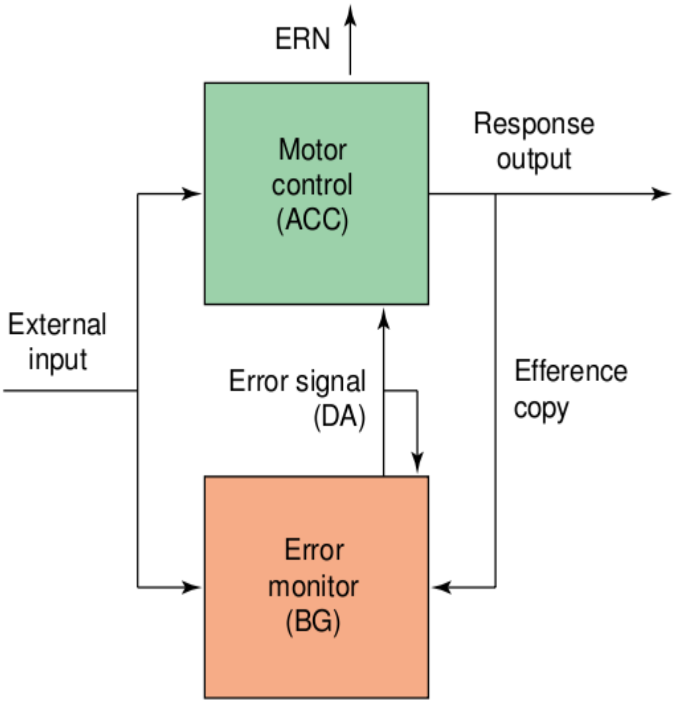


Figure 4.

Model of RL-theory (adapted from Holroyd and Yeung, 2003). The red box indicates mechanisms that monitor performance; the green box indicates mechanisms that map external input into response output. Abbreviations: ACC, anterior cingulate cortex, BG, basal ganglia; DA, mesencephalic dopamine system.

The mathematical base of RL-theory is the so-called temporal difference learning (TD learning) account. TD learning is a prediction method mostly used for solving reinforcement learning problems. It combines a Monte Carlo method, i.e. the model learns by sampling according to a specific strategy, with bootstrapping methods, which means that current estimates are based on approximations of previously acquired estimates (Sutton & Barto, 1998). Schultz and colleagues (1997) reported that the firing rate of dopaminergic neurons in the ventral tegmental area (VTA) and the substantia nigra (i.e., parts of the mesencephalic dopamine system) seems to imitate the TD learning algorithm in monkeys, thus relating the TD learning account to neuroscience. Following this notion, Schultz (1998) trained a monkey to associate a specific cue with a rewarding stimulus. When the monkey was initially confronted with the rewarding stimulus its dopamine cells of the aforementioned areas increased their firing rates, thus indicating differences between the expected and the actual reward. After several repeated presentations the increase in firing rate was related to the specific cue indicating the reward. Finally, when the monkey had completely learned the cue-response contingency, no increase in firing rate for the presentation of the rewarding stimuli was observable any more.

It should be emphasised that the function of the TD learning account is not simply reflecting the difference between a received reward and a predicted one, as the Rescorla/Wagner model[2] does (Rescorla & Wagner, 1972), which renders the TD learning account so useful. Instead, information about the next prediction is incorporated in the error signals (Bertsekas & Tsitsiklis, 1996), which is stated more precisely in the following simplified formula by Montague and colleagues (2004; reward prediction error hypothesis):

Current TD error = current reward + δ * next prediction – current prediction

> δ reflects a coefficient between 0 and 1 weighting the relative influence of the next prediction

Some evidence supporting the RL-theory is also revealed by neuropsychology. Prefrontal and basal ganglia lesions disrupt the emergence of an ERN, thus suggesting that the fronto-striato-thalamo-cortical loops proposed by Holroyd and Coles (2002) are necessary pre-requisites to generate this ERP component (Ullsperger & von Cramon, 2006).

Theories regarding the Violation of Expectations

Recently, Oliveira and colleagues (2007) postulated an expectancy deviation hypothesis to

---

[2] The Rescorla/Wagner model (1972) is an account for Pavlovian conditioning. The model predicts that differences between expected and unexpected outcomes are evident only during the first observations.

explain neuronal processes involved in decision making. The authors suggested in their theoretical account that it is not the unfavourable feedback per se which evokes the mentioned error ERP components. In their opinion, it is the violation of previously built-up expectations that could be detected by the ACC which is responsible for the error-related ERPs. Furthermore, they explain larger ERN/FRN amplitudes after an error or negative feedback by an overoptimistic bias which strongly influences our performance estimates and resulting expectations (Miller & Ross, 1975). Comparable conclusions were drawn by Wu and Zhou (2009) in their study investigating prediction errors. According to the authors the FRN might reflect a reward prediction error which is not only defined in terms of valence. They suggested that the FRN indicates also information about whether or not a pre-established expectancy about an event is full-filled, irrespective whether experimentally introduced block-wise or trial-wise.

In line with the RL-theory Yasuda and co-workers (2004) reported the FRN responds more to negative reward prediction error signals (i.e., the events that were worse than expected) than to the error feedback per se. They speculated that the FRN might merely indicate a neuronal signal modifying the behavioural response strategy or that the FRN enhancement after unexpected negative outcomes might be due to this surprising outcome. However, Hajcak, Holroyd et al. (2005) reported the FRN being equally large for expected and unexpected feedback. Holroyd et al. (2006) extended the RL-theory and claimed that the FRN is indicating whether a goal has been achieved or not.

ERN and FRN indicating motivationally salient Events

Following the evolution of theoretical concepts of the ERN/FRN, interpretations changed over the time.

Gehring and Willoughby (2002) and Luu and colleagues (2003) came to the agreement that the FRN is either reflecting the affective significance or the emotional valence of the eliciting stimuli. Surprisingly, Yeung and colleagues (2005) and Donkers and colleagues (2005) observed an FRN component even in paradigms where no overt motor response was required. According to their finding they state that the ERN/FRN is sensitive to task demands. For example, the ERN amplitude is correlated with subjective judgements of response accuracy (Scheffers & Coles, 2000). Furthermore, the ERN amplitude is enhanced when response accuracy is emphasized over speed (Falkenstein et al., 1995;

Gehring et al., 1993). And the ERN amplitude is reduced with incorrect responses to infrequent stimuli, e. g. in conditions in which errors are particularly likely (Holroyd & Coles, 2002). Following Yeung and colleagues (2005) notion, the FRN is also numerically larger with overt responses required than without any responses.

2. 1. 4. Correct Response Negativity (CRN)

There is another controversy regarding ERP components and feedback processing. Are there specific ERP components following correct reactions or positive feedback? Vidal and colleagues (2000, 2003) reported a small response-locked negative-going ERP component on correct trials. The authors labelled this component 'Correct Response Negativity (CRN)'. They also assumed that the CRN shares morphological and topographical properties with the ERN, which would imply that both components might index response monitoring processes. In line with this assumption, Ridderinkhof and colleagues (2004) hypothesized that CRN amplitude might be related to response control performance measures. And indeed they observed increased accuracy and reduced reaction time interference from incompatible stimuli in trials following those with large CRN amplitudes. Thus, it might well be that ERN and CRN are comparable medio-frontal negativities related to evaluation during response monitoring. Coles and colleagues (2001) related an observed CRN in correct trails either to the possibility that error-processing is also present with correct responses, or to methodological issues deriving from the response-locked average.

Later on, Hajcak, Moser and colleagues (2005) suggested that there possibly is a functional difference between ERN and CRN, since their data showed no increase in CRN amplitude during more significant correct trials, but an increase in ERN amplitude during more significant error trials. Applying a gambling paradigm, Hajcak and co-workers (2006) observed small negative ERP deflections after feedback indicating a reward (i.e., a correct response choice). This observation found was supported by another study by Hajcak and colleagues (2007), where once again small FRNs were found after rewarding feedback.

At the same time, also a positive ERP component in the ERN/FRN latency range after positive feedback has been claimed by other researchers (Holroyd et al., 2003; 2008; Potts et al., 2006). Potts and colleagues (2006) used a passive S1-S2 reward prediction paradigm where they observed positive-going components. Holroyd and colleagues (2008) calculated difference waves of oddball and feedback conditions which also yielded positive ERP deflections. Thus, both studies have only restricted value for comparison to the remaining majority of ERN and FRN studies since they used rather rare experimental paradigms.

2.2 P300

2.2.1 Theoretical Background of the P300

The P300 component is another valuable indicator of performance monitoring. In general, the P300 wave (also P3 or classical P3b) is described as a positive deflection of the ERP, peaking between 300-600 ms after stimulus onset at posterior electrode sites (Duncan-Johnson & Donchin, 1977; Johnson & Donchin, 1980). Originally, the P300 was investigated with acoustic oddball paradigms. With this kind of paradigm subjects were required to respond to (e. g., pressing a button, raising a finger, or count) infrequent target tones and withhold that response with frequent target tones. The distinction of the two tones was possible due to different sound level or pitch. The P300 component appears to be dependent on various factors, such as categorical stimulus probability (Johnson & Donchin, 1980; Kutas et al., 1977), stimulus quality, attention (Polich & Kok, 1995), task relevance of the stimulus (Coles et al., 1995), or motivational significance of the presented stimuli (Yeung & Sanfey, 2004). Furthermore, the P300 amplitude is also an index of task demands, task complexity and resource allocation (Israel et al., 1980). Most researchers consider the P300 to play an important role in recognition and memory-updating processes. Donchin and Coles (1988) suggested that the P300 amplitude is an indicator for context-updating in working memory. Only recently, some researchers reported a relation of P300 and motivational stimulus significance (Nieuwenhuis, Aston-Jones et al., 2005; Yeung & Sanfey, 2004). Varying P300 amplitudes were found in tasks requiring decisions or outcome evaluations (Hajcak et al., 2005; 2007; Luu et al., 2009; Sato et al., 2005; Yeung & Sanfey, 2004; Yeung et al., 2005). All these authors assumed that P300 amplitude modulation might reflect the evaluation of the functional significance of the stimulus at hand that passed on external feedback. Supporting this view, Nieuwenhuis, Aston-Jones and colleagues (2005) proposed that the P300 might code the motivational significance of rewarding stimuli along a good versus bad dimension. This coarse categorization of stimulus evaluation was also attributed to the FRN (Hajcak et al., 2006).

However, the results regarding the P300 as an indicator of stimulus valence are still controversial; contradicting observations were reported, that larger P300 amplitudes might be related to negative valence (Cohen et al., 2007; Ito et al., 1998), to positive valence (Bellebaum & Daum, 2008; Hajcak et al., 2007), or to be uncorrelated with stimulus valence, at all (Sato et al., 2005; Yeung and Sanfey, 2004).

2.2.2 Neuronal Generators of the P300

The temporo-parietal junction (TPJ) has been suggested as potential generator structure for the P300 (Smith et al., 1990), as well as several other areas including parietal and frontal sites (Ardekani et al., 2002).

3. Psychopathology and Feedback Processing

Psychopathology is associated with a variety of cognitive and affective deficits. Common phenomena in psychopathology are impairments in action monitoring. For example, they manifest in prolonged response inhibition time (Enticott, Ogloff & Bradshaw, 2008) and reduced ERN amplitudes (Alain, McNeely, He, Christensen & West, 2002) in patients suffering from schizophrenia, as well as enhanced ERN amplitudes in obsessive-compulsive disorder (Gehring, Himle & Nisenson, 2000).

A construct generally related to psychopathology is the antisocial personality disorder (DSM-IV [American Psychiatric Association]; but see also dissocial personality disorder of the IDC-10 [World Health Organization]), and the construct of psychopathy.

3.1 Antisocial (Dissocial) Personality Disorder

Although the diagnosis of an antisocial personality disorder (DSM-IV) is considered to be the corresponding diagnosis to the dissocial personality disorder (ICD-10), some discrepancies arise when comparing the two. To start with, both diagnostic manuals share the following characteristics of the disorder: (i) lack of respect for social norms, obligations and irresponsibility, (ii) reckless, irritable, violent and aggressive behaviour, and (iii) lack of remorse and guilt. However, lack of empathy and the inability of maintaining lasting relationships is only incorporated into the ICD-10 diagnostic scheme, whereas the facets of repeated lying and conning others for personal benefit and pleasure, high levels of impulsivity, and reckless disregard for safety (both for oneself and for others) are only listed in the DSM-IV diagnostic scheme (Rodrigo et al., 2010).

Psychological measures

In the present PhD-thesis the personality questionnaire "Persönlichkeits-Stil und Störungs Inventar" (PSSI, Kuhl & Kazén, 1997) was administered to each participant prior to EEG data collection. The PSSI is a self-assessment tool comprising 14 scales assessing the relative manifestation of 14 non-clinical personality traits that cover the non-pathological diagnostic criteria of personality disorders of the DSM-IV as well as the ICD-10. In particular, the first sub-scale of the PSSI, the so-called antisociality (AS) scale was of interest for the present project. Reliability of this specific sub-scale (Cronbach's α = 0.86) as well as its validity are reported to be satisfactory. The scale consists of ten items, characterizing individuals according to self-determined and inconsiderate behavior to achieve their individual goals. Furthermore, individuals scoring high on this sub-scale are described to act self-confidently, offending, and humiliating during the interaction with others, as well as having problems in adjusting to social and legal norms. Each item has to be rated in the range of four options: 'does not apply – 'applies in some ways' - 'applies predominantly' – 'applies completely'. The following is an example item, "*If people turn against me I can wear them down*".

The AS sub-scale is not supposed to be sensitive to clinical levels of antisociality though (Kuhl & Kazén, 1997), thus identifying only individuals with moderate levels of antisocial personality characteristics, but it does not account for pathological symptoms of the personality disorder.

3.2 Psychopathy

Psychopathy is mostly described as personality disorder, although it is not included in the present version of the ICD-10 or DSM-IV. However, the construct of psychopathy is described by various behavioural tendencies and personality characteristics which can be found in diagnostic manuals. A combination of superficial charm, persistent instrumental antisocial behaviour, marked sensation-seeking, poor ability for reflection, blunted empathy and shallow emotional experiences is thought to represent a prototypical psychopathic individual (Hare, 2003). The idea of the construct itself was raised by Cleckley who gave a clinical description of the construct in his book *The Mask of Sanity* (1941). Cleckley mainly described apparently good functioning persons with a covered disturbance. The description of these individuals included interpersonal (egocentricity, lovelessness, impersonal sexuality and superficial charm), and emotional (affect impairments, lack of nervousness and guiltlessness) characteristics, as well as disinhibited or antisocial behaviour (see Fowles & Dindo, 2009), thereby being a great burden for society and acquainted individuals. The concept has been developed further over the last decades.

Psychological measures

Hare provided a renewed conceptualization as well as an instrument to measure psychopathy (Psychopathy Checklist-Revised, PCL-R; Hare, 2003). The PCL-R is a semi-structural interview combining information of charts and professional ratings. Contrary to the concept of dissocial personality disorder and antisocial personality disorder, Hare's concept of psychopathy combines specific personality traits and antisocial behavioural tendencies (Hare & Neumann, 2008). The ICD-10 and DSM-IV, however, characterize these personality disorders only on the behavioural level. There are several other measures to capture psychopathic personality traits, such as the Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996), which is a self-description inventory to assess psychopathy-related characteristics in non-criminal samples though. There is also a German version of this inventory available (Psychopathic Personality Inventory-Revised, PPI-R; Alpers & Eisenbarth, 2008).

Factor-analytic approaches have revealed a two-factor solution regarding the theoretical concept behind the PCL-R (Hare, 2003). Factor 1 is characterized by emotional-interpersonal deficits indicating core features of psychopathy (e. g., superficial charm, manipulative behavior, lack of remorse or guilt, shallow affect, etc.) - “primary psychopathy”, Factor 2 is characterized by impulsive antisocial behavior (e. g., poor behavioral control, impulsivity, irresponsiblity, etc.) - “secondary psychopathy”. Both PCL-R factors correlated moderately with each other (Hare, 2003). Lilienfeld and Andrews (1996) administered the PPI to a large population of non-criminal individuals. Factor analysis of this data revealed two factors (PPI-I - “fearless dominance”, PPI-II - “impulsive antisociality”) parallel to those of the PCL-R, although these two were not correlated. Fowles and Dindo (2006) as well as Patrick (2007) have stressed theoretical implications of the two-factor model of psychopathy. Factor 1 might be associated with a pattern of low fear and anxiety, and strong reward motivation that leads to reward-seeking behavior lacking fear of consequences or concerns for others. In comparison, Factor 2 might be merely associated with impulsivity and disinhibition leading to chronic antisocial behavior and antagonism towards others. Furthermore, Patrick (2007) suggested a relation of Factor 2 with externalizing psychopathology, a heritable personality dimension in young adults which is considered to be a risk factor for antisocial personality disorder, as well as substance and alcohol abuse.

To summarize, the diagnosis of psychopathy is associated with increased scores on items related to both factors.

3.3 Biological Origins of Psychopathy

Recent findings suggest viewing psychopathy as a developmental disorder (Lynam et al., 2007) with first symptoms emerging in childhood. Two current theoretical accounts stress either attentional or emotional dysfunctions in individuals scoring high on psychopathy measures.

In 1983, Hare and Jutai first linked psychopathy to attentional abnormalities. Recently, an attention-based model of psychopathy was proposed by Lorenz and Newman (2002). The authors observed reduced response modulation in psychopaths. They interpreted this as a result of inadequate processing of the meaning of peripheral or incidental information, which is not in the attention focus of the psychopathic individual. In line with this proposal is an observation by Raine and Venables (1988) when administering a visual performance task to psychopaths and healthy controls and recording EEG. The psychopathic individuals displayed larger and delayed P300 amplitudes. Thus, Raine and Venalbes (1988) assumed that these results might indicate that psychopaths showed enhanced ability to attend to stimuli of interest. This observation points towards the possibility of performance proficiencies of psychopaths under specific circumstances (Raine, 1989). Recently, Sadeh and Verona (2008) summarized that Factor 1 psychopathic traits were associated with this over-focussed attention on motivationally salient stimuli. Furthermore, these authors showed that Factor 2 psychopathic traits were associated with deficits in cognitive control as indexed by deficits in working memory functions (Sadeh & Verona, 2008).

Since psychopathy is linked to emotional dysfunctions and antisocial behaviour, the second theoretical account – emotional dysfunctions in psychopathy – has gained growing support (Blair et al., 2005; Frick & Morsee, 2006; Lykken, 1995). Blair (2010) subsumes in his review on neuroimaging and psychopathy that a disruption of the functioning of the amygdala, the superior temporal cortex, as well as the orbitofrontal cortex (OFC) has been repeatedly associated with psychopathic tendencies (de Oliveira-Souza et al., 2008; Tiihonen et al., 2008). Furthermore, Blair (2010) claims that these observed dysfunctions were specific for individuals with psychopathic traits and cannot be observed in any other patient group. Further studies on the biological basis of psychopathy report severe difficulties in aversive conditioning and instrumental learning in psychopaths (Blair, 2001; Patrick, 1994), and deficits to share the emotions ‘fear’ and ‘sadness’ in others (Blair, 2001). All these considerations may be subsumed under the low-fear model of psychopathy (Fowles & Dindo, 2006; Patrick, 2007).

To address the relationship between psychopathic and antisocial personality characteristics, about 30% of individuals suffering from antisocial personality disorder also meet the criteria of psychopathy (Hart & Hare, 1996). Recent results by Coid and Ullrich (2010) confirm the percentage of 30% comorbidity of antisocial personality disorder and psychopathy. Furthermore, the authors suggest a dimensional approach and postulate that antisocial personality disorder and psychopathy were disorders on a diagnostic continuum with symptom overlap.

3.4 The Relation of Antisociality, Psychopathy, and Feedback Processing

The present PhD-thesis focussed on sub-clinical manifestations of antisocial personality traits. Several studies investigated error and feedback processing in psychopathy and related constructs, such as antisocial personality traits, impulsivity, externalizing psychopathology and low socialization level with psychophysiological methods. For example, Munro and colleagues (2007a) studied activation differences between inmates of a forensic institution classified as psychopaths by the PCL-R (Hare, 2003) and healthy controls. The authors used a classical letter flanker paradigm (a common paradigm to evoke the ERN)[3] (neutral condition) as well as an emotional face flanker paradigm (affective condition). Participants had to distinguish the emotions anger and fear to judge whether the flanking stimuli were the same as the one in the middle or not. The authors found reduced ERN amplitudes only for the affectively-loaded flanker paradigm. The neutral stimuli led to no error-related activation differences between the two groups. Altogether, Munro and colleagues (2007b) hypothesized that effects of psychopathy might be more likely in cases where response monitoring involves either affectively-based stimuli (such as emotional faces), or affectively-charged situations (such as rewarding situations, punishment). These results are partly in line with Brazil and colleagues (2009) who found no group differences concerning the ERN amplitudes between inmates of a psychiatric institution scoring high on the PCL-R and matched controls with a comparable flanker task. However, Brazil and co-workers (2009) reported a reduction in the error positivity (Pe) amplitude as well as reduced signalling of error rates in the psychopathy group. The Pe is known to be a positive deflection succeeding the ERN (Gehring et al., 1993; Falkenstein et al., 1991). The Pe peaks between 200 to 400 ms after the onset of the incorrect response, thus, it is assumed to reflect later stages of error processing (Falkenstein et al., 1991). These later stages of error processing were linked to conscious error awareness (Falkenstein et al., 2000). According to Brazil and colleagues (2009) psychopaths showed intact early error processing as indexed by unchanged ERN amplitudes, but later stages of error processing might be impaired as indexed by Pe amplitude reduction. This would imply that psychopathic individuals have difficulties to effectively use internal error information to change and adapt their future behaviour. von

---

[3] The classical letter flanker task by Eriksen and Eriksen (1974) consists of five-letter strings where the two letters at each end were either congruent (SSSSS, HHHHH) to the central presented letter or incongruent (SSHSS, HHSHH). Participants' task is to indicate whether the central letter is S or H by two different mortor responses. Since fast motor responses are required, response errors are more common in incongruent trials.

Borries and colleagues (2009) investigated error-related learning deficits in individuals with psychopathy. The authors reported deficits in internal error processing, but no ERP amplitude modifications during external error processing. Chang and colleagues (2010) tried to relate error monitoring in healthy individuals to those with antisocial personality characteristics; they observed that depressive symptoms as well as antisocial characteristics were valuable predictors for ERN amplitudes in a flanker task. However, antisocial personality traits alone were the only predictor for Pe amplitude. In contrast to Brazil and co-workers (2009), Chang and colleagues (2010) chose a correlative approach to investigate error monitoring with psychopathy and antisociality. Furthermore, Brazil's subjects were inpatients of a psychiatric institution, whereas Chang's subjects were college students.

Regarding P300 amplitudes and antisociality and psychopathy, Gao and Raine (2009) reported mixed results in their review (c. f., Raine & Venalbes, 1988 vs. Bernat et al., 2007). Furthermore, they suggested that results in antisocials may not apply to results obtained in psychopaths. Nevertheless, reduced P300 amplitudes and longer P300 latencies were repeatedly associated with antisocial behavior, thus reflecting inefficient allocation of neuronal resources (Gao & Raine, 2009).

4. Research Questions – Aim of the Project

The aim of the present study was two-fold:

The first question focused on the processing of unexpected feedback outcomes, in particular on unexpected positive feedback outcomes. In line with the literature presented above, two ERP components are likely to occur with unexpected positive feedback. Firstly, an intermediate negative-going ERP deflection during the FRN time interval may occur as proposed by Hajcak and colleagues (2007) and, secondly, also a positive-going ERP deflection in the same time interval as reported by Holroyd and colleagues (2003, 2008) and Potts and colleagues (2006) .

However, according to the monitoring function of the ACC proposed in the RL-theory we suggested that unexpectedness as well as negative valence alone would induce a reward prediction error signal, which could appear as a negative-going ERP deflection on the scalp. Therefore, we expected in line with Hajcak and co-workers (2007) unexpected positive feedback to elicit a distinct negative ERP deflection during the FRN time range, which should nevertheless be smaller than the ERP component after unexpected negative feedback. Regarding the P300, we were interested in effects related to probability of occurrence and stimulus valence. In line with Johnson and Donchin (1980), we expected larger P300 amplitudes with unexpected compared to expected feedback stimuli, but no amplitude modulation caused by stimulus valence (Sato et al., 2004; Yeung & Sanfey, 2005). These research questions will be addressed in the first manuscript.

The second research question focused on the relationship of antisocial personality characteristics and feedback processing. We studied the impact of non-clinical antisocial personality characteristics on FRN and P300 amplitudes. Since antisocial personality traits were associated with the concept of psychopathy – which is among others characterized by deficits in emotional processing (Cleckley, 1976; Blair et al., 2004) – we expected individuals with higher values of antisociality to display smaller FRN amplitudes than individuals with lower values of antisociality. Furthermore, we expected smaller P300 amplitudes in antisocial subjects (Gao & Raine, 2009). These research questions will be addressed in the second manuscript.

5. Material and Methods

Two EEG studies were conducted at the Brain Research Laboratory of the Faculty of Psychology, University of Vienna. Twenty healthy volunteers participated in each study. The experimental gambling paradigms used are described in detail in each methods section of the two following manuscripts. The first EEG-study used numerical feedback stimuli that directly indicated real monetary gain or loss in each trial (Study Monetary Feedback). The data of this study are presented in *Article I*.

Feedback stimuli were different in the second study. Photographs of human posers depicting happy and angry facial expressions were presented to indicate positive (happy) or negative (angry) performance feedback in each trial (Study Facial Feedback). Although participants of the second study also received a financial bonus after completion of the experiment, the facial feedback stimuli indicated monetary reward only indirectly compared to the first study. Data of both studies are included in *Article II*.

Participants' financial remuneration was paid by a scholarship of the University of Vienna awarded to DMP in 2008. (Förderstipendium StudFG).

6. Article I

Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials

Daniela M. Pfabigan, Johanna Alexopoulos, Herbert Bauer, Uta Sailer

Faculty of Psychology, University of Vienna, Liebiggasse 5, A–1010 Vienna, Austria

Running head: FRN, P300, and Prediction Errors Signals

Descriptors: FRN, P300, reward prediction error signals

Abstract

The feedback-related negativity (FRN) has been hypothesized to be most sensitive to unexpected negative feedback. The present study investigated feedback expectancy and valence using a probabilistic gambling paradigm where subjects encountered expected or unexpected positive and negative feedback outcomes. In line with previous studies FRN amplitude reflected a negative reward prediction error, but to a minor extent also a positive reward prediction error. Moreover, the P300 amplitude was largest after unexpected feedback, irrespective of valence.

We propose to interpret the FRN in terms of a reinforcement learning signal which is detecting mismatch between internal and external representations indexed by the ACC to extract motivationally salient outcomes.

Introduction

Outcome evaluation has been of research interest for several years now. In event-related potential (ERP) studies, two ERP components have been found to be sensitive to different aspects of performance outcomes. One of the ERP components sensitive to outcome evaluation is the feedback-related negativity (FRN), which is a negative-going deflection in the ERP occurring after external negative feedback (Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997; Nieuwenhuis, Holroyd, Mol, & Coles, 2004). Peaking around 250 ms over frontal midline electrode-sites, the FRN is thought to reflect an early outcome evaluation, based on the binary classification of good vs. bad (Hajcak, Moser, Holroyd, & Simons, 2006), or of whether a goal has been achieved or not (Holroyd, Hajcak, & Larsen, 2006). The medial frontal negativity (MFN), an ERP that is elicited after monetary losses (Gehring & Willoughby, 2002; Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak et al., 2006), has comparable scalp distribution and latency. Nieuwenhuis, Yeung, Holroyd, Schurger, and Cohen (2004) showed that a negative ERP deflection could be elicited by either monetary or performance feedback when feedback contained information about both dimensions at the same time, depending on which dimension was emphasized during task presentation. Therefore, the terms FRN and MFN are interchangeable in the present study.

The FRN was originally interpreted by Miltner et al. (1997) in terms of operations of an error-processing system. In contrast to that, Botvinick, Braver, Barch, Carter, and Cohen (2001) as well as Yeung, Botvinick, and Cohen (2004) claimed to integrate the FRN into their concept of the anterior cingulate cortex (ACC) as a monitor for response conflict. Another theory accounting for the FRN is the reinforcement-learning theory (RL-theory) proposed by Holroyd and Coles (2002). The RL-theory states that events are evaluated by a monitoring system within the basal ganglia, making predictions whether or not events will turn out to be successful. If events are worse than expected, the inhibitory impact of dopaminergic neurons in the prefrontal cortex on the ACC is reduced by a phasic decrease of the dopaminergic level (Schultz, Dayan, & Montague, 1997). This so called temporal difference error (Holroyd & Coles, 2002) leads to a disinhibition of neurons in the ACC triggering the FRN. This negative prediction signal is used to optimize the acquisition of new action-outcome relations. Moreover, Nieuwenhuis, Holroyd, et al. (2004) reported that the amplitude of the FRN was depending on the relation between expected and actual outcome, and that it was most negative after unexpected negative feedback. Indeed, the cortical generator of the FRN has been identified in the ACC and adjacent cortical regions (Gehring & Willoughby, 2002; Holroyd et al., 2004; Miltner et al., 1997). Whereas Holroyd and Coles (2002) interpreted the FRN purely as reinforcement signal, Gehring and Willoughby (2002) stated that the FRN might reflect the motivational impact of ongoing events. Yeung, Holroyd, and Cohen (2005) extended the RL-theory by adding the notion that the FRN could reflect the reward signal alone without requiring an overt action. Subsequently, the ACC would use this reward signal to learn contingencies of the external environment.

However, predictions of the RL-theory regarding positive prediction signals are ambiguous. RL-theory would be consistent with the assumption that unexpected positive feedback (i) is associated with a small FRN (see Hajcak, Moser, Holroyd, & Simons, 2007), but also (ii) is not associated with an FRN, but with a positive ERP deflection in the time range of the FRN (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Potts, Martin, Burton, & Montague, 2006). The present study explicitly tested the hypothesis that the amplitude of the FRN is sensitive to the expectedness of the feedback as well as to feedback valence. More precisely, we predicted (i) that negative feedback would elicit an FRN, with larger FRN amplitudes for unexpected negative feedback than for expected or control negative feedback, and (ii) that unexpected positive feedback would also induce a distinct negative ERP deflection in the interval of the FRN.

The second component sensitive to performance outcomes is the P300, a component of the ERP peaking around 300 – 600 ms after stimulus presentation at posterior sites, which is primarily sensitive to stimulus significance and probability of occurrence (Duncan-Johnson & Donchin, 1977; Johnson & Donchin, 1980). Recent studies found modulated P300 amplitudes during decision and outcome evaluation tasks (Hajcak et al., 2005; Hajcak et al., 2007; Luu, Shane, Pratt, & Tucker, 2009; Sato et al., 2005; Yeung and Sanfey, 2004; Yeung et al., 2005), probably reflecting the evaluation of the functional significance of the feedback stimuli. In line with these P300 amplitude modulations, Nieuwenhuis, Aston-Jones, and Cohen (2005) stated that the P300 would also code the motivational significance of a reward along a good vs. bad dimension. However, the results are still controversial. Based on Yeung and Sanfey (2004) and Sato et al. (2005), we expected larger P300 amplitudes for less expected feedback, regardless of feedback valence. Furthermore, we explored the latencies of both ERP components.

Materials and Methods

Participants

Twenty right-handed subjects – ten women, ten men – participated in this study (mean age

26.6 +/- 3.27 years). Handedness was assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects were students of the University of Vienna, had normal or adequately corrected vision, were free of neurological diseases, and had no psychiatric history. The study was conducted in accordance with the *Declaration of Helsinki* (1973, revised in 1983) and local guidelines and regulations of the University of Vienna. Prior to participation, written informed consent was obtained from each participant. Subjects received an individual financial remuneration (between 10 and 25 Euros) depending on individual performance in the experimental task. Data from all 20 participants were subjected to statistical analysis.

Task

Prior to the experimental session, subjects had to complete a personality questionnaire (PSSI; Kuhl and Kazén, 1997) and a social attribution questionnaire (FKK; Krampen, 1991), the results of which will not be the subject of the current article.

The participants were comfortably seated about 70 cm in front of a 19-inch cathode ray tube monitor in a sound-attenuated room. Stimulus presentation and synchronization with the electroencephalogram (EEG) data collection was controlled by E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA). The experiment started with 48 training trials, where subjects (i) learned specific cue-response-contingencies necessary for the forthcoming experiment, and (ii) made themselves familiar with the experimental task. Trial numbers and reward probabilities per condition are presented in Table 1.

TABLE 1. Reward probabilities in training and experimental sessions, classification of conditions, and probability of occurrence

	Probability of positive feedback			
Cue-Response-Combination	Training vs. Experiment	Condition	Number of trials	Probability of occurrence
Cue 1+ Button 1	100% 75%	exp-pos	225/900	25%
		unexp-neg	75/900	8.3%
Cue 2 + Button 2	75% 75%	control-pos	225/900	25%
		control-neg	75/900	8.3%
Cue 3 + Button 1/2	0% 25%	unexp-pos	75/900	8.3%
		exp-neg	225/900	25%

The assignment of the three visual cues to the experimental conditions was counterbalanced across participants.

Each trial started with a black fixation cross on a gray screen, having a random duration between 2200 to 2700 ms. Subsequently, a black line drawing of a geometrical figure (a circle, a triangle, or a star, each presented 16 times; 10.5 x 10.5 cm in size; Bates et al., 2000) was presented for 500 ms as an imperative cue. This cue was followed by a black question mark for another 2000 ms. During its presentation subjects had to choose one of two buttons on a response pad (Psychology Software Tools, Inc.). During the training session, cue ‘one’ was associated with 100% reward probability for button one. Cue ‘two’ was associated with 75% reward probability for button two. Cue ‘three’ was not rewarded at all during training trials; no matter which button was selected. After a rewarded button choice, the German word for correct (RICHTIG) was presented in black letters on the screen. After a non-rewarded choice, the German word for incorrect (FALSCH) was presented. Feedback lasted for 1500 ms. In the experimental session (900 trials in total, six blocks) subjects were asked to search for more extended button press response patterns based on the previously learned cue-response mappings, e. g. pressing button one twice, and button two three times in five consecutive trials. This instruction was used to sustain subjects' expectations regarding the different reward probabilities for the three cues during the whole experiment, and to remind subjects of the training reward probabilities at the beginning of each new block. However, unknown to the subjects, no such button press response pattern existed. We assumed that subjects would maintain the original reward probabilities of the training session because they were not able to find any advanced response patterns. Although the instruction to search for meta-rules might have induced working memory and monitoring processes during the decision phase, this instruction was nevertheless necessary to make the occurrence of unexpected feedback plausible. During the experimental trials participants were presented with one of the three imperative cues each. The assignment of the three different cues to the conditions was counterbalanced across subjects. During the presentation of the question mark, subjects had to choose between button one or two by pressing the corresponding button. Subsequently, feedback was provided for 1500 ms after a short delay of 350 ms to minimize interference by movement-related potentials due to the button press. A correct answer was indicated by the number 15 in green color presented in the middle of the screen (2 x 1.5 cm in size), equivalent to a gain of 15 Euro-cents. After an incorrect answer, the number 15 was displayed in red color which indicated a loss of 15 Euro-cents. If no choice was made during the response epoch, subjects were informed that they had missed the response, and also lost 15 Euro-cents. After a block of 150 trials, overall performance feedback was provided.

After selecting the previously learned buttons for cue ‘one’ and ‘two’, subjects were provided with positive feedback in 75% of the trials. Feedback after cue ‘three’ was positive during 25% of the trials. Contrasting these new reward probabilities with those in the training trials, subjects encountered trials where a gain was highly expected (cue ‘one’), but a loss occurred, i.e., feedback was worse than expected. Likewise, subjects encountered trials where a loss was highly expected (cue ‘three’), but a gain occurred, i.e., feedback was better than expected. In contrast to the cues ‘one’ and ‘three’, cue ‘two’ was presented with unchanged probabilities for gains and losses (75% probability for gain, and 25% probability for loss) in the experimental session. Since subjective expectancy levels were not manipulated with cue ‘two’, it served as control condition. After three blocks subjects got paid the amount of money they had already gained in a five minutes break. This procedure was chosen to maintain the subjects' motivation. At the end of the experiment subjects were asked to estimate the reward frequencies of the three cues in a short questionnaire. Finally, they were rewarded with the money they had gained in the last three blocks. Including a seed capital of five Euros, subjects gained 19.21 +/- 4.09 Euros on average. At the end they were debriefed that no button press response pattern had existed. The whole experiment took about 70 minutes

Rationale of the task

In order to explicitly test the hypothesis that the amplitude of the FRN might be sensitive to the expectedness of feedback as well as to feedback valence, subjects had to be confronted with expected and unexpected positive as well as negative and control feedback (see Table 1). To manipulate subjective expectations participants were first asked to learn specific cue-response mappings to gain reward, which subsequently could be used to reinforce or violate established expectations.

Electroencephalographic recording

The EEG was recorded via 61 Ag/AgCl electrodes equidistantly embedded in an elastic electrode cap (EASYCAP GmbH; model M10, Herrsching, Germany). A balanced sterno-vertebral site, above the seventh vertebra and the right sterno clavicular joint, served as reference site for EEG recordings (Stephenson & Gibbs, 1951). Vertical and horizontal electrooculogram (EOG) was recorded with a bipolar setting to allow off-line eye movement correction. Electrodes were placed 1 cm above and below the right eye, and on the outer canthi. Electrode impedances were controlled by a skin scratching procedure at each electrode site prior to EEG recordings. A sterile single-use needle was used to slightly remove dead skin cells (Picton & Hillyard, 1972). Afterwards, degassed electrode gel (Electrode-Cap International, Inc., Eaton, OH) was filled into each electrode. All electrode impedances were kept below 2 kΩ, as checked with an impedance meter. All signals were recorded within a frequency range of 0.1 to 125 Hz and sampled at 250 Hz for digital storage.

Data analysis

EEG data

Prior to analysis, subject- and channel-specific weighting coefficients for vertical and horizontal eye movement artifacts were calculated as the ratio of the covariance between each EEG channel and the EOG, and the variance within the EOG channels. These parameters were obtained in two pre-experimental calibration trials where subjects performed guided vertical and horizontal eye movements (Bauer & Lauber, 1979). Subsequently, the weighted actual EOG signals were subtracted from the EEG in the experimental trials. Using a template matching procedure, blink coefficients were calculated and subtracted off-line from each EEG channel trial-by-trial (Lamm, Fischmeister, & Bauer, 2005, for a detailed description).

Off-line analysis was carried out using EEGLAB 6.03b (Delorme & Makeig, 2004), implemented in Matlab 7.5.0 (The MathWorks, Inc., Natick, MA). EEG data were low-pass filtered with a cut-off frequency at 30 Hz (roll-off 6 dB/octave) and epoched for each trial, starting 200 ms before feedback onset and lasting for 1200 ms. The 200 ms interval preceding stimulus onset served as baseline. Thereafter, extended infomax independent component analysis (ICA, Bell & Sejnowski, 1995; Lee, Girolami, & Sejnowski, 1999) was applied to the single-subject data to correct for residual eye movement-related activity, as outlined by Delorme, Sejnowski, and Makeig (2007). For each subject, individual components were screened for maps with a symmetric frontal topography, accounting for eye blinks and vertical eye movements. As suggested by Delorme et al. (2007), these components were discarded from further analysis by performing a back projection of the remaining components to the voltage time series. Subsequently, a semi-automatic artifact removal procedure was applied to the back transformed data. Artifact-afflicted trials that met the following criteria were labelled and finally rejected after visual inspection: voltage values exceeding +/-75 µV in any channel or a voltage drift of more than 75 µV. Due to the experimental setup, expected and control positive trials were presented more often than unexpected and control negative trials. To adjust the signal-to-noise-ratios, the number of trials of unexpected positive, unexpected negative, and control negative feedback was drawn randomly from each expected positive, expected negative, and control positive feedback condition for each subject to approximate the number of trials in all conditions. Thereafter, each condition contained 33.18 +/- 4.48 trials on average.

ERP data analysis

Artifact-free trials were averaged per subject and per condition, and grand averages of the six conditions were generated. Data were grouped into six conditions with the factors expectation and valence: (i) expected positive feedback (cue 'one'; EXP-POS), (ii) expected negative feedback (cue 'three'; EXP-NEG), (iii) unexpected positive feedback (cue 'three'; UNEXP-POS), (iv) unexpected negative feedback (cue 'one'; UNEXP-NEG), (v) control positive feedback (cue 'two'; CONTROL-POS), and (vi) control negative feedback (cue 'two'; CONTROL-NEG). Subsequently, the peak-to-peak voltage differences between the most negative peak 200–350 ms after feedback onset and the preceding positive peak at each of the electrode sites Fz, FCz, and Cz were calculated (Holroyd et al., 2003). This procedure was chosen to gain a more veridical account of neuronal activation related to feedback processing, as argued by Picton et al. (2000), since the FRN is superimposed on the slow-going P300 wave. If no FRN peak was apparent, the difference score was set to zero (which occurred in 3.33% of all cases, mostly after expected positive and control positive conditions). Furthermore, we added the factor experimental half with the levels 'first' and 'second' to the analysis, corresponding to the first 50% and the last 50% of the trials in the experiment. If subjects ignored the experimental instruction, diminished ERPs after unexpected conditions could be expected in the second half of the experiment (Rescorla & Wagner, 1972).[4] The peak-to-peak measures were subjected to a 2x3x3x2 repeated measures analysis of variance (ANOVA) with the factors HALF (first, second), LOCATION (electrode sites Fz, FCz, and Cz), EXPECTATION (expected, unexpected, control), and VALENCE (positive, negative).

For P300 analysis, peak-to-peak voltage differences between the most positive value at Pz in the time range of 300-500 ms and the preceding negative peak (i.e., N200) were calculated. These peak difference values were subjected to a 2x3x2 repeated measures ANOVA with the factors HALF (first, second), EXPECTATION (expected, unexpected, control), and VALENCE (positive, negative).

Peak latencies were measured from feedback onset to the corresponding peak amplitudes of the FRN (largest at FCz) and of the P300 (measured at Pz). The mean latencies of each subject were subjected to two separate 2x3x2 repeated measures ANOVAs with the factors HALF, EXPECTATION, and VALENCE. The level of significance was set at $p<.05$ for all tests. If necessary, degrees of freedom were adjusted using Greenhouse-Geisser correction. Significant interactions were explored with Tukey's HSD post-hoc tests. To demonstrate the effect size of the experimental manipulation, partial eta-squared (ηp^2) is reported, where 0.05 represents a small effect, 0.10 equals a medium effect, and 0.20 represents a large effect (i.e., describing at least 20% of the variance; Cohen, 1973).

---

[4] The model proposed by Rescorla & Wagner (1972) describes Pavlovian conditioning, i. e. the basis of the reinforcement learning account of the FRN. The model predicts that a difference between ERPs after expected and unexpected feedback should only be observable in the first experimental trials.

Results

Behavioral Results

During the pre-experimental training session, subjects responded in 92.2% of the cue ‘one’-trials with button one (100% reward probability), and in 79.4% of the cue ‘two’-trials with button two (75% reward probability). Subjects were never rewarded after the presentation of cue ‘three’, regardless of their response. Since this feedback provided no information for learning a cue-response mapping, they chose button one in 45% and button two in 49.1% of all cases. In the remaining 5.9%, subjects were too slow to make a choice. Pair-wise McNemar tests showed significant differences in button press preferences. Subjects chose button one significantly more often after cue ‘one’ than after cue ‘two’ ($\chi^2(1)=11.72$, $p<.001$), or cue ‘three’ ($\chi^2(1)=82.39$, $p< .001$). Moreover, subjects chose button two significantly more often after cue ‘two’ than after cue ‘one’($\chi^2(1)=11.72$, $p<.001$), or cue ‘three’ ($\chi^2(1)=29.85$, $p<.001$). Thus, subjects learned a strong association between cue ‘one’ and button one, as well as a weaker association between cue ‘two’ and button two, but they did not develop any button preference for cue ‘three’. In the post-experimental questionnaire, subjects estimated the reward frequencies of the three imperative cues. Ratings for positive feedback after cue ‘one’ ranged from 60-90% probability of occurrence (median=71.65). Likewise, ratings for positive feedback after cue ‘two’ were between 20-85% probability of occurrence (median=70). After the presentation of cue ‘three’, estimations for positive feedback lay in the range of 1–70% probability of occurrence (median=20). As can be seen the actual frequencies were underestimated in all three cases. A Wilcoxon signed-ranks test revealed that subjects expected positive feedback significantly less often after cue ‘three’ than cue ‘one’ ($Z=-3.92$, $p<.001$), or cue ‘two’ ($Z=-3.93$, $p<.001$). Thus, subjects were aware of the fact that positive feedback was presented more often after cues ‘one’ and ‘two’ than after cue ‘three’ which is the premise for the perception of the reward contingencies.

ERP data

ERPs elicited by the six feedback conditions are displayed in Figure 1 for electrode locations FCz at which the FRN was largest, and for Pz at which the P300 was measured. In Figure 2, the grand mean difference waves (negative minus positive feedback, merged over the whole experiment; see Picton et al., 2000) are plotted for expected, unexpected and control outcomes to visualize FRN scalp topography and waveforms. Figure 3 depicts mean FRN amplitudes at FCz with standard errors of the six conditions in a bar graph.

FIGURE 1. Grand average ERPs

Grand averages of the six conditions at electrode sites FCz (upper panel) and Pz (lower panel) for half 1 (left column) and half 2 (right column; n=20). Negative is drawn upwards per convention. Feedback presentation started at 0 ms, which is marked by a ticked vertical line.

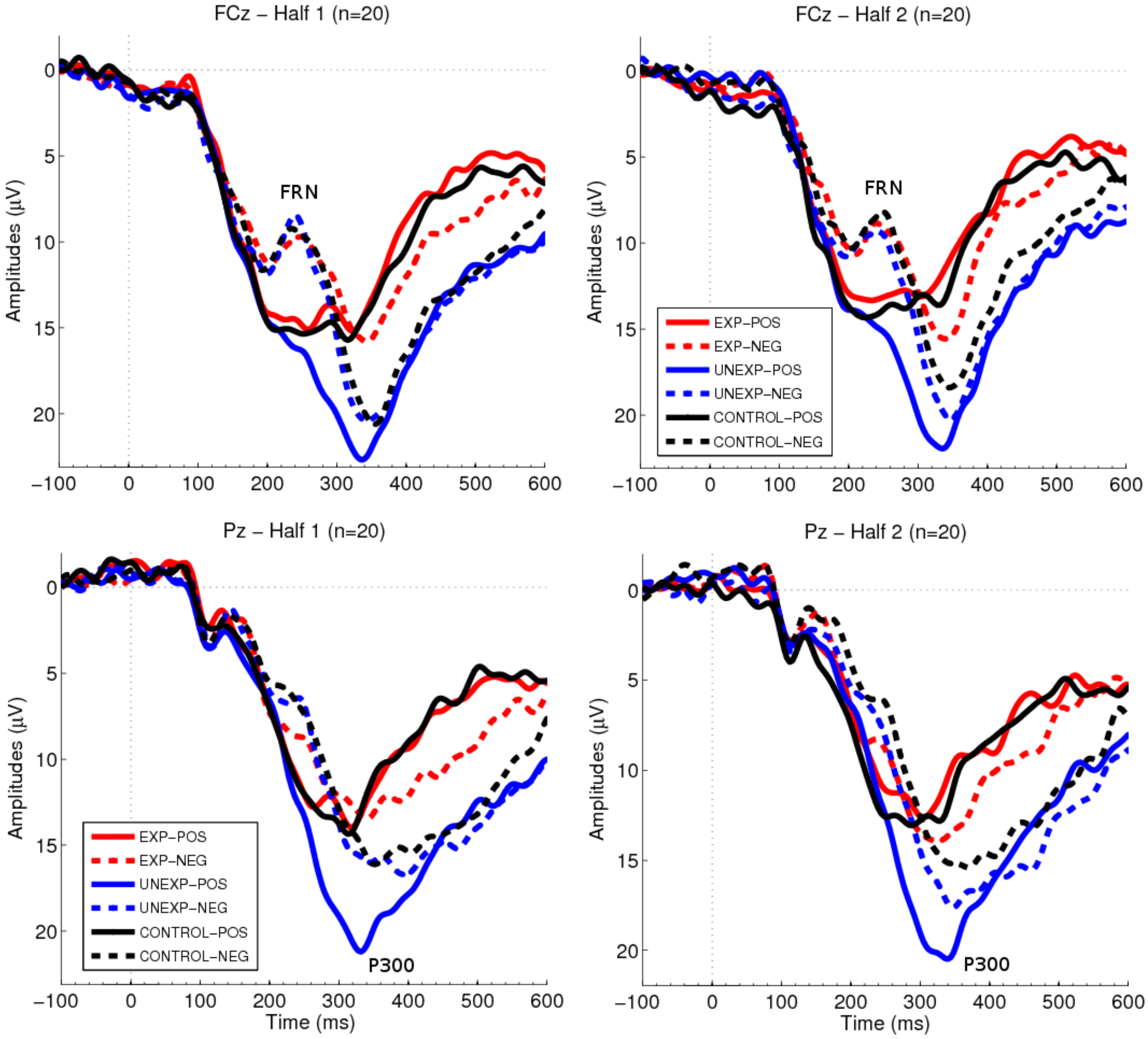
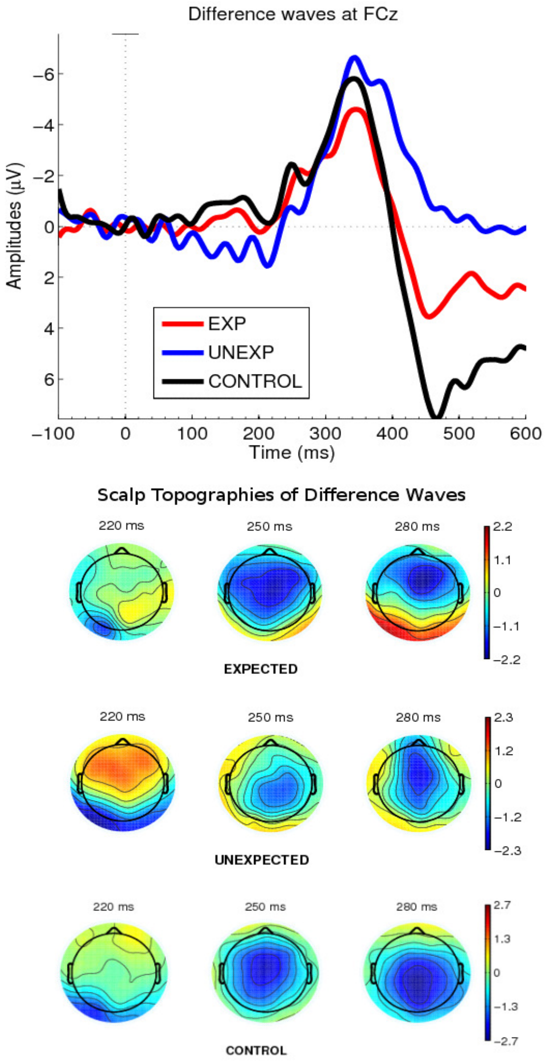
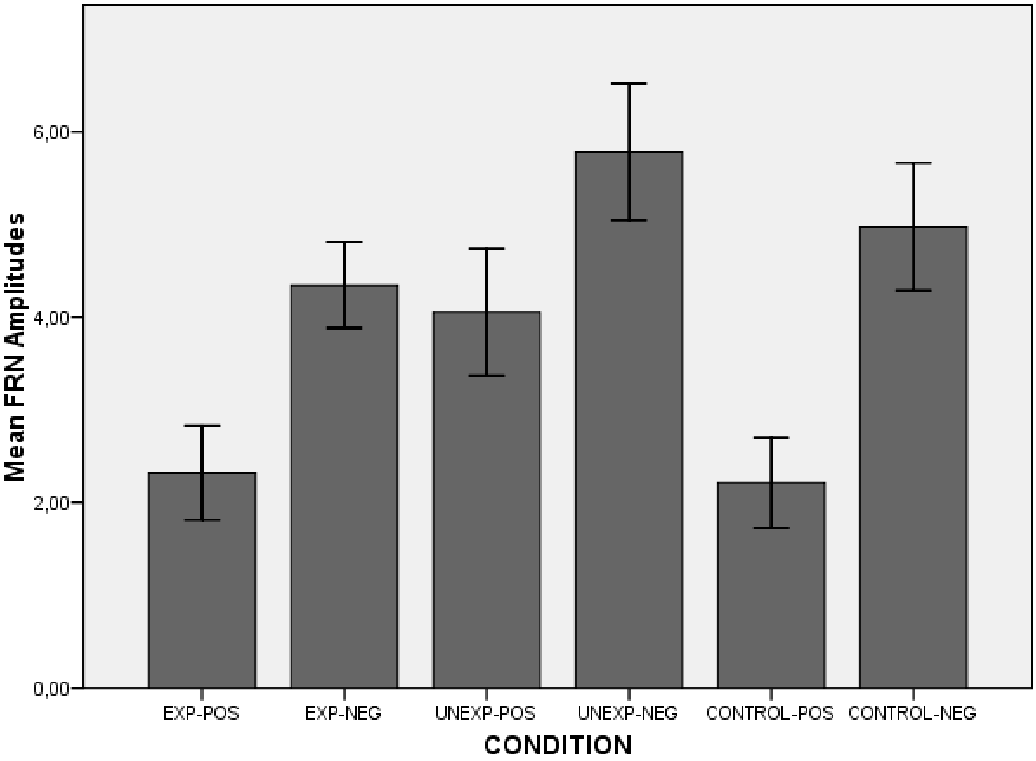


FIGURE 2. Difference wave amplitude courses and scalp topography



Upper panel: Amplitude courses of the voltage differences between negative and positive outcomes for expected, unexpected and control feedback (n=20) at FCz. Negative is drawn upwards. Lower panel: Scalp topographies of the above plotted difference waves. Timings are given relative to the onset of the feedback stimuli. Note that half 1 and 2 were merged together here.

FIGURE 3. Mean FRN amplitude values at Fcz



Error bars indicate one standard error.

FRN – Amplitude and latency

The ANOVA of the FRN amplitude showed significant main effects for LOCATION, F(1.2, 22.4)=5.63, p<.05, ηp^2=0.23, EXPECTATION, F(1.1, 21.7)=8.84, p<.01, ηp^2=0.31, and VALENCE, F(1, 19)=20.88, p<.001, ηp^2=0.52. Furthermore, a significant interaction was observed between LOCATION and VALENCE, F(1.3, 23.7)=6.78, p<.05, ηp^2=0.26). Post-hoc tests revealed that FRN amplitude was comparable for Fz and FCz (ns.), but was smaller at Cz than at FCz (p<.01). Moreover, FRN amplitudes were larger for negative compared to positive feedback at all electrode locations (all p's<.05) with less pronounced amplitude differences at Cz. FRN amplitude was also larger after unexpected than expected and control conditions (both p's<.05). Mean FRN peak-to-peak difference values and peak latencies for both halves and all conditions at FCz are displayed in Table 2. Experimental half had no effect on FRN amplitudes. We assumed therefore that subjects had obeyed the instruction to search for meta-rules and thus kept the training reward contingencies established during the training phase present throughout the entire experiment.[5] Regarding the FRN latency analysis at FCz, no significant effects or interactions of any of the factors emerged (all p's>.10).

---

[5] The use of the peak-to-peak-to-peak measure suggested by Yeung & Sanfey (2004) yielded comparable results. A main effect of EXPECTATION, F(1.2, 22.4)=25.14, p<.001, η_p^2=0.57, a main effect of VALENCE, F(1, 19)=34.52, p<.001, η_p^2=0.65, as well as interactions between LOCATION and EXPECTATION, F(2.4, 45.2)=3.42, p<.05, η_p^2=0.15, and LOCATION and VALENCE, F(1.5, 28.2)=4.04, p<.05, η_p^2=0.18, and an interaction between EXPECTATION and VALENCE, F(2, 37.3)=6.49, p<.01, η_p^2=0.25, were observed.

TABLE 2. Mean peak-to-peak amplitude values and mean latencies and corresponding standard deviation (SD) of the FRN at FCz, and the P300 at Pz (both n=20)

	FRN (FCz)				**P300 (Pz)**			
Condition Half 1	*Mean amplitudes*	*SD*	*Mean latency*	*SD*	*Mean amplitudes*	*SD*	*Mean latency*	*SD*
exp-pos	2.32	1.93	246	35.67	-12.22	4.71	329	76.49
exp-neg	4.37	2.25	263	37.53	-12.19	3.77	357	63.10
unexp-pos	4.01	3.36	261	39.95	-17.68	5.91	350	69.65
unexp-neg	5.88	3.47	256	22.29	-16.62	5.72	389	75.49
control-pos	1.81	1.68	265	49.37	-12.25	4.54	321	73.50
control-neg	4.99	3.28	262	21.62	-14.63	4.30	397	70.45

Condition Half 2	*Mean amplitudes*	*SD*	*Mean latency*	*SD*	*Mean amplitudes*	*SD*	*Mean latency*	*SD*
exp-pos	2.32	3.09	248	33.49	-9.93	4.37	336	80.72
exp-neg	4.32	3.16	251	18.38	-10.90	4.25	353	56.93
unexp-pos	4.10	3.48	258	31.75	-17.55	6.56	343	60.46
unexp-neg	5.69	3.74	262	36.3	-16.00	5.25	396	76.51
control-pos	2.62	3.19	259	35.29	-9.95	5.50	316	64.59
control-neg	4.97	4.33	262	30.83	-15.42	5.27	395	71.70

P300 – Amplitude and latency

The analysis of peak-to-peak amplitude differences at electrode site Pz revealed a main effect for EXPECTATION, F(2, 38)= 40.82, $p<.001$, $\eta p^2=0.68$, and a significant interaction between EXPECTATION and VALENCE, F(2, 38)=11.61, $p<.001$, $\eta p^2=0.38$. A post-hoc test indicated that unexpected feedback, no matter whether positive or negative, elicited larger P300 amplitudes than expected feedback (all p's<.05). Furthermore, both unexpected feedback conditions elicited larger P300 amplitudes than the positive control condition (both p's<.001). A valence effect on expectation level was only observable for the control condition with larger P300 amplitudes for negative control feedback compared to positive control feedback (p<.05) which can be explained by the lower probability of occurrence of negative control feedback than positive control feedback. Experimental half had no effect on P300 amplitudes.

For P300 latency significant main effects for EXPECTATION, F(2, 38)=4.60, $p<.05$, $\eta p^2=0.20$, and VALENCE, F(1, 19)=34.11, $p<.001$, $\eta p^2=0.64$ were observed, which were subsumed under a significant EXPECTATION x VALENCE interaction, F(2, 38)=7.59, $p<.01$, $\eta p^2=0.29$. The post-hoc test indicated that positive feedback yielded shorter P300 latencies after unexpected and control feedback (both p's<.05), but not after expected feedback.

Mean P300 peak-to-peak difference values and latencies are shown in Table 2.

Discussion

The present study investigated reward-related feedback processing using a probabilistic gambling paradigm in which participants encountered expected, unexpected, and neutral (i.e., control) positive and negative feedback. Subjects' expectations were built-up in a training session and then manipulated in the experimental session. In line with previous studies, negative valence was a good predictor for large FRN amplitudes. However, expectation level was another valuable predictor for FRN amplitude, i. e., the FRN amplitude was larger after unexpected as compared to control and expected feedback. These two factors - unexpectedness and negative valence - had comparable effects regarding the processing of decision outcomes. Thus, the FRN indicated mainly a negative reward prediction error, but to a lesser extent also a positive reward prediction error. The P300 was largest after unexpected and control negative feedback, thus coding unexpected events.

Our findings are in line with Wu and Zhou (2009) who suggested that the prediction error reflected by the FRN is not only defined in terms of valence, but also in terms of whether a pre-established expectancy is fulfilled or not, regardless of whether expectancy was induced trial-by-trial-wise or block-wise. Thus, the conclusion that the FRN represents the evaluation of the motivational impact of outcomes (Gehring & Willoughby, 2002) via factors such as valence and expectedness is highly plausible.

The FRN and prediction errors

The most pronounced FRN amplitudes of the data at hand were observed after unexpected negative outcomes, which were most unfavourable for the subjects. These results corroborate the assumption of the RL-theory of the FRN amplitude depending on the relation between actual and expected outcome (Nieuwenhuis, Holroyd et al., 2004; Wu & Zhou, 2009). Nevertheless, findings about the effect of expectation level on the FRN amplitude and the effect of positive feedback are not conclusive, as several studies reported divergent results. Shortly after Holroyd and Coles (2002) announced the RL-theory, the FRN was interpreted in terms of the absence of a negative deflection after positive feedback (Hajcak et al., 2005; Holroyd et al., 2003; Holroyd, Pakzad-Vaezi, & Krigolson, 2008). This interpretation changed recently, and by now research claims that a small negativity can be found after positive feedback (Hajcak et al., 2007; Moser & Simons, 2009). For example, inducing trial-by-trial expectations, Hajcak et al. (2007) did find small FRN amplitudes after unexpected reward, indicating a positive reward prediction error comparable to the present study. Similarly, Oliveira, McDonald, and Goodman (2007) postulated in their expectancy-deviation hypothesis that deviations from previously built-up expectations serve as strong agent eliciting FRN-like components. Their subjects encountered unexpected false-positive feedback in a time estimation task. However, subjects may have thought of a malfunction of the experimental set-up and interpreted the false-positive feedback as error feedback. Along these lines, a study by Ehlis, Herrmann, Bernhard, and Fallgatter (2005) showed that so called “PC-errors” (i.e., errors that were declared as computer-generated) elicited smaller error-related ERPs than “person-related errors” (i.e., errors that were generated by the subjects themselves), which could explain the negative deflection after false positive feedback in the Oliveira study (2007). To avoid the interpretation of unexpected positive feedback as PC-error, the present study used the aforementioned task instruction of the meta-rule search. Also, Wu and Zhou (2009) observed enhanced FRN amplitudes for violations of expectancy regarding the magnitude of the reward during their gambling task. The authors conclude that prediction errors are not solely defined in terms of valence, but also in terms of whether the actual outcome fits a pre-established expectancy or not.

In the present study all feedback stimuli were motivationally salient, indicating monetary gain or loss, and they had to be processed in the context of the previously built-up response contingencies. We propose that the FRN is indicating a mismatch between internal and external representations indexed by the ACC to extract positive and negative motivational salient outcomes. Unexpected feedback can also be described as a discrepancy (i.e., conflict) between internal and external feedback representations, which will require behavioural adaptations. Furthermore, the more predictable the feedback is, the less motivational value it contains for acquiring new behavioural strategies (e. g. less conflict is present). The low motivational value of expected feedback can explain the nearly absent FRN after expected positive and control positive feedback and the tendency of smaller FRNs after expected negative and control negative feedback, as compared to unexpected negative feedback, in the present data. However, another explanation for negative ERP deflections after unexpected positive feedback is offered by Holroyd et al. (2008). The authors claim that conflict-related processes are reflected in enhanced N200 amplitudes. This conflict-related negative deflection could be reduced after unexpected positive feedback because of simultaneous increase in mesencephalic dopamine activity due to unpredicted feedback since both processes are likely to be mediated by the ACC (Botvinick et al., 2001).

Extending the notion of Yeung et al. (2005) of the FRN indicating motivational salient events, we propose that such events might be discrepant or mismatching, and that they may comprise the dimensions of negative valence (Hajcak et al., 2006; Holroyd et al., 2003; Miltner et al., 1997), of unexpectedness (Bellebaum & Daum, 2008; Holroyd & Coles, 2002; Nieuwenhuis, Holroyd et al., 2004), or of differing reward magnitude (Goyer, Woldorff, & Huettel, 2008). In line with this argumentation, Goyer et al. (2008) observed larger FRNs after negative feedback in Gehring and Willoughby's (2002) gambling paradigm also for unchosen options. The authors hypothesized that the FRN was also influenced by contextual factors such as prior outcome history, and concluded that the motivational significance of the outcome is most important for error processing, since even early ERP signals are sensitive to the degree of an error, i. e. the representational mismatch it is causing.

Assuming that unexpected outcomes imply more motivational significance than expected outcomes via the mismatch, more pronounced outcome-related ERPs would be predicted for them. Indeed, this assumption is corroborated by observations where amplitudes of error signals were altered by manipulating contextual task properties leading to changes of motivational outcome aspects. For example, studies emphasizing speed over accuracy, thus devaluating the subjective significance of errors, have shown a decrease of error-related ERPs (Falkenstein, Hohnsbein, & Hoorman, 1995; Gehring, Goss, Coles, Meyer, & Donchin, 1993). Decreased error-related ERPs were also observed in studies where subjects scoring high on negative emotionality tests lost motivation in the course of the experiment (Luu, Collins, & Tucker, 2000). In line with these arguments, Holroyd and Coles (2008) reported that error-related amplitudes rather reflected the subjective value of a previous action like the original cue-response-contingencies of the present study, as opposed to simple good vs. bad-evaluation per se. Likewise, Moser and Simons (2009) suggested that the

FRN signals the integration of information about current and past actions and emotions.

The P300 and feedback processing

The P300 wave, a positive-going ERP component, is typically investigated in oddball-paradigms, where it is evoked when participants are attending to infrequent target stimuli. Its amplitude is maximal over parietal electrode sites (Snyder & Hillyard, 1976). The amplitude of the P300 increases with decreasing stimulus frequency, i.e., decreasing probability of the stimulus (Donchin & Coles, 1988). In the present study, the P300 was most prominent during less frequent feedback conditions (unexpected positive and negative, control negative), which were also subjectively less probable for participants. Therefore, subjective reward probability is the likely candidate to have induced larger P300 amplitudes after more unexpected feedback regardless of valence. In accordance with the context updating hypothesis (Donchin & Coles, 1988), unexpected events require updating of representations in working memory and therefore elicit larger P300s.

Although visual inspection might indicate that unexpected positive feedback induced the largest P300 amplitudes, no significant difference to unexpected negative feedback was observed. Therefore, the present results support the a priori hypothesis of the P300 being insensitive to feedback valence (Yeung & Sanfey, 2004). In contrast to Wu and Zhou (2009) feedback probability was coded by the FRN as well as the P300 in the present study. The authors explained their missing P300 probability effect via the possibility that they induced expectations about reward magnitude and not about reward valence. Hence, they concluded that the P300 might encode only the most significant feedback property when there are conflicting levels of relevance. Since only feedback expectancy and valence competed for attention in the present study, it might be the observed inconsistency between the presented cue and expectations after the button press that led to working memory updating processes. Whether or not the P300 is involved in coding the motivational significance of rewards as suggested by Nieuwenhius et al. (2005) remains a question of debate then.

The observed differences in P300 latencies can be explained in terms of stimulus evaluation processes. The P300 latency is thought to be modulated by stimulus classification demands; it is delayed if stimulus or distractor features are ambiguous (Kutas, McCarthy, & Donchin, 1977). In general, positive feedback, e. g. a gain of 15 Euro-cents, yielded shorter P300 latencies than negative feedback after unexpected and control feedback. This indicates that stimulus evaluation was easiest in these conditions. In line with parts of these results, Yeung et al. (2005) reported longer P300 latencies for losses; so it might be possible that the combination of lower probability and negative valence had a significant influence on the stimulus evaluation and delayed it via top-down processes.

Conclusions

To sum up, the present data indicate that feedback attributes such as expectancy and valence are coded by the FRN which is sensitive mostly to unexpected negative feedback, but also to unexpected positive feedback. For the data at hand, the FRN can be described in terms of a reinforcement learning signal indicating a mismatch between internal and external representations regardless of stimulus valence or expectedness. This mismatch is likely to be indexed by the ACC to extract positive and negative motivationally salient outcomes.

Acknowledgments

The participants' financial remuneration was funded by a scholarship of the University of Vienna awarded to D. M. P. (Förderstipendium StudFG). Preliminary results of parts of this study were presented at the 6th FENS Forum in Amsterdam (2008), and at the XXIXth ICP in Berlin (2008).

Address reprint requests to: Daniela M. Pfabigan, Department of Clinical, Biological, and Differential Psychology, Faculty of Psychology, University of Vienna, Liebiggasse 5, A-1010 Vienna, Austria. E-mail: daniela.pfabigan@univie.ac.at

References

Bates, E., Federmeier, K., Herron, D., Iver, G., Jacobsen, T., Pechmann, et al. (2000). Introducing the CRL International Picture-Naming Project (CRL-JPNP). *Center for Research in Language Newsletter*, *12*, 1-14.

Bauer, H., & Lauber, W. (1979). Operant conditioning of brain steady potential shifts in man. *Biofeedback and Self-Regulation*, *4*, 145-154.

Bell, A. J., & Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7, 1129-1159.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience*, *27*, 1823-1835.

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

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurements, 33*, 107-111.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9-21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*, 1443-1449.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Science*, *11*, 355-372.

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

Ehlis, A.-C., Herrmann, M. J., Bernhard, A., & Fallgatter, A.J. (2005). Monitoring of Internal and External Error Signals. *Journal of Psychophysiology*, *19*, 263-269.

Falkenstein, M., Hohnsbein, J., & Hoorman, J. (1995). Event-related potential correlates of errors in reaction tasks. In G. Karmos, M. Molnar, V. Csepe, I. Czigler, & J. E. Desmedt (Eds.), *Perspectives of event-related potentials research* (pp. 287-296), Amsterdam: Elsevier.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385-390.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *14*, 593-602.

Goyer, J. P., Woldorff, M. G., & Huettel, S. A. (2008). Rapid Electrophysiological Brain Responses are Influenced by Both Valence and Magnitude of Monetary Rewards. *Journal of Cognitive Neuroscience*, *20*, 2058-2069.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161-170.

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*, 148-154.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*, 905-912.

Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Reviews*, *109*, 679-709.

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*, 2481-2484.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L. E., Mars, R. B., Coles, M. G. H., & et al. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 497-498.

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*, 93-101.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*, 688-697.

Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behaviour. *Cortex*, *44*, 548-559.

Johnson, R. Jr., & Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, *17*, 167-178.

Krampen, G. (1991). *Fragebogen zu Kompetenz- und Kontrollüberzeugungen (FKK)*. Hogrefe Verlag, Göttingen.

Kuhl, J., & Kazén, M. (1997). *Persönlichkeits-Stil und Störungs-Inventar (PSSI)*. Hogrefe Verlag, Göttingen.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: P300 as a measure of stimulus evaluation time. *Science*, *197*, 792-795.

Lamm, C., Fischmeister, F., & Bauer, H. (2005). Individual differences in brain activity during visuo-spatial processing assessed by slow cortical potentials and low resolution electromagnetic tomography. *Cognitive Brain Research*, *25*, 900-912.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput*ations, *11*, 417-441.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General*, *129*, 43-60.

Luu, P., Shane, M., Pratt, N. L., & Tucker, D. M. (2009). Corticolimbic mechanisms in the control of trial and error learning. *Brain Res*earch, *1247*, 100-113.

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task: Evidence for a “Generic” Neural System for Error Detection. *Journal of Cognitive Neuroscience*, *9*, 788-798.

Moser, J. S., & Simons, R. F. (2009). The neural consequences of flip-flopping: the feedback-related negativity and salience of reward prediction. *Psychophysiology*, *46*, 313-320.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004a). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*, 441-448.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004b). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, *14*, 741-747.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus norepinephrine system. *Psychological Bulletin*, *131*, 510-532.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia*, *9*, 97-113.

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviations and the Representations of Action-Outcome Associations. *Journal of Cognitive Neuroscience*, *19*, 1994-2004.

Picton, T. W., & Hillyard, S. A. (1972). Cephalic skin potentials in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, *33*, 419-424.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. JR., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127-152.

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When Things are Better or Worse than Expected: The Medial Frontal Cortex and the Allocation of Processing Resources. *Journal of Cognitive Neuroscience*, *18*, 1112-1119.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black, W.F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99), New York: Appleton-Century-Crofts.

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & et al. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport*, *4*, 407-411.

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

Snyder, E., & Hillyard, S. A. (1976). Long-latency evoked potentials to irrelevant, deviant stimuli. *Behavioral Biology*, *16*, 319-331.

Stephenson, W., & Gibbs, F. A. (1951). A balanced non-cephalic reference electrode. *Electroencephalography and Clinical Neurophysiology*, *3*, 237-240.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114-122.

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience*, *24*, 6258-6264.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931-959.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*, 535-544.

Citation of this manuscript:

Pfabigan, D. M., Alexopoulos, J., Bauer, H. and Sailer, U. , Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. Psychophysiology, no. doi: 10.1111/j.1469-8986.2010.01136.x (access 15-10-2010)

Author contributions:

Concept and design: DMP. Execution: DMP, JA. Data analyses: DMP, US. Support of materials and analysis tools: DMP, JA. Manuscript: DMP, JA. HB, US.

7. Article II

Antisocial personality traits modulate event-related potentials that accompany feedback processing

Daniela M. Pfabigan, Johanna Alexopoulos, Herbert Bauer, Uta Sailer

Faculty of Psychology, University of Vienna, Liebiggasse 5, A–1010 Vienna, Austria

Running head: Antisocial personality traits and feedback processing

Abstract

This study investigated the relationship between feedback processing and antisocial personality traits. While event-related potentials were recorded, participants encountered expected and unexpected feedback during a gambling task. Feedback consisted either of emotional faces or numerical stimuli indicating monetary gain or loss.

When emotional faces served as feedback stimuli (experiment 1), only antisocial subjects showed enhanced P300 amplitudes after unexpected feedback. When numerical stimuli served as feedback stimuli (experiment 2), the feedback-related negativity (FRN) after losses tended to be enhanced in antisocial subjects. Furthermore, P300 latency was prolonged after expected feedback in antisocials. These results suggest that external feedback is salient to antisocials, and moreover, that emotional reactivity is intact or even enhanced in antisocial subjects. Apparently, antisocials seem to care about external feedback when unexpected emotional expressions or monetary reinforcers are involved.

Keywords: antisocial personality, feedback processing, FRN, P300

1. Introduction

Deficits in action monitoring are common phenomena in psychopathology. Prolonged response inhibition (Enticott, Ogloff, & Bradshaw, 2008) and reduced error-related amplitudes (e. g., Alain, McNeely, He, Christensen, & West, 2002) were observed in schizophrenia, as well as increased error-related activation in patients suffering from obsessive-compulsive disorder (OCD; Gehring, Himle, & Nisenson, 2000; Hajcak & Simmons, 2002). A related construct to psychopathology is psychopathy, which is a personality construct described by a variety of affective abnormalities, such as callousness, lack of empathy, lack of remorse, and antisocial personality traits (Cleckley, 1964; Hare, 1991). Comparable to other forms of psychopathology, psychopathy is associated with arousal-based deficits, e. g. the disrupted processing of emotional facial expressions, in particular fearful expressions (Blair et al., 2004; Montagne et al., 2005), or reduced physiological responsiveness to aversive conditioning stimuli (Birbaumer et al., 2005). These deficits have been associated with altered anterior cingulate cortex (ACC) activation. The ACC is a brain region central for the integration of cognitive, affective, and visceral information (Allman, Hakeem, Erwin, Nimchinsky, & Hof, 2001; Critchley, 2005; Thayer & Lane, 2000). This region is thought to be involved in action monitoring and the assessment of the motivational significance of external stimuli (Devinsky, Morrell, & Vogt, 1995), as well as the processing of affective information (Bush, Luu, & Posner, 2000).

An event-related potential (ERP) component related to performance monitoring and ACC function is the so-called ‘Feedback-Related Negativity’ (FRN, Miltner, Braun, & Coles, 1997; Nieuwenhuis, Holroyd, Mol, & Coles, 2004; Yeung, Botvinick, & Cohen, 2004a). The FRN is a negative going deflection over frontal electrode sites occuring 200 to 300 ms after external feedback on negative performance or monetary loss. It is thought to be generated in or near the ACC (Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Miltner et al., 1997). Holroyd and Coles (2002) postulated that the FRN is a reinforcement signal induced by the mesencephalic dopamine system which is conveyed to the ACC to optimize new action-outcome relations. Furthermore, these authors assumed that events worse than expected would elicit the largest amplitude deflections. In contrast, Gehring and Willoughby (2002) stated that the FRN might rather reflect the subjective negative evaluation of self-relevant information than the commission of an error per se. Following their hypothesis, it has been proposed to view the FRN as a reinforcement signal which is detecting mismatch between internal and external representations to emphasize motivationally salient outcomes (Yeung, Holroyd, & Cohen, 2005; Pfabigan, Alexopoulos, Sailer, & Bauer, 2009). Since psychopathology is known to affect motivation, the amplitude of the FRN is prone to also be affected. For example, reduced FRN amplitudes have been observed in depressive subjects (Foti & Hajcak, 2009). Similarly, a trend towards abnormal FRN amplitudes has been reported in OCD patients (Gründler, Cavanagh, Figueroa, Frank, & Allen, 2009; Nieuwenhuis, Nielen, Mol, Hajcak, & Veltmann, 2005). Investigating a personality construct related to psychopathology, Hirsh and Inzlicht (2008) explored the influence of Neuroticism on feedback processing. The authors assumed that neurotic subjects showed enhanced responses to uncertainty due to enhanced emotional responsiveness of the ACC (Bush et al., 2000), which actually was reflected in the FRN amplitude. To date, only one study related feedback processing and psychopathy (von Borries, Brazil, Bulten, Buitelaar, Verkes, & de Bruijn, 2009). These authors found impaired learning in a group of psychopaths during a probabilistic learning task, but FRN amplitudes comparable to a control sample.

The P300 has been associated with antisocial behavior. Peaking around 300 to 600 ms after stimulus onset at posterior recording sites, the P300 is reported to be sensitive to stimulus significance and probability of occurrence (Duncan-Johnson & Donchin, 1977; Johnson & Donchin, 1980). Increased P300 amplitude is thought to reflect enhanced stimulus processing capability by indexing the allocation of neural resources (Polich, 2003). P300 amplitude modulation was found in decision and outcome evaluation tasks, probably reflecting the functional or motivational significance of the feedback stimuli (Hajcak, Holroyd, Moser, & Simmons, 2005; Luu, Shane, Pratt, & Tucker, 2009; Yeung and Sanfey, 2004b; Yeung et al., 2005). Ambiguous results have been reported regarding the relationship between P300 amplitude and antisocial behavior since various P300 paradigms were used to assess emotional deficits in antisocial subjects (see Gao & Raine, 2009). Recently, Hicks et al. (2007) suggested that P300 reduction might be specifically associated with antisocial characteristics of psychopathy.

Since antisocial personality traits are a core symptom in psychopathy, the relationship between non-pathological antisocial personality traits and feedback processing was investigated in the current study by means of FRN and P300 in detail for the first time. Subjects were presented with a gambling task using external feedback. The feedback stimuli varied in their emphasis on emotional or motivational content across experiments. In the first experiment feedback was given via emotional face expressions indirectly indicating gain or loss, thus emphasizing the emotional content of the feedback. Emotional faces are valid social cues which act as social reinforcers (Rolls, 2000). In the second experiment, numerical feedback directly indicated monetary gain or loss, thus stressing more the motivational aspect of the feedback. Money is a well-known secondary reinforcer bearing high motivational value. In general, we expected negative and unexpected feedback to evoke the largest FRN amplitudes as suggested by Holroyd and Coles (2002) and Nieuwenhuis et al. (2004). In particular, we expected antisocial subjects to be less sensitive to these external feedback stimuli than social subjects because of the former groups' deficits in emotional reactivity (Cleckley, 1976; Blair et al., 2004), which should be reflected in smaller FRN amplitudes. With respect to the P300, larger amplitudes with unexpected than expected feedback were expected (Duncan-Johnson & Donchin, 1977). Moreover, smaller P300 amplitudes in antisocial than social subjects were expected since P300 reduction has been suggested to be a biological marker of vulnerability to externalizing disorders, e. g. problems in impulse control (Hicks et al., 2007; Krueger, 1999). Furthermore, P300 latencies were expected to be prolonged in antisocial subjects (Gao & Raine, 2009).

2. EXPERIMENT 1

In a gambling paradigm antisocial and social participants encountered expected and unexpected positive and negative feedback outcomes. The feedback stimuli consisted of social cues, i. e., emotional faces, indicating correct or incorrect responses. This way we attempted to test the hypothesis that antisocial subjects would show reduced FRN and P300 amplitudes than more social subjects when processing emotionally salient social feedback stimuli.

2.1 Methods

2.1.1 Participants and measures

Initially, twenty right-handed female students of the University of Vienna participated. The data of two participants had to be excluded from further analysis due to severe artifacts (continuous excessive alpha rhythm). The mean age of the remaining 18 subjects was 23.4 +/- 4.0 years. Handedness was assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Participants had normal or corrected-to-normal vision and were free of neurological or psychiatric diseases. The study was conducted in accordance with the *Declaration of Helsinki* and local guidelines of the University of Vienna. Informed written consent was obtained from each participant prior to participation. At the end of the experiment each participant received a bonus of 15 Euros for participation. Prior to EEG data collection, subjects completed a personality questionnaire (Persönlichkeits-Stil und Störungs-Inventar, PSSI; Kuhl & Kazén, 1997). The PSSI is a self-assessment tool covering the relative manifestation of 14 personality traits as the non-pathologic personality representations of personality disorders described in the DSM-

IV (American Psychiatric Association) and the ICD-10 (World Health Organization) diagnostic criteria. For this study, in particular the so-called AS (antisociality)-scale of the PSSI (self-determined personality and antisocial personality disorder) was in focus. Its reliability (Cronbach's $\alpha = 0.86$ - AS-scale) and validity are reported to be satisfactory. High T-values on the AS-scale, which consists of ten items (e.g., *"If people turn against me, I can wear them down."*), characterize people with self-determined and inconsiderate behaviour to achieve individual goals. They are described to act self-confidently, offending and humiliating while interacting with others, and furthermore to have problems adjusting to social and legal norms. Participants' average score on the AS-scale was T=50.33 (SD=11.11), ranging from 31 to 72. Participants were divided into three groups based on whether their T-values lay approximately below, above, or within two thirds of the sample's standard deviation. Six subjects formed the 'social' group (mean T=39.33, SD=4.55, range 31-42), six subjects the 'middle' group (mean T=49.00, SD=3.69, range 45-54), and the remaining six subjects constituted the 'antisocial' group (mean T=62.67, SD=7.45, range 56-72). Only the six social and the six antisocial subjects were considered for analysis. The T-values of these two groups differed significantly from each other (independent samples t-test: $t(10)=-6.55$, $p<.001$).

2.1.2. Task

Stimulus presentation and synchronous multi-channel electroencephalogram (EEG) recordings were controlled by a Pentium IV 3.00 GHz computer using E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, Pennsylvania). Subjects were comfortably seated 70 cm in front of a cathode ray tube monitor. The paradigm used was identical to that described in Pfabigan et al. (2009). An experimental session began with a training run of 48 trials where subjects learned specific cue-response-contingencies for the forthcoming experimental task. Each trial started with a black fixation cross on a grey screen, followed by an imperative cue consisting of a black line drawing of a simple figure (circle, triangle, or star, each presented 16 times; 10.5 x 10.5 cm in size; Bates et al., 2000). During the subsequent presentation of a black question mark, subjects had to choose one of two buttons on a response pad. Feedback was provided afterwards. The imperative cue remained on the screen for 500 ms; the question mark appeared immediately following the cue offset and remained on the screen until the participant responded or 2000 ms had elapsed. About 350 ms after the offset of the question mark the feedback stimulus appeared on the screen for 700 ms. During the inter-trial-interval, the fixation cross was presented again for a random duration of 2200 to 2700 ms. In the training session, one of the three imperative cues was associated with 100% reward probability for button one (cue ‘one’), another cue was associated with 75% reward probability for button two (cue ‘two’). Irrespective of button choice the third cue was not rewarded at all (cue ‘three’). The German word for correct (RICHTIG) was presented after reward-linked button choices and that for incorrect (FALSCH) with all other choices. The assignment of the three cues to the different reward probabilities was counterbalanced across subjects. After having learnt these simple cue-response-mappings the experimental session consisting of 900 trials started. Subjects were now asked to search for more complex button press response patterns on the basis of these simple cue-response-mappings (e.g., pressing button one thrice, and button two twice in five consecutive trials). This instruction was chosen to sustain participants’ expectations regarding the different reward probabilities for the three cues during the whole experiment. However, unknown to the participants, no such button press response pattern existed. Although this instruction to search for meta-rules might have induced monitoring and working memory processes during the decision phase, it was indispensable for making the occurrence of unexpected feedback plausible.

Feedback stimuli consisted of pictures of faces with emotional expressions taken from the standardized Ekman series (Ekman & Friesen, 1976; 4 x 5 cm in size). Two male and two female faces showing the emotions ‘happiness’ and ‘anger’ were used as positive (‘happy’ face) and negative (‘angry’ face) feedback stimuli; valence and gender were balanced across experimental trials. Participants were made familiar with the emotional faces during task instruction. Correct choices were indicated by the central presentation of a happy face, incorrect choices by an angry face. Subjects were informed that they could earn 10 to 15 Euros depending on their task performance, i.e., the number of correct responses. In contrast to the training session, subjects were now provided with positive feedback in 75% of the trials where they selected the previously learned buttons for cue ‘one’ and ‘two’. With cue ‘three’ subjects were provided with positive feedback in 25% of these trials. This contrast between the new reward contingencies and those of the training session ensured that participants encountered trials where a gain was highly expected (cue ‘one’), but a loss occurred, i.e., feedback was worse than expected. Likewise, subjects encountered trials where a loss was highly expected (cue ‘three’), but a gain occurred, i.e., feedback was better than expected. The data corresponding to cue ‘two’ were not further analysed since subjective expectation levels had not changed with this cue (75% probability for gain during the training and the experimental session). Nevertheless, cue ‘two’ was essential in this experimental paradigm - otherwise the occurrence of unexpected feedback stimuli would not have been plausible to the subjects (see Table 1).

	Probability of positive feedback				
Cue-Response-Combination	Training	Experiment	Condition	Number of trials	Probability of occurence
Cue 1+ Button 1	100%	75%	exp-pos	225/900	25%
			unexp-neg	75/900	8.3%
Cue 2 - Button 2	75%	75%	——		
Cue 3 + Button 1/2	0%	25%	unexp-pos	75/900	8.3%
			exp-neg	225/900	25%

The assignment of the three visual cues to the experimental conditions was counterbalanced across participants.

Table 1. Reward probabilities in training and experimental sessions, classification of conditions, and probability of occurrence in both studies.

After each of the six experimental blocks participants were given an overall performance feedback by means of the number of correct responses. Subsequently, they were instructed to search for a new button press response pattern during the next block. A five minute break took place after three blocks to allow subjects a short period of rest. After the six blocks subjects were asked to estimate in a brief questionnaire the obtained reward frequencies per cue. Afterwards the participants were told that they had performed extremely well – and regardless of their points accomplished - all were paid the full amount of money. Finally, they were debriefed about the external feedback manipulation. The whole experiment took about 70 minutes.

2.1.3. Data Acquisition and Preprocessing

The EEG was recorded via 61 Ag/AgCl ring electrodes, arranged equidistantly in an elastic electrode cap (EASYCAP GmbH, Herrsching, Germany; model M10). A balanced sterno-vertebral reference was used (Stephenson & Gibbs, 1951). Vertical and horizontal electrooculograms (EOG) were recorded bipolarily with electrodes placed 1 cm above and below the left eye and on the outer canthi, respectively to enable off-line eye movement artifact correction. During two pre-experimental calibration trials, subjects performed vertical and horizontal eye movements. These data were used to calculate subject- and channel-specific coefficients for eye movement correction (Bauer & Lauber, 1979). Skin scratching at each recording site (see Picton & Hillyard, 1972) and degassed conductance gel ensured electrode impedances below 2 kΩ. All signals were recorded within a frequency range of 0.016 to 125 Hz and sampled at 250 Hz for digital storage.

Off-line and prior to analysis the weighted EOG signals were subtracted from the EEG signals accordingly. Subsequently, blink coefficients were calculated using a template matching procedure and blink artifacts were also subtracted from the EEG signals (see Lamm et al., 2005). EEGLAB 6.03b (Delorme & Makeig, 2004) was used for further analysis. A low-pass filter (cut-off frequency 30 Hz, roll-off 6dB per octave) was applied to the EEG data. For ERP analysis signal epochs started 200 ms before feedback onset and lasted 900 ms; the mean of the first 200 ms serving as baseline. Before applying extended (infomax) independent component analysis (ICA, Bell & Sejnowski, 1995; Lee, Girolami, & Sejnowski, 1999) trials contaminated by muscular or movement artifacts were rejected based on visual inspection. ICA was performed to remove residual ocular artifacts, as described in Delorme, Sejnowski, & Makeig (2007), and afterwards a semi-automatic artifact removal procedure was done to eliminate epochs containing voltage values exceeding +/-75 µV in any channel.

Due to the experimental set-up the data sets per subject consisted of three times as many expected feedback trials than unexpected feedback trials. Therefore, numbers of trials per condition were equalized per subject in order to adjust for the signal-to-noise ratio of the ERPs. According to the total number of unexpected positive feedback trials, the same number out of all expected positive feedback trials was randomly drawn per subject. The same procedure was applied to the unexpected and expected negative feedback trials. As a result, each experimental condition contained on average 60.78 +/- 5.17 trials per person.

2.1.4. Data Analysis

Artifact-free epochs were averaged separately for each subject and each of the following four conditions: (1) expected positive feedback (exp-pos; cue ‘one’), (2) expected negative feedback (exp-neg; cue ‘three’), (3) unexpected positive feedback (unexp-pos; cue ‘three’), and (4) unexpected negative feedback (unexp-neg; cue ‘one’).To assess the FRN amplitudes at electrode sites Fz, FCz, and Cz voltage differences between the most negative voltage peak between 200 and 400 ms after feedback onset (FRN) and the average voltage value of its immediately preceding and following positive peaks were calculated (Yeung & Sanfey, 2004b). This procedure was chosen to achieve a more reliable account of neuronal activation in relation to feedback processing, as argued by Picton et al. (2000), because of the FRN being superimposed on the slow positive going P300. P300 amplitudes were obtained by searching for local maxima between 300 and 500 ms after feedback onset at electrode site Pz.

FRN amplitude differences were analyzed by means of a mixed 2x3x2x2 ANOVA with the between-subject factor *group* (social, antisocial), and the within-subject factors *location* (Fz, FCz, Cz), *expectation* (expected, unexpected), and *valence* (positive, negative). P300 latency was defined as the time elapsed between feedback onset and the P300 peak amplitude. The effect of factors *group*, *expectation*, and *valence* on the P300 peak amplitude and latency at Pz was analyzed by means of a mixed 2x2x2 ANOVA. Significant interactions were further analyzed by Tukey HSD post-hoc tests. If necessary, the degrees of freedom were adjusted using the Greenhouse-Geisser correction for repeated measures. To demonstrate the effect size of the ANOVA models, partial eta-squared (η_p^2) is reported. Small effects are represented by scores < 0.05, medium effects by scores around 0.10, and large effects by scores > 0.20 (Cohen, 1973).

2.2 Results

2.2.1. Behavioral Results

In the training session, subjects responded in 84% of cue ‘one’-trials with button one (100% reward probability), and in 78.8% of cue ‘two’-trials with button two (75%). With cue ‘three’-trials, subjects chose button one in 43.6%, and button two in 48.1% of all cases. In the remaining 8.4%, subjects were too slow to respond. Pair-wise McNemar tests showed significant differences in button press preferences. Button one was chosen significantly more often after cue ‘one’ than cue ‘three’ ($\chi^2_{(1)}$ =52.41, p<.001). In the post-experimental questionnaire, subjects estimated the probability of occurrence of positive feedback after cue ‘one’ with a median of 73% [60;90], after cue ‘two’ with a median of 70% [50;80], and after cue ‘three’ with a median of 33% [2;40]. A Wilcoxon signed-ranks test revealed that positive feedback was expected significantly more often after cue ‘one’ than cue ‘three’ (Z=-3.74, p<.001).

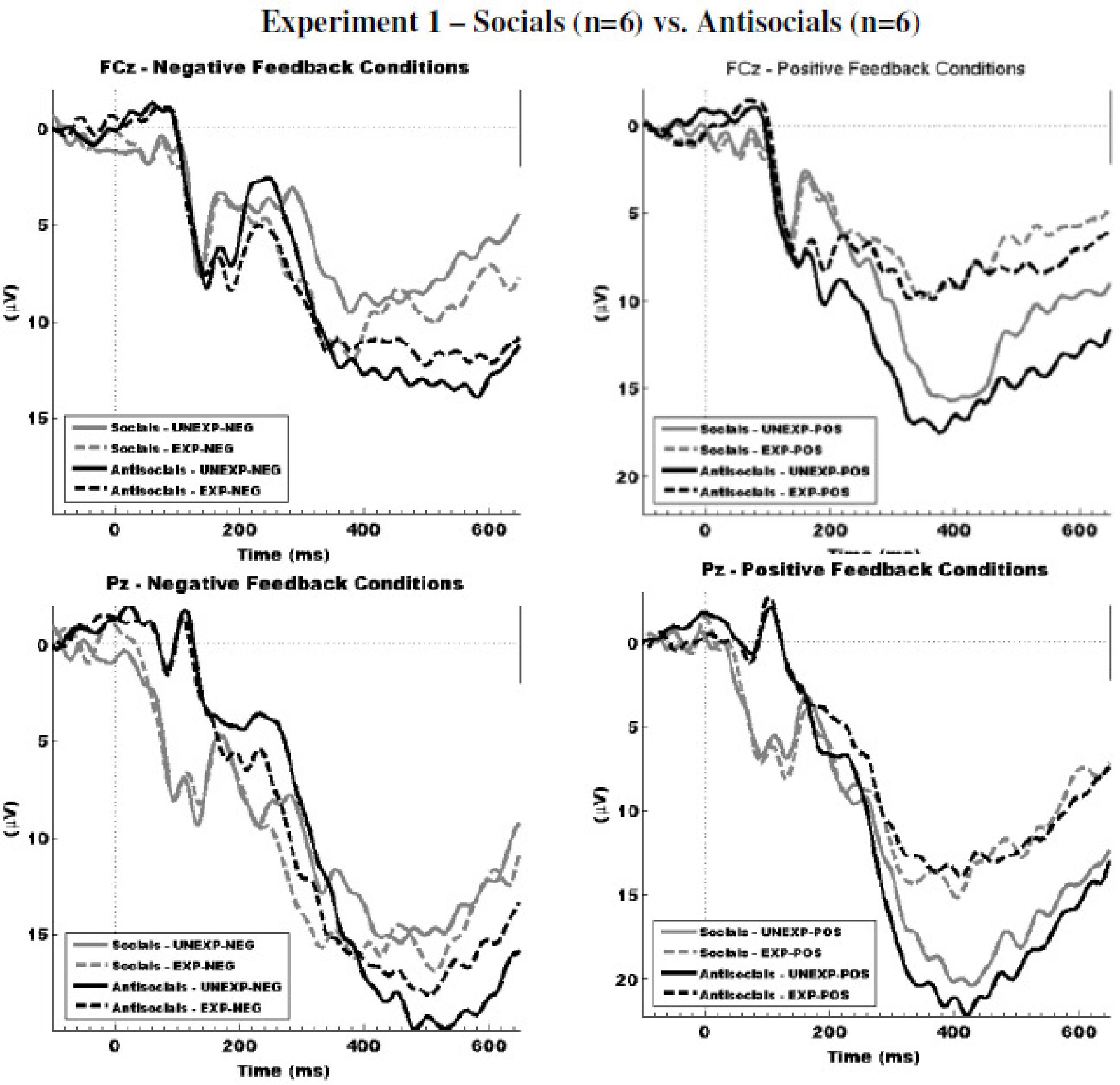
2.2.2. FRN

Main effects for *location* (F(2,12)=17.06, p<.0001, η_p^2=0.63), *expectation* (F(1,10)=14.06, p<.05, η_p^2=0.58), as well as a two-way interaction of *location* and *expectation* (F(2,13)=5.09, p<.05, η_p^2=0.34), and a three-way interaction of *group*, *location*, and *expectation* (F(2,13)=7.74, p<.05, η_p^2=0.44) were observed. For social subjects, larger FRN amplitudes for unexpected than expected feedback were apparent at electrode sites Fz (p<0.05) and FCz (p<0.05); whereas antisocial subjects showed larger FRN amplitudes after unexpected than expected feedback at all three electrodes (all p’s<.001). However, no group differences regarding valence or expectation emerged. The factor *valence* just missed significance (F(1,10)=4.79, p=.054, η_p^2=0.32), with larger FRN amplitudes for negative feedback. Figure 1 depicts amplitude courses of the two groups.

2.2.3. P300

For the factor *expectation*, a main effect (F(1,10)=25.25, p<.01, η_p^2=0.72), an interaction with *valence* (F(1,10)=9.67, p<.05, η_p^2=0.49), as well as an interaction with *group* (F(1,10)=6.69, p<.05, η_p^2=0.40) were observed. Unexpected positive (p<0.05) and unexpected negative (p<0.05) feedback elicited larger P300 amplitudes than expected positive feedback in all subjects. Additionally and regardless of valence, the P300 amplitudes of antisocial subjects were significantly larger for unexpected than for expected feedback (p<.05). In contrast, the P300 amplitudes of social subjects were not different for unexpected and expected feedback. No group effects regarding valence emerged. Also P300 latency did not show any significant differences.

Figure 1. Grand average ERPs of Experiment 1



Grand averages at electrode sites FCz (upper panel) and Pz (lower panel) for negative (right column) and positive (left column) feedback conditions differentiating social and antisocial subjects for Experiment 1. Negative is drawn upwards per convention, feedback presentation started at 0 ms.

2.3 Discussion

With emotional feedback stimuli, enhanced FRN amplitudes were observed after unexpected feedback in all subjects. Negative feedback conditions tended to elicit larger FRN amplitudes in all subjects. However, as opposed to our hypothesis, antisocial subjects did not show any FRN amplitude reduction in comparison to social subjects. In general, P300 amplitudes were larger after unexpected feedback in antisocial subjects but not in social subjects regardless of valence. Neither amplitude reduction in antisocials nor differences in processing speed were apparent for the present data.

In general, the larger FRN amplitudes after unexpected compared to expected feedback may indicate that the facial stimuli used were motivationally salient to all subjects in a comparable way. In this experiment, correct responses resulted in positive feedback stimuli. Subjects were told that monetary incentives were dependent on the number of correct responses accumulated over all trials. Therefore, the coupling of the feedback stimuli to monetary gain or loss was indirect. That might explain why negative feedback may have been less salient for subjects and did not lead to a significant FRN enhancement.

Interestingly, unexpected feedback conditions yielded larger P300 amplitudes only in the antisocial group, although unexpected stimuli have been reported to generally evoke larger P300 amplitudes due to their lower expectancy of occurrence (Duncan-Johnson & Donchin, 1977). Working memory processes might be a key to explain this result: according to Kok (2001) high memory load is accompanied by reduced P300 amplitudes which would imply that antisocial subjects processed unexpected feedback stimuli with less cognitive effort. However, FRN results do not support this notion because no FRN group differences emerged. Additionally, the feedback stimuli itself were the same for expected and unexpected conditions, thus ruling out the possibility of stimulus properties being related to working memory load.

Recent studies discussed whether or not P300 amplitude might reflect the evaluation of functional significance of feedback stimuli (Hajcak et al., 2005; Luu et al, 2009; Yeung & Sanfey, 2004b). If the P300 amplitude is indicating motivational significance, larger amplitudes would indicate larger stimulus salience. The P300 amplitude might emphasize the fact that antisocials cared more about unexpected stimuli because of reduced frustration tolerance. Indeed, a reduction in frustration tolerance was reported as a well-known symptom of psychopathic individuals (Hare, 1999).

To sum up, the assumption of emotional processing deficits in psychopathy indexed by FRN and P300 amplitudes was not corroborated by this experiment, although emotional stimuli were used. Since antisocial individuals are reported to show active and inconsiderate goal-directed behavior to obtain monetary incentives (Cornell et al., 1996), we hypothesized that the FRN and P300 amplitudes effects for social and antisocial subjects would disappear after enhancing the motivational feedback content while simultaneously reducing the emotional content. Hence, more emphasis was put on the motivational stimuli content using feedback stimuli that more directly indicated monetary reinforcement. This assumption was tested in experiment 2.

3. EXPERIMENT 2

In the second experiment, we used the same paradigm as in experiment 1, but with numbers instead of emotional faces as feedback stimuli. All changes from experiment 1 are reported.

3.1. Methods

3.1.1. Participants

Initially, twenty right-handed students of the University of Vienna (thereof ten women) participated in this study. The data of one male participant had to be excluded due to data acquisition problems. The mean age of the remaining 19 subjects was 26.3 +/- 3.1 years. The study was conducted in accordance with the *Declaration of Helsinki* and local guidelines of the University of Vienna. Informed written consent was obtained from each participant prior to participation. At the end of the experiment subjects received an individually adjusted bonus depending on their performance in the experimental task (between 10 and 25 Euros). Again, the PSSI questionnaire was administered before EEG data collection. Participants scored with an average T-value of 48.84 (SD=9.90) on the AS-scale, individual T-values ranged from 34 to 66. Based on the distribution of these individual T-scores, participants were separated into three groups (approximately below, above and within two thirds of the sample's standard deviation). Five subjects formed the 'social' group (mean T=36.40, SD=2.51, range 34-40; two females), seven subjects formed the 'middle' group (mean T=47.57, SD=4.86, range 42-54; four females), and seven subjects constituted the 'antisocial' group (mean T=59.00, SD=3.83, range 55-66; four females). There was no influence of sex on the individual scores on the AS-scale (independent samples t-test: t(17)=-1.00, p>.30). No differences of the individual AS-scale scores were observed when comparing both experiments either (independent samples t-test; t(35)=-0.43, p>.60). Only the five social and the seven antisocial subjects were considered for data analysis. The T-values of these two groups differed significantly (independent samples t-test: t(10)=-11.47, p<.001).

Parts of these data have been submitted for publication with focus on reward prediction error signals.

3.1.2. Task

In contrast to experiment 1, subjects were presented with numerical feedback stimuli. After completing 48 training trials, participants started with the first experimental block of 150 trials. As in experiment 1, they had to search for button press response patterns different to those in the training session during each block (see Table 1). A correct choice was indicated by the central presentation of the number 15 in green colour (2 x 1.5 cm in size), announcing a gain of 15 Eurocents. An incorrect choice was indicated by the number 15 in red colour, announcing a loss of 15 Eurocents. If subjects had missed the response interval they were informed about it and also lost 15 Eurocents; the respective trials were discarded from further analysis. After a block of 150 trials, subjects were provided with overall performance feedback about how much money they had won. Afterwards, they were instructed to search for a new button press response pattern in the next block. After three blocks, a five minute break took place, where subjects were paid with the amount of money they had already gained to maintain their motivation. The experiment ended after six blocks. Afterwards, subjects were asked to estimate the subjectively perceived reward frequencies of the three cues in a brief questionnaire. Finally, they were rewarded with the remaining money won in the last three blocks. Including a seed capital of five Euros, participants gained on average 19.69 +/- 3.57 Euros. Subjects were debriefed about the external feedback manipulation afterwards.

3.1.3. Data Acquisition and Preprocessing

The same data acquisition procedure was applied as described in experiment 1. To accommodate the different trial numbers, the same procedure was applied as in experiment 1. Finally, each condition contained 61.61 +/- 8.4 trials on average per person.

3.1.4. Data Analysis

Subject- and condition-wise averages were calculated for the four conditions (1) expected positive feedback (exp-pos; cue 'one'), (2) expected negative feedback (exp-neg; cue 'three'), (3) unexpected positive feedback (unexp-pos; cue 'three'), and (4) unexpected negative feedback (unexp-neg; cue 'one'). Subsequently, FRN and P300 peaks were extracted using the same criteria as in experiment 1. For FRN analysis, data were subjected to a mixed 2x3x2x2 ANOVA with the between-subject factor *group* (social, antisocial), and the within-subject factors *location* (Fz, FCz, Cz), *expectation* (expected, unexpected), and *valence* (positive, negative). For the P300 peak and latency analysis, only *group*, *expectation*, and *valence* were considered as factors.

3.2. Results

3.2.1. Behavioral Results

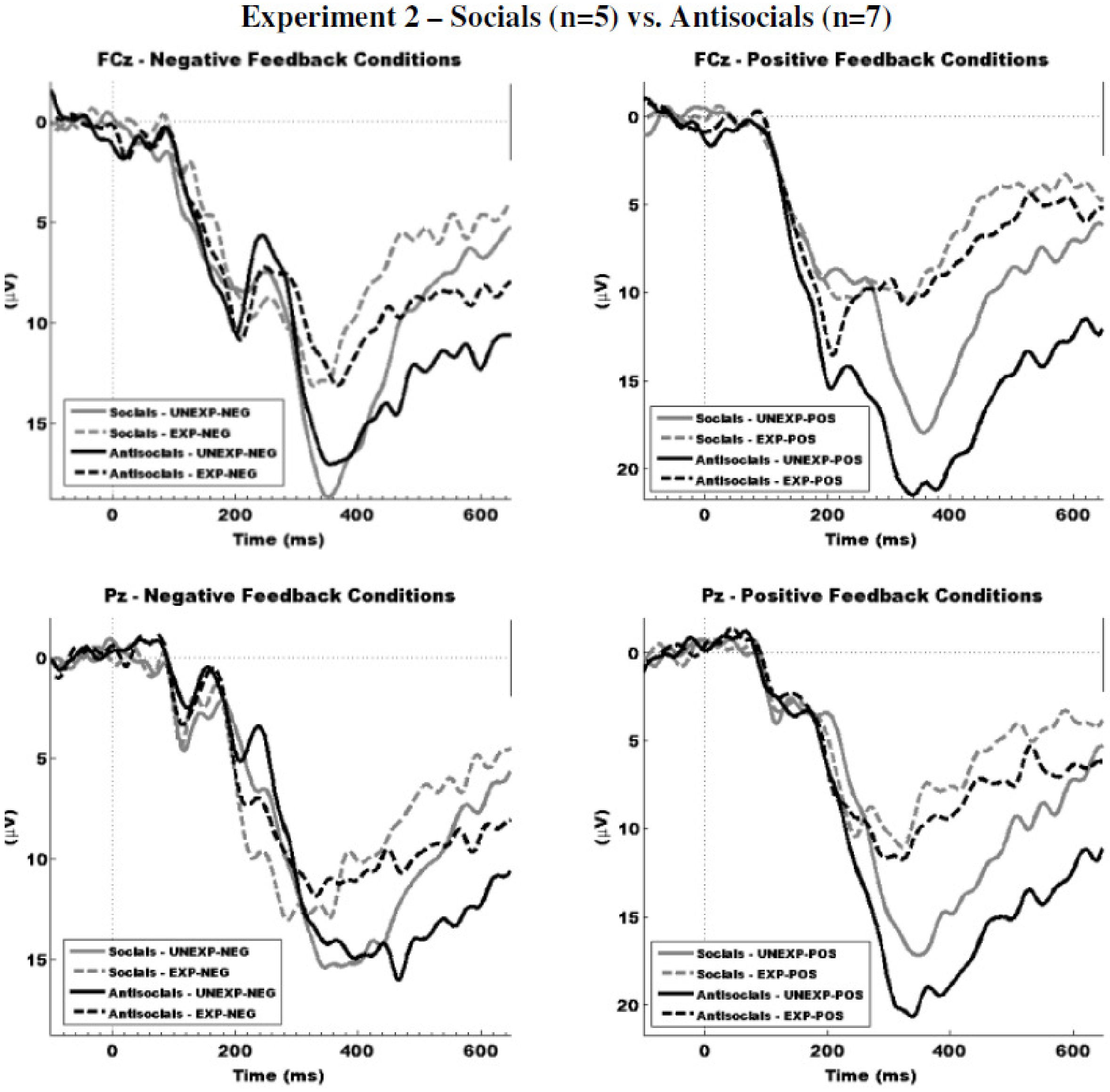
During the pre-experimental training session, subjects responded in 92.8% of cue 'one'-trials with button one (100% reward probability), and in 79.9% of cue 'two'-trials with button two (75%). In cue 'three'-trials subjects chose button one in 45.1% of all cases, and button two in 48.7%. For the remaining 6.3%, subjects were too slow to respond. Pair-wise McNemar tests showed significant differences in button press preferences. Button one was chosen significantly more often after cue 'one' than cue 'three' ($\chi^2_{(1)}$=81.74, p<.001). In the post-experimental questionnaire, subjects estimated the probability of occurrence of positive feedback after cue 'one' with a median of 70% [60;90], after cue 'two' with a median of 70% [20;85], and after cue 'three' with a median of 20% [1;70]. A Wilcoxon signed-ranks test revealed that positive feedback was expected significantly more often after cue 'one' than cue 'three' (Z =-3.83, p<.001).

3.2.2. FRN

Figure 2 displays feedback-locked average ERPs for expected and unexpected, positive and negative feedback conditions for the social and antisocial group. Main effects of *expectation* (F(1,10)=22.46, p<.01, η_p^2=0.69) and *valence* (F(1,10)=19.08, p<.01, η_p^2=0.66) indicated larger FRN amplitudes after unexpected and after negative feedback stimuli. Furthermore, the analysis revealed a two-way interaction of *valence* and *group* (F(1,10)=6.39, p<.05, η_p^2=0.39). Visual inspection indicated larger FRN amplitudes in antisocial than social subjects for negative feedback, although the post-hoc analysis

Figure 2. Grand average ERPs of Experiment 2

Grand averages at electrode sites FCz (upper panel) and Pz (lower panel) for negative (right column) and positive (left column) feedback conditions differentiating social and antisocial subjects for Experiment 2. Negative is drawn upwards per convention, feedback presentation started at 0 ms.



did not reach significance.

3.2.3. P300

A main effect of *expectation* (F(1,10)=27,71, p<.001, η_p^2=0.74) indicated larger P300 amplitudes for unexpected than for expected stimuli. The two-way interaction with *valence* (F(1,10)=4.59, p=.058, η_p^2=0.32) approached significance suggesting that unexpected positive feedback elicited the most positive P300 amplitudes. Antisocial and social subjects did not differ in their P300 amplitude (all p's>0.2). Regarding P300 latency, a main effect of *expectation* F(1,10)=13.15, p<.01, η_p^2=0.57) and an interaction of *group* and *expectation* F(1,10)=6.81, p<.05, η_p^2=0.41) emerged. Only social subjects tended to show a shorter latency for expected compared to unexpected feedback (p=.07).

3.3. Discussion

As expected, motivationally salient feedback stimuli directly indicating monetary gain or loss led to larger FRN amplitudes after both negative and unexpected stimuli for all subjects. However, not supporting our hypothesis, antisocial subjects did not show diminished FRN amplitudes. In contrast, antisocials showed a tendency for enhanced FRN amplitudes after monetary loss. Larger P300 amplitudes in all subjects after unexpected feedback supported the general P300 hypothesis; nevertheless, the proposed reduction in P300 amplitude for antisocials was not apparent.

Again, the current results contradict the assumption that deficits in emotional reactivity in antisocials are reflected by diminished feedback processing ERPs. The feedback stimuli of the present paradigm represented monetary gain or loss. Thus, we assumed that they were interpreted as incentives and were therefore motivationally salient to subjects.

In general, increased FRN amplitudes can be interpreted as error signals (Miltner et al., 1997) and response conflict signals (Botvinick, Braver, Barch, Carter, & Cohen, 2001), or as indicator for events worse than expected (Holroyd & Coles, 2002; Nieuwenhuis et al., 2004). This would imply that antisocial subjects experienced either more cognitive conflict after negative feedback, or that they experienced negative feedback in general as more unexpected than positive feedback. Further FRN interpretations emphasize subjective stimulus evaluation and motivational salience. Regarding the first, increased FRN amplitudes would indicate negative evaluation of self-relevant information (Gehring & Willoughby, 2002). Regarding the latter, increased FRN amplitudes would indicate substantial mismatch between internal and external stimulus representations. Since such a mismatch indicates the need for behavioral modification, it instantly renders the stimulus at hand motivationally salient (Yeung et al., 2005). Additionally, reduced frustration tolerance might have led to an increase in FRN amplitude via representing the motivational significance of an unexpected or negative event. In favor of this hypothesis, the present results might indicate that antisocial subjects perceived negative feedback as more salient. Therefore, they evaluated negative feedback more negatively than social subjects, and experienced a greater mismatch between the external feedback stimuli and their internal beliefs. Furthermore, these increased FRN amplitudes for antisocial subjects could imply that antisocial subjects cared more than social subjects about loosing money. In general, it may be more important to antisocial than social subjects to maximize their monetary gain. Corroborating this interpretation, Cornell et al. (1996) reported that antisocial individuals show inconsiderate goal-directed behavior to obtain monetary incentives as well as an increase in social status.

The missing group effect for the P300 peak amplitude results indicated that both groups allocated a comparable amount of processing capacity to the feedback stimuli. The P300 latency results point to its potential modification by antisocial personality traits. Whereas social subjects tended to show shorter P300 latencies and therefore, an accelerated processing speed for expected compared to unexpected stimuli, the P300 latencies of antisocial subjects did not differ for expected and unexpected feedback stimuli. This might indicate reduced speed of information processing and poorer attentional processing of expected target stimuli in antisocials (Gao & Raine, 2009). It has been suggested that the P3 only occurs after an event has been fully categorized or evaluated (Donchin & Coles 1988). According to this interpretation, the evaluation of expected compared to unexpected feedback stimuli took longer in antisocials than in social subjects. Therefore, it might be that antisocials need more time to categorize expected compared to unexpected stimuli. However, when the categorization process is finalized, the antisocials' representation of the stimulus is updated to the same extent as in social subjects, as indicated by comparable P300 amplitude.

To conclude, it may be possible that antisocials show high instead of low emotional reactivity to reinforcers. However, this may be specific to this particular kind of monetary incentives.

4. General Discussion

The main result of the present study is that no emotional processing deficit (Cleckely, 1976; Blair et al., 2004) was apparent during external feedback processing in antisocial subjects. On the contrary, antisocial subjects showed partly even enhanced feedback-related potentials. In experiment 2, the motivationally salient feedback stimuli differentiated antisocial from social subjects via the processing of negative stimuli. Antisocials showed a tendency towards enhanced FRN amplitudes after negative feedback. In contrast, valence did not differentiate between antisocial and social subjects in experiment 1, because no group or interaction effects emerged. One explanation might be the fact that feedback stimuli were only indirectly indicating monetary incentives in experiment 1. Thus, it is possible that facial feedback stimuli, i. e., social cues, were less salient to antisocial subjects than the numerical stimuli because the former were not directly associated with monetary reinforcers.

Furthermore, dissociation is apparent when comparing the present results to studies investigating internal error processing using error-related potentials in individuals with psychopathy and related constructs. Since the FRN can be described as the feedback-locked variant of the response-locked error-related negativity (ERN; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993), the present results should be comparable to these studies. The ERN is a negative ERP peaking around 50-100 ms after error commission (Falkenstein et al., 1991; Gehring et al., 1993). Hall, Bernat & Patrick (2007) reported reduced error-related amplitudes for subjects scoring high on the externalizing construct, which is thought to reflect an underlying vulnerability factor for conduct disorder, substance-use disorders, and antisocial behavior (Krueger, 1999; Krueger, McGue, & Iacono, 2001). The authors argued that lack of intrinsic motivation might be reflected in these reduced ERN amplitudes. In contrast, Munro et al. (2007a) and Brazil et al. (2009) reported no error processing differences between psychopaths and controls during reaction time tasks using neutral stimuli. Nevertheless, Munro et al. (2007a) found reduced error-related activation in psychopaths when using emotionally loaded stimuli in a reaction time task. Munro et al. (2007a) suggested therefore that psychopathy interacts with performance monitoring potentials (either response- or feedback-locked) more likely when affectively based stimuli (e. g., emotional faces) or affectively charged situations (e. g., reward or punishment) were involved. Munro's assumption is partly transferable to the present data. Differences between antisocial and social subjects in experiment 2 could be explained due to the reward/punishment scenario. However, Munro et al. (2007b) would have predicted reduced feedback-related amplitudes in antisocials during this affectively charged situation. To summarize, affectively charged situations can either lead to a reduction of error-related amplitudes during internal error processing in individuals with psychopathy as reported by Munro et al. (2007a), or to an increase of error-related amplitudes during external error processing in antisocial individuals as reported by the present study. This may suggest dissociation between internal and external error processing in individuals with psychopathy or antisocial personality traits. In line with this assumption, von Borries et al. (2009) reported deficits in internal, but not in external error processing in individuals with psychopathy. Nevertheless, this dissociation might be explainable by differing motivational systems. Internal error processing might be driven by intrinsic motivation. Thus, the amplitude of the error-related negativity decreases when intrinsic motivation decreases. In contrast, external error processing might be driven by extrinsic incentives (e. g., money). Thus, the amplitude of the FRN increases in the presence of valuable external incentives. Since monetary reinforcers affect antisocial or psychopathic individuals (Cornell et al., 1996), FRN amplitude enhancement after negative feedback in experiment 2 might be explainable.

The present P300 results were mixed and add to the ambiguous literature. Other studies found reduced P300 amplitudes in violent offenders (Bernat, Hall, Steffen, & Patrick, 2007), but also enhanced P300 amplitudes in psychopaths (Raine & Venables, 1988), or even no difference at all in P300 amplitudes between the two groups (Jutai, Hare, & Connolly, 1987). In particular, the data at hand disagree with the assumption of Hicks et al. (2007) of reduced P300 amplitudes in antisocials. The authors reported P300 reduction in relation to the concept of externalizing. Hicks et al. (2007) explained the P300 reduction based on diminished vigilance behavior during task presentation and decreased pre-stimulus preparation. Either the present data contradict Hicks et al.'s assumption of attention deficits in antisocials, or the non-pathological characteristic of antisocial personality traits can not be pulled together with the concept of externalizing. Disregarding the literature on P300 and psychopathy, the present data support the view that heightened P300 amplitudes indicate less expected events (Duncan-Johnson & Donchin, 1977), and incorporate high motivational significance (Yeung & Sanfey, 2004b) in all subjects.

One has to keep in mind that most studies regarding psychopathy investigated male inmate subject samples whereas the present study investigated a healthy student sample comprising at least to one half of women. Although no gender differences were apparent in participants' scorings on the AS-scale, future investigations should consider sex as factor, since men are at a significant higher risk than women to develop an antisocial personality disorder (Grant et al., 2004). Furthermore, inmate populations were often not differentiated in violent offenders with high aggression scores, or individuals with psychopathy, or antisocial personality traits. Antisocial personality traits were often found within psychopaths, but are not exclusively incorporated. Only about 30% of individuals with antisocial personality disorder meet the criteria of psychopathy (Hart & Hare, 1996).

5. Conclusion

To conclude, it might be possible that antisocial and social subjects vary in their processing of expectancy, since differences in the P300 component emerged in both experiments between the two groups. Nevertheless, the present data suggest that emotional reactivity is intact or even enhanced in antisocial subjects. Apparently, antisocials experience external feedback as motivationally salient when presented with unexpected emotional expressions or monetary reinforcers.

Author Notes

Participants' financial remuneration was funded by a scholarship of the University of Vienna to D. P. (Förderstipendium StudFG). Parts of this article were presented at the International Symposium on the Neural Basis of Decision Making, Groesbeek, Netherlands, April 2009, and at the 49th annual meeting of the Society for Psychophysiological Research, Berlin, Germany, October, 2009.

Author contributions:

Concept and design: DMP. Execution: DMP, JA. Data analyses: DMP, US. Support of materials and analysis tools: DMP, JA. Manuscript: DMP, JA. HB, US.

References

Alain, C., McNeely, H. E., He, Y., Christensen, B. K., & West, R. (2002). Neurophysiological evidence of error-monitoring deficits in patients with schizophrenia. *Cerebral Cortex*, *12(8)*, 840-846.

Allman, J. M., Hakeem, A. Y., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex: the evolution of an interface between emotion and cognition. *Annals of the New York Academy of Science*, *935*, 107-117.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders –DSM-IV*. Washington D. C.: American Psychiatric Press.

Bates, E., Federmeier, K., Herron, D., Iver, G., Jacobsen, T., Pechmann, et al. (2000). Introducing the CRL International Picture-Naming Project (CRL-JPNP). *Center for Research in Language Newsletter*, *12*, 1-14.

Bauer, H. & Lauber, W. (1979). Operant conditioning of brain steady potential shifts in man. *Biofeedback and Self-Regulation*, 4, 145-154.

Bell. A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7, 1129-1159.

Bernat, E.M., Hall, J.R., Steffen, B.V., & Patrick, C.J. (2007). Violent offending predicts P300 amplitude. *International Journal of Psychophysiology*, *66*, 161–167.

Birbaumer, N., Veit, R., Lotze, M. Erb, M., Herrmann, C., Grodd, W., et al. (2005). Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. *Archives of General Psychiatry*, *62*, 799-805.

Blair, R. J. R., Mitchell, D. G. V., Perschardt, K. S., Colledge, E., Leonard, R. A., Shine, J. H., et al. (2004). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences*, *37*, 1111-1122.

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

Brazil, I. A., de Bruijn, E. R. A., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J. D. M., Buitelaar, J. K. et al. (2009). Early and Late Components of Error Monitoring in Violent Offenders with Psychopathy. *Biological Psychiatry*, *65*, 137-143.Bush, G., Luu, P. & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215-222.

Cleckley, H. (1964). *The Mask of Sanity*. 4th ed. London, UK: The C. V. Mosby Company.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, *33*, 107-111.

Cornell, D.G., Warren, J., Hawk, G., Stafford, E., Oram, G., & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. *Journal of Consulting and Clinical Psychology*, *64*, 783-790.

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *The Journal of Comparative Neurology*, *493*, 154-166.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9-21.

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*, 1443-1449.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, *118*, 279-306.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? Behavioral Brain Science, 11, 355-425.

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

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Enticott, P. G., Ogloff, J. R. P., & Bradshaw, J. L. (2008). Response inhibition and impulsivity in schizophrenia. *Psychiatry Research*, *157*, 1-3.

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*, 447-455.

Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biological Psychology*, *81*, 1-8.

Gao, Y., & Raine, A. (2009). P3 event-related potential impairments in antisocial and psychopathic individuals: A meta-analysis. *Biological Psychology*, *82*, 199-210.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error-detection and compensation. *Psychological Science*, *4*, 385-390.

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, *11*, 1-6.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *14*, 385-602.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J., et al. (2004). Prevalence, Correlates, and Disability of Personality Disorders in the United States: Results From the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*, *65*, 948-958.

Gründler, T. O. J., Cavanagh, J. F., Figueroa, C. M., Frank, M. J., & Allen, J. J. B. (2009). Task-related dissociation in ERN amplitude as a function of obsessive-compulsive symptoms. *NeuroPsychologia*, *47*, 1978-1987.

Hajcak, G., & Simmons, R. F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. *Biological Psychiatry*, *110*, 63-72.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161-170.

Hall, J. R., Bernat, E. M., & Patrick, J. C. (2007). Externalizing Psychopathology and the Error- Related-Negativity. *Psychological Science*, *18,* 326-333.

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised.* Toronto, Ontario: Multi-Health Systems.

Hart, S. D., & Hare, R. D. (1996). Psychopathy and antisocial personality disorder. *Current Opinion in Psychiatry*, *9*, 129-132.

Hicks, B. M., Bernat, E., Malone, S. M., Iacono, W. G., Patrick, C. J., Krueger, R. F., et al. (2007). Genes mediate the association between P3 amplitude and externalizing disorders. *Psychophysiology*, *44*, 98–105.

Hirsh, J. B., & Inzlicht, M. (2008). The Devil You Know: Neuroticism Predicts Neural Response to Uncertainty. *Psychological Science*, *19*, 962-967.

Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Reviews*, *109*, 679-709.

Johnson, R. Jr., & Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, *17*, 167-178.

Jutai, J. W., Hare, R. D., & Connolly, J. F. (1987). Psychopathy and event-related brain potentials (ERPs) associated with attention to speech stimuli. *Personality and Individual Differences*, *8*, 175–184.

Kok, A. (2001). On the utility of P3 amplitudes as a measure of processing capacity. *Psychophysiology*, *38*, 557-577.

Krueger, RF. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, *56*, 921–926.

Krueger, R. F., McGue, M., & Iacono, W. G. (2001). The higher-order structure of common DSM mental disorders: Internalization, externalization, and their connections to personality. *Personality and Individual Differences*, *30*, 1245-1259.

Kuhl, J. & Kazen, M. (1997). *Persönlichkeits-Stil und Störungs-Inventar (PSSI).* Göttingen, Germany: Hogrefe Verlag.

Lamm, C., Fischmeister, F., & Bauer, H. (2005). Individual differences in brain activity during visuo-spatial processing assessed by slow cortical potentials and low resolution electromagnetic tomography. *Cognitive Brain Research*, 25, 900-912.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, *11*, 417-441.

Luu, P., Shane, M., Pratt, N. L., & Tucker, D. M. (2009). Corticolimbic mechanisms in the control of trial and error learning. *Brain Res*earch, *1247*, 100-113.

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task: Evidence for a „Generic“ Neural System for Error Detection. *Journal of Cognitive Neuroscience*, *9*, 788-798.

Montagne, B., van Honk, J., Kessels, R. P. C., Frigerio, E., Burt, M., van Zandvoort, M. J. E., et al. (2005). Reduced efficiency in recognising fear in subjects scoring high on psychopathic personality characteristics. *Personality and Individual Differences*, *38*, 5-11.

Munro, G. E. S., Dywan, J., Harris, G. T., McKee, S., Unsal, A., & Segalowitz, S. J. (2007a). ERN varies with degree of psychopathy in an emotion discrimination task. *Biological Psychology*, *76*, 31-42.

Munro, G. E. S., Dywan, J., Harris, G. T., McKee, S., Unsal, A., & Segalowitz, S. J. (2007b). Response inhibition in psychopathy: The frontal N2 and P3. *Neuroscience Letters*, *418*, 149-153.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience & Biobehavioral Reviews*, *28*, 441-448.

Nieuwenhuis, S., Nielen, M. M., Mol, N., Hajcak, G., & Veltman, D. J. (2005). Performance monitoring in obsessive-compulsive disorder. *Psychiatry Research*, *134*, 111-122.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh Inventory. *NeuroPsychologia*, 9, 97-113.

Pfabigan, D., Alexopoulos, J., Sailer, U., & Bauer, H. (2009). Antisocial personality traits and feedback processing – An ERP study. *Psychophysiology*, *46*, S78.

Picton, T. W., & Hillyard, S. A. (1972) Cephalic skin potentials in electroencephalography. *Electroencephalography and Clinical Neurophysiol*ogy, *33*, 419-424.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. JR., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127-152.

Polich, J. (2003). Overview of P3a and P3b. In J. Polich (Ed.), *Detection of Change: Event-related Potential and fMRI Findings* (pp. 83-98), Boston: Kluwer Academic Press.

Raine, A., & Venables, P.H. (1988). Enhanced P3 evoked potentials and longer P3 recovery times in psychopaths. *Psychophysiology, 25*, 30–38.

Rolls, E. T. (2000). The Orbitofrontal Cortex and Reward. *Cerebral Cortex*, *10*, 284-294.

Stephenson, W. & Gibbs, F. A. (1951) A balanced non-cephalic reference electrode. *Electroencephalography and Clinical Neurophysiol*ogy, *3*, 237-240.

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*, 201-216.

Von Borries, A. K. L., Brazil, I. A., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & de Bruijn, E. R. A. (2009). Neural correlates of error-related learning deficits in individuals with psychopathy. *Psychological Medicine*, doi: 10.1017/S0033291709992017.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders.* Geneva, Swiss: World Health Organization.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004a). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931-959.

Yeung, N., & Sanfey, A. G. (2004b). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience*, *24*, 6258-6264.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*, 535-544.

8. Conclusion and Outlook

Both studies aimed to provide further insight into the process of decision making and feedback utilization.

The first article provided further support for the notion of Gehring and Willoughby (2002) viewing the FRN component as an indicator of the evaluation of the motivational impact of feedback outcomes. Furthermore, we propose that the FRN is a neuronal signal, which incorporates several distinct feedback attributes such as feedback expectancy or feedback valence. Thus, the FRN could be interpreted as a reinforcement learning signal indicating a mismatch between internal and external feedback representations. Most likely, this mismatch is detected by the ACC to extract positive and negative motivationally salient outcomes. Future research should aim to develop a theoretical model behind this suggestion of the FRN indexing mismatch between internal and external representations. Botvinick and colleagues (2001) have provided a valuable starting point for this challenge with their conflict monitoring account of the FRN. However, a noteworthy suggestion might be to try to integrate the conflict monitoring account of the FRN (Botvinick et al., 2001) with the RL-theory (Holroyd & Coles, 2002) to achieve more valuable and reliable model predictions regarding the behaviour of the FRN amplitude.

The refinement of the experimental paradigm used would be another valuable contribution in future research. The topic of participants' instruction regarding the search for meta-rules has been raised by several anonymous reviewers. Although the FRN results were comparable during the two halves of the experiment, it would be interesting to omit this specific instruction regarding meta-rules and let participants find their own explanations for expected and unexpected feedback presentations. This research question is currently addressed in an on-going diploma thesis. Furthermore, reward probabilities of 50% (i.e., no predictions can be learned) were not included in the present project. These reward contingencies would be comparable to feedback presentation in Miltner's time estimation task (Miltner et al., 1997), since his participants were also not able to predict feedback valence.

The second article, on the other hand, has more heuristic value compared to the first manuscript. Although the general results supported the aforementioned suggestion of the FRN indicating representational mismatch, the results regarding personality characteristics are somewhat ambiguous. In opposition to the *a priori* hypothesis, no deficits in emotional processing indexed via feedback processing were apparent in antisocial subjects. Furthermore, no P300 amplitude reduction in antisocial was observed either.

Since the second manuscript is not published yet, it is certainly going to be revised in several ways. A major concern of this manuscript is the rather small sample size for each experiment. Although effect sizes are satisfactory, it might be possible that the observed effects of antisocial personality characteristics on FRN and P300 amplitudes are biased by the small sample size. Therefore, a future revision of this manuscript will include an increase in participants to strengthen the presented results. There is another aspect which is going to be taken care of in future manuscript revisions. The present manuscript focussed on the association between antisocial personality characteristics and the concept of psychopathy. The research question and hence deducted research hypotheses were embedded in the theoretical framework of psychopathy. However, participants of the present project were all healthy, young, and high-functioning college students – in comparison to older inmate or psychiatric participants of most of the cited studies. Furthermore, antisocial personality characteristics are reported to be associated with Factor 2 of psychopathy rather than with Factor 1 (Fowles & Dindo, 2006). In line with this theoretical dual-deficit account of psychopathy by Fowles and Dindo (2006) is the notion that primary psychopaths were unmotivated to modulate maladaptive behavior due to emotional processing deficits (Blair et al., 2005)- Secondary psychopaths were unable to modulate maladaptive behavior due to their impulsiveness and reduced capability of exerting cognitive control (Morgan & Lilienfeld, 2000). In regard to the cognitive deficits of individuals scoring high on Factor 2 of psychopathy, Wilkowski and colleagues (2008) observed an error adjustment deficit in secondary psychopaths. These subjects showed a reduction in their tendency to slow down their behavioral performance after an erroneous trial in different reaction time tasks. Thus, Wilkowski and co-workers (2008) suggested that secondary psychopaths were prone to errors to a greater extent than primary psychopaths. The authors related this deficit to deficient ACC functions. Since error processing and error monitoring are associated with ACC functions (Botvinick et al., 2001), deficits in feedback processing might nevertheless be observable in antisocial individuals. In light of these findings, the theoretical framework of the second manuscript is going to be revised and will incorporate this distinction between the two factor models of psychopathy.

To further the understanding of the association between antisocial personality characteristics and the two-factor model of psychopathy, future studies should apply psychological measures to assess both constructs.. Regarding stimulus conceptualisation, it would be interesting to further investigate the processing of motivationally salient and neutral stimuli indicating gain or loss. However, the physical appearance of feedback stimuli should be considered as a possible disruption factor and should be operationalized carefully (e. g., picture complexity, colour, luminance, emotional content, etc.).

The interpretation of the present P300 results should also be expanded. Larger P300 amplitudes for emotional feedback stimuli in antisocial subjects might indicate that antisocial participants had to recruit more cognitive resources to evaluate the unexpected and affective feedback stimuli compared to social ones. Thus, the task might have been more difficult for antisocial compared to social participants. This assumption could be related to behavioral data such as button choice behavior. Furthermore, antisocials might have been more motivated by the potential win-situation than social subjects. This would imply that unexpected feedback stimuli might have recruited more processing resources. Future investigations should take this assumption into account and collect additional data regarding the subjective rating of motivational impact of the feedback stimuli used.

Another interesting research question concerns the relationship between antisocial personality characteristics, psychopathy, feedback processing, and empathy. Hare's (2003) definition of psychopathy already includes the term 'lack of empathy'. Decety and Moriguchi (2007) also describe empathy deficits in individuals with antisocial personality disorder. Blair and colleagues (1995) suggested these by observing reduced ability to feel other people's emotional states in antisocials. Recently, Ali and colleagues (2009) reported differing deficits in empathy processing in primary and secondary psychopaths. Applying an emotion recognition task with happy, sad, and neutral stimuli, the authors observed that individuals scoring high on primary psychopathy were not processing sad facial expressions properly. Individuals scoring high on secondary psychopathy showed inappropriate responses to neutral facial expressions. Primary psychopaths were experiencing sad expressions as neutral ones; secondary psychopaths were experiencing neutral expressions as negative ones. These findings raise the question whether or not individuals scoring high in either of the two psychopathy factors experience negative and positive feedback in different ways.

9. Abstract

The processing of external feedback cues is crucial for decision making processes. Recent neuroscientific research has mainly focused on the processing of negative feedback events. The present project investigated neuronal processes related to positive feedback cues and personality characteristics. Two components of the event-related potential (ERP), the Feedback-Related Negativity (FRN) and the P300 were investigated in the context of positive and negative expected and unexpected feedback outcomes. Two EEG-studies were conducted applying a gambling paradigm where feedback stimuli consisted either of numbers indicating direct monetary gain and loss or of happy and angry faces indicating indirect monetary gain and loss.

Two research questions were addressed in separate scientific manuscripts. The first manuscript addressed the questions whether unexpected positive feedback elicits a negative ERP deflection in the latency range of the FRN component. Feedback expectancy and feedback valence were manipulated in the experimental paradigm. Results indicate that expectancy as well as valence had comparable impact on FRN amplitude modulation. FRN amplitudes were larger after unexpected compared to expected, and after negative compared to positive feedback. P300 amplitudes were modulated by expectancy – unexpected feedback conditions yielding largest P300 amplitudes – but not by valence. Thus, the proposal is made to interpret FRN amplitude modulation in terms of a reinforcement learning signal which is indicating motivationally salient outcomes.

The second manuscript addressed the question whether antisocial personality characteristics influence FRN amplitude modulation related to feedback expectancy and feedback valence. The effect of numerical versus emotional feedback stimuli was investigated in individuals scoring low and high on a psychological measure of antisociality. Results indicate that it is the dimension of feedback expectancy and not of valence that differentiates social from more antisocial individuals.

Future research on feedback processing should try to integrate the different theoretical frameworks and recent findings to promote the understanding of the underlying cognitive processes.

Zusammenfassung

Die Verarbeitung externer Feedbackreize ist essentiell für das Treffen von Entscheidungen. Die aktuelle neurowissenschaftliche Forschung bezüglich Entscheidungsfindung befasste sich bis dato hauptsächlich mit negativem Feedback und seinen Konsequenzen. Die vorliegende Dissertation beschäftigte sich mit neuronalen Prozessen in Zusammenhang mit positiven Feedbackreizen und Persönlichkeitseigenschaften. Zwei Komponenten des ereigniskorrelierten Potentials (EKP), die Feedback-Related Negativity (FRN) und die P300 Komponente wurden im Kontext von positiven und negativen, sowie erwarteten und unerwarteten Feedbackreizen untersucht. Es wurden zwei Elektroenzephalogramm (EEG) Studien durchgeführt, in denen eine Spielaufgabe als Experimentalparadigma vorgegeben wurde. Die Feedbackreize bestanden aus Zahlen, die direkt einen Geldgewinn oder -verlust andeuteten, sowie aus fröhlichen und ärgerlichen Gesichtern, die indirekt einen Geldgewinn oder -verlust anzeigten.

In getrennten Manuskripten wurde der Klärung zweier Forschungsfragen nachgegangen. Das erste Manuskript beschäftigte sich mit der Frage, ob unerwartetes positives Feedback eine vergleichbare negative Auslenkung des EKPs hervorruft wie es bei der FRN nach der Präsentation von negativem Feedback zu beobachten ist. Deshalb wurden die Feedbackdimensionen Erwartung und Valenz experimentell manipuliert. Die daraus resultierenden Ergebnisse deuten darauf hin, dass sowohl Erwartung als auch Valenz einen vergleichbaren Einfluss auf die Amplitudenmodulation der FRN haben. Eben diese FRN Amplituden waren erhöht nach unerwartetem und negativem Feedback. P300 Amplituden wurden hingegen nur durch die Erwartung des Reizes moduliert, nicht durch dessen Valenz. Daraus resultiert die wissenschaftliche Hypothese, die FRN Amplitudenmodulation als Signal des Verstärkungslernens zu betrachten, welches saliente Ereignisse anzeigt.

Das zweite Manuskript beschäftigt sich mit der Frage ob antisoziale Persönlichkeitseigenschaften die FRN Amplitudenmodulation in Bezug auf Feedbackerwartung und – valenz beeinflussen. Es wurde der Einfluss von numerischen im Gegensatz zu emotionalen Feedbackreizen in jenen Versuchspersonen untersucht, die entweder hohe oder niedrige Werte auf einer psychologischen Skala zur Erfassung von antisozialen Persönlichkeitseigenschaften aufwiesen. Die Ergebnisse deuten darauf hin, dass es die Dimension der Feedbackerwartung und nicht der -valenz ist, die zwischen sozialern und antisozialeren Individuen unterscheidet.

Zukünftige Forschung zu Feedbackverarbeitung sollte versuchen die theoretischen Ansätze mit aktuellen Forschungsergebnissen in Einklang zu bringen, um das Verständnis zugrunde liegender Prozesse von Entscheidungsfindung besser verstehen zu können.

10. References

Alain, C., McNeely, H. E., He, Y., Christensen, B. K., & West, R. (2002). Neurophysiological evidence of error-monitoring deficits in patients with schizophrenia. *Cerebral Cortex*, *12(8)*, 840-846.

Ali, F., Amorim I. S., & Chamorro-Premuzic, T. (2009). Empathy deficits and trait emotional intelligence in psychopathy and Machiavellianism. *Personality and Individual Differences*, *47*, 758-762.

Allman, J. M., Hakeem, A. Y., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The anterior cingulate cortex: the evolution of an interface between emotion and cognition. *Annals of theNew York Academy of Science*, *935*, 107-117.

Alpers, G. W., & Eisenbarth, H. (2008). *PPI-R. Psychopathic Personality Inventory-Revised* (German Version). Goettingen: Hogrefe.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders – DSM- IV.* Washington D.C.: American Psychiatric Press

Andreassi, J. L. (2007). *Psychophysiology. Human Behavior & Physiological Response*. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Ardekani, B. A., Choi, S. J., Hossein-Zadeh, G. A., Porjesz, B., Tanabe, J. L., Lim, K. O., et al. (2002). Functional magnetic resonance imaging of brain activity in the visual oddball task. *Cognitive Brain Research*, *14*, 347-356.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience, 27, 1823-1835.*

Bernat, E.M., Hall, J.R., Steffen, B.V., & Patrick, C.J. (2007). Violent offending predicts P300 amplitude. *International Journal of Psychophysiology, 66, 1*61–167.

Bertsekas, D. P. & Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming.* Athena Scientific: Belmont, Massachusetts.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*, *57*, 1-29.

Blair, R. J. R. (2001). Neuro-cognitive models of aggression, the Antisocial Personality Disorder and Psychopathy. *Journal of Neurology, Neurosurgery and Psychiatry*, *71*, 727-731.

Blair, R. J. R., Mitchell, D. G. V., Perschardt, K. S., Colledge, E., Leonard, R. A., Shine, J. H., et al. (2004). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences*, *37*, 1111-1122.

Blair, R. J. R., Mitchel, D. G. V., & Blair, K. S. (2005). *The Psychopath: Emotion and the Brain.* Blackwell: Oxford.

Blair, R. J. R. (2010). Neuroimaging of Psychopathy and Antisocial Behavior: A Targeted Review. *Current Psychiatry Reports*, *12*, 76-82.

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

Brazil, I. A., de Bruijn, E. R. A., Bulten, B. H., von Borries, A. K. L., van Lankveld, J. J. D. M., Buitelaar, J. K. et al. (2009). Early and Late Components of Error Monitoring in Violent Offenders with Psychopathy. *Biological Psychiatry*, *65*, 137-143.

Bush, G., Luu, P. & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215-222.

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*, 747-749.

Cleckley, H. (1941, 1976). *The Mask of Sanity.* St. Louis, MO: Mosby.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. NeuroImage, 35, 968-978.

Coid, J., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, *51*, 426-433.

Coles, M. G. H., Smid, H., Scheffers, M., Otten, L., 1995. *Mental chronometry and the study of human information processing*. In: Rugg., M., Coles, M. (Eds.), Electrophysiology of the Mind. Oxford University Press, New York, pp. 94-95.

Coles, M. G. H., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology*, 56, 173-189.

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *The Journal of Comparative Neurology*, *493*, 154-166.

Decety, J., & Moriguchi, Y. (2007). The empathic brain and its dysfunction in psychiatric populations: implications for intervention across different clinical conditions. *BioPsychoSocial Medicine*, *1*:22, doi: 10.1186/1751-0759-1-22.

de Olivaira-Souza, R., Hare, R. D., Bramati, I. E., Garrido, G. J., Ignacio, F. A., Tovar-Moll, F., & Moll, J. (2008). Psychopathy as a disorder of the moral brain: fronto-temporo-limbic grey matter reductions demonstrated by voxel-based morphometry. *Neuroimage*, *40*, 1202-1213.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, *118*, 279-306.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral Brain Science*, *11*, 355-425.

Donkers, F. C. L., Nieuwenhuis, S., & van Boxtel, G. J. M. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, *25*, 777-787.

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

Enticott, P. G., Ogloff, J. R. P., & Bradshaw, J. L. (2008). Response inhibition and impulsivity in schizophrenia. *Psychiatry Research*, *157*, 1-3.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143-149.

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*, 447-455.

Falkenstein, M., Hohnsbein, J., & Hoorman, J. (1995). Event-related potential correlates of errors in reaction tasks. In G. Karmos, M. Molnar, V. Csepe, I. Czigler, & J. E. Desmedt (Eds.), *Perspectives of event-related potentials research* (pp. 287-296), Amsterdam: Elsevier.

Fowles, D. C., & Dindo, L. (2009). Temperament and Psychopathy. A Dual-Pathway Model. *Current Opinions in Psychological Science, 18* (3), 179-183.

Frick, P. J., & Marsee, M. A. (2006). Psychopathy and developmental pathways to antisocial behaviour in youth. In C. J. Patrick (Ed.). *Handbook of Psychopathy*, pp. 353-374. Guilford: New York.

Gehring, W. J. (1992). *The error-related negativity: Evidence for a neural mechanism for error-related processing.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385-390.

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, *11*, 1- *Handbook of Psychopathy*6.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *14*, 593-602.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161-170.

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*, 151-160.

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*, 148-154.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*, 905-912.

Hare, R. D., & Jutai, J. W. (1983). Criminal history of the male psychopath: some preliminary data. In K. T. Van Dusen and S. A. Mednick (Eds.). *Prospective Studies of Crime and Delinquency*, pp. 225-236. Kluwer-Nijhoff: Boston.

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised* (2^{nd} edition). Toronto, ON: Mulit-Health Syst.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology*, 4, 217-246.

Hart, S. D., & Hare, R. D. (1996). Psychopathy and antisocial personality disorder. *Current Opinion in Psychiatry*, *9*, 129-132.

Herrmann, M. J., Roemmler, J., Ehlis, A. C., Heidrich, A., & Fallgatter, A. J. (2004). Source localization (LORETA) on the error-related-negativity (ERN/Ne) and positivity (Pe). C*ognitive Brain Research*, *20*, 294-299.

Holroyd, C. B., & Coles, M. G. H. (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Reviews*, *109*, 679-709.

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*, 2481-2484.

Holroyd, B.B., & Yeung, N. (2003). Alcohol and error processing. *TRENDS in Neuroscience*, *26*, 402-404.

Holroyd, C. B., Niewenhuis, S., Yeung, N., Nystrom, L. E., Mars, R. B., Coles, M. G. H., & Cohen, J. D. (2004). Dorsal anterior cingulated cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 497-498.

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: Electrophysiological responses to feedback stimuli. *Brain Research*, *1105*, 93-101.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*, 688-697.

Israel, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, *27*, 259-273.

Ito, T., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*, 887-900.

Johnson, R. Jr., & Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, *17*, 167-178.

Johnson, T. M., Otten, L. J., Boeck, K., & Coles, M. G. H. (1997). Am I too late? The neural consequences of missing a deadline. *Psychophysiology*, *34*, S48.

Kuhl, J. & Kazen, M. (1997). *Persönlichkeits-Stil und Störungs-Inventar (PSSI).* Göttingen, Germany: Hogrefe Verlag.

Kuhl, J. (2001). *Motivation und Persönlichkeit. Interaktionen psychischer Systeme*. Göttingen: Hogrefe.

Kutas, M., McCarthy, G., Donchin, E., 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792-795.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self report measure of psychopathic personality traits in noncriminal populations. J*ournal of Personality Assessment*, *66*, 488-524.

Lorenz, A. R., & Newman, J. P. (2002). Deficient response modulation and emotion processing in low-anxious Caucasian psychopathic offenders: results from al lexical decision task. *Emotion*, *2*, 91-104.

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science*, *14*, 47-53.

Luu, P., Shane, M., Pratt, N. L., & Tucker, D. M. (2009). Corticolimbic mechanisms in the control of trial and error learning. *Brain Res*earch, *1247*, 100-113.

Lykken, D. T. (1995). *The antisocial personalities.* Mahaw, N. J.: L. Erlbaum, U.S.

Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*, *116*, 155-165.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or Fiction. *Psychological Bulletin*, *82*, 213-225.

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task *Handbook of Psychopathy*: Evidence for a “Generic” Neural System for Error Detection. *Journal of Cognitive Neuroscience*, *9*, 788-798.

Montagne, B., van Honk, J., Kessels, R. P. C., Frigerio, E., Burt, M., van Zandvoort, M. J. E., et al. (2005). Reduced efficiency in recognising fear in subjects scoring high on psychopathic personality characteristics. *Personality and Individual Differences*, *38*, 5-11.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*, 760-767.

Morgan, A. B., & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*, 113-136.

Munro, G. E., Dywan, J., Harris, G. T., McKee, S., Unsal, A., & Segalowitz, S. J. (2007a). ERN varies with degree of psychopathy in an emotion discrimination task. *Biological Psychology, 76,* 31-42.

Munro, G. E. S., Dywan, J., Harris, G. T., McKee, S., Unsal, A., & Segalowitz, S. J. (2007b). Response inhibition in psychopathy: The frontal N2 and P3. *Neuroscience Letters*, *418*, 149-153.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004) . Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, *14*, 741-747.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus norepinephrine system. *Psychological Bulletin*, *131*, 510-532.

Nieuwenhuis, S., Slagter, H. A., Alting von Geusau, N. J., Heslenfeld, D. J., & Holroyd, C. B. (2005). Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, *22*, 3161-3168.

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviations and the Representations of Action-Outcome Associations. *Journal of Cognitive Neuroscience*, *19*, 1994-2004.

Patrick, C. J. (1994). Emotion and psychopathy: startling new insights. *Psychophysiology*, *31*, 319-330.

Patrick, C. J. (2007). Antisocial personality disorder and psychopathy. In W. O'Donohue, K. A. Fowler, & S. O. Lilienfeld (Eds.), *Handbook of personality disorders* (pp. 109-166). New York: Sage.

Polich, J. & Kok, A., 1995. Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, 41, 103-46.

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When Things are Better or Worse than Expected: The Medial Frontal Cortex and the Allocation of Processing Resources. *Journal of Cognitive Neuroscience*, *18*, 1112-1119.

Raine, A., & Venables, P.H. (1988). Enhanced P3 evoked potentials and longer P3 recovery times in psychopaths. *Psychophysiology, 25*, 30–38.

Raine, A. (1989). Evoked potentials and psychopathy. *International Journal of Psychophysiology*, *8,* 1-16.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: A.H. Black, W.F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99), New York: Appleton-Century-Crofts.

Ridderinkhof, K. R., Nieuwenhuis, S., Hajcak, G., van den Wildenberg, W., & Burle, B. (2004). Suboptimal action monitoring in mediofrontal cortex results in performance declines. Poster presented at the 11th annual meeting of the Cognitive Neuroscience Society, San Francisco.

Rodrigo, C., Pajapakse, S., & Jayananda, G. (2010). The 'antisocial' person: an insight in to biology, classification and current evidence on treatment. *Annals of General Psychiatry*, *9,* 1-31.

Sadeh, N., & Verona, E. (2008). Psychopathic Personality Traits Associated with Abnormal Selective Attention and Impaired Cognitive Control. *Neuropsychology*, *22*, 669-680.

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & et al. (2005). Effects of value and reward magnitude on feedback negativity and P300. *NeuroReport*, *4*, 407-411.

Scheffers, M. K., Coles, M. G. H., Bernstein, P., Gehring, W. J., & Donchin, E. (1996). Event-related brain potentials and error-related processing: An analysis of incorrect response to Go and NoGo stimuli. *Psychophysiology, 33*, 42-53.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 141-151.

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

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

Smith, M. E., Halgren, E., Sokolik, M., Baudena, P., Musolini, A., Liegeois-Chauvel, C., et al. (1990). The intracranial topography of the P3 event-related potential elicited during auditory oddball. *Electroencephalography and Clinical Neurophysiology*, *76*, 235-248.

Sutton, R., & Barton, A. (1998). *Reinforcement learning.* MIT Press: Cambridge.

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*, 201-216.

Tiihonen, J., Rossi, R., Laakso, M. P., Hodgins, S., Testa, C., Perez, J. et al. (2008). Brain anatomy of persistent violent offenders: more rather than less. *Psychiatry Research*, *163*, 201-212.

Ullspgerger, M. & von Cramon, D. Y. (2003). Error monitoring using external feedback: Specific roles of the habenular complex, the reward system, and cingulated motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *23*, 4308-4314.

Ullsperger, M., & von Cramon, D. Y. (2006). The role of intact forntostriatal circuits in error processing. *Journal of Cognitive Neuroscience, 18*(4), 651-664.

Van Veen, V., Holroyd, C. B., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2004). Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. *Brain and Cognition*, *56*, 267-276.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological Psychology,* 51, 109-128.

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: A reexamination of available data. *Biological Psychology*, 64, 265-282.

Vogt, B. A. (2005). Pain and emotion. Interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience*, *6*, 533-544.

Vogt, B. A. (2009). *Cingulate Neurobiology and Disease.* University Press: Oxford.

Von Borries, A. K. L., Brazil, I. A., Bulten, B. H., Buitelaar, J. K., Verkes, R. J., & de Bruijn, E. R. A. (2009). Neural correlates of error-related learning deficits in individuals with psychopathy. *Psychological Medicine*, doi: 10.1017/S0033291709992017.

Wilkowski, B. M., & Robinson, M. D. (2008). Putting the brakes on antisocial behavior: Secondary psychopathy and post-error adjustments in reaction time. *Personality and Individual Differences*, *44*, 1807-1818.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders.* Geneva, Swiss: World Health Organization.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114-122.

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

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931-959.

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *Journal of Neuroscience*, *24*, 6258-6264.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*, 535-544.

11. Supplements

Instruction Study Monetary Feedback

Vielen Dank für Ihre Teilnahme an diesem EEG-Experiment zum Thema Entscheidungsfindung.

Ihre Aufgabe in den folgenden Trainingsdurchgängen ist folgende: Versuchen Sie herauszufinden, welche Symbole welchen Tasten der Responsebox zugeordnet sind. Es gibt drei Symbole, diese werden Ihnen nun kurz zum Kennenlernen vorgestellt.

Presentation of the three visual cues: star, triangle, circle.

Wie Sie gerade gesehen haben, handelt es sich um drei unterschiedliche Symbole, als Antworttasten stehen Ihnen aber nur die Tasten '1' und '2' zur Verfügung.

Das Training sieht nun folgendermaßen aus: Sie werden für 500 ms eines der drei Symbole

präsentiert bekommen, sobald anschließend ein Fragezeichen erscheint, drücken Sie bitte entweder Taste '1' oder '2'. Kurz danach erhalten sie eine Rückmeldung ob Sie richtig getippt haben oder nicht. Zwischen den Trainingsaufgaben erscheint jeweils ein kleines Fadenkreuz. Bei manchen Symbolen ist es möglicherweise einfacher eine in hohem Maße verlässliche Zuordnung zu finden als bei anderen! Bitte starten Sie das Training mit einer beliebigen Taste der Responsebox!

Forty-eight training trials were presented.

Jetzt startet das eigentliche Experiment! Sie werden wieder kurz jeweils eines der bekannten Symbole sehen, bei der Präsentation des Fragezeichens, welches maximal 2000 ms am Schirm zu sehen sein wird, antworten Sie bitte entweder mit Taste '1' oder '2' . Kurz danach erhalten Sie erneut Rückmeldung. Für jede richtige Antwort, angezeigt durch die Zahl 15 in grüner Farbe werden Ihnen 15 Cent gutgeschrieben, bei nicht passenden Antworten (15 in roter Farbe) werden Ihnen 15 Cent abgezogen.

Im Experiment müssen Sie die Symbol-Tasten-Kombinationen des Trainings als Antwortgrundlage verwenden. Allerdings werden nun die von Ihnen geforderten Zuweisungen wesentlich schwieriger werden, da es ab jetzt zusätzlich hochkomplexe Tastendrucksequenzen zu entdecken und anzuwenden gilt! In jedem der sechs Durchgänge wird eine andere Tastendrucksequenz gesucht! Ihr Startkapital beträgt 2,50 €, es ist ein Gewinn bis zu 25€ möglich. Nach jeweils 150 Trials erhalten Sie zusätzlich Feedback, wie viel Geld Sie bereits erspielt haben und wie sich Ihr Gesamtguthaben entwickelt. Hier können Sie jeweils eine kurze Pause einlegen. Sollten Sie einmal auf den Tastendruck vergessen bzw. nicht innerhalb der geforderten 2000 ms mit einem Tastendruck antworten, so läuft das Programm automatisch weiter, es werden Ihnen jedoch jeweils 15 Cent abgezogen. Nach der Hälfte der Aufgaben ist eine längere Pause geplant. Hier erhalten Sie Ihren bis dahin erspielten Geldbetrag ausbezahlt, den Rest am Ende des Experiments! Es wird auf alle Fälle ein Gewinn von mindestens 10 Euro sein! Viel Erfolg beim Experiment! Sie starten es mit einem beliebigen Tastendruck!

Instruction Study Facial Feedback

Vielen Dank für Ihre Teilnahme an diesem EEG-Experiment zum Thema Entscheidungsfindung.

Ihre Aufgabe in den folgenden Trainingsdurchgängen ist folgende: Versuchen Sie herauszufinden, welche Symbole welchen Tasten der Responsebox zugeordnet sind. Es gibt drei Symbole, diese werden Ihnen nun kurz zum Kennenlernen vorgestellt…

Presentation of the three visual cues: star, triangle, circle.

Wie Sie gerade gesehen haben, handelt es sich um drei unterschiedliche Symbole, als Antworttasten

stehen Ihnen aber nur die Tasten '1' und '2' zur Verfügung. Das Training sieht nun folgendermaßen aus: Sie werden für 500 ms eines der drei Symbole präsentiert bekommen, sobald anschließend ein

Fragezeichen erscheint, drücken Sie bitte entweder Taste '1' oder '2'. Kurz danach erhalten sie eine

Rückmeldung ob Sie richtig getippt haben oder nicht. Zwischen den Trainingsaufgaben erscheint jeweils ein kleines Fadenkreuz. Bei manchen Symbolen ist es möglicherweise einfacher eine in hohem Maße verlässliche Zuordnung zu finden als bei anderen! Bitte starten Sie das Training mit einer beliebigen Taste der Responsebox!

Forty-eight training trials were presented.

Jetzt startet das eigentliche Experiment! Sie werden wieder kurz jeweils eines der bekannten

Symbole sehen, bei der Präsentation des Fragezeichens, welches maximal 2000 ms am Schirm zu

sehen sein wird, antworten Sie bitte entweder mit Taste '1' oder '2' . Kurz danach erhalten Sie erneut

Rückmeldung. Jede korrekte Antwort wird mit einem Gesicht mit einem positiven Gesichtsausdruck zurückgemeldet. Bei falschen Antworten werden Sie ein Gesicht mit einem negativen Gesichtsausdruck erkennen. Im Experiment müssen Sie die Symbol-Tasten-Kombinationen des Trainings als Antwortgrundlage verwenden. Allerdings werden nun die von Ihnen geforderten

Zuweisungen wesentlich schwieriger werden, da es ab jetzt zusätzlich hochkomplexe Tastendruck-

sequenzen zu entdecken und anzuwenden gilt! Diese beginnen immer mit einer der zuvor gelernten

Zuweisungen! In jedem der sechs Durchgänge wird eine andere Tastendrucksequenz gesucht! Nach jeweils 150 Trials erhalten Sie zusätzlich Feedback, wie viele Antworten Sie richtig hatten.

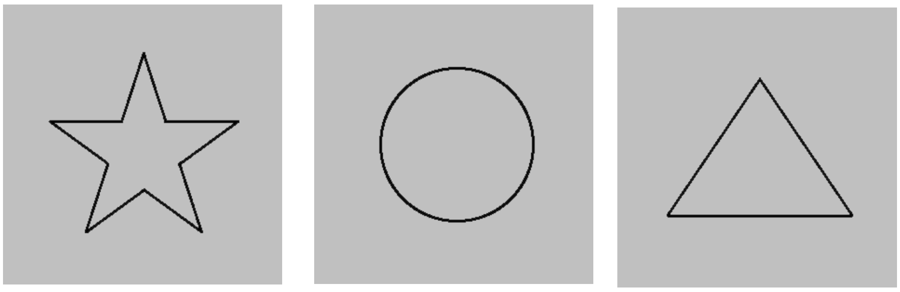
Hier können Sie jeweils eine kurze Pause einlegen. Sollten Sie einmal auf den Tastendruck vergessen bzw. nicht innerhalb der geforderten 2000 ms mit einem Tastendruck antworten, so läuft das Programm ohne Konsequenzen automatisch weiter. Nach der Hälfte der Aufgaben ist eine längere Pause geplant. Am Ende des Experiments werden Sie eine leistungsbezogene finanzielle Entschädigung erhalten.

Auf den nächsten Folien lernen Sie die verwendeten Feedbackstimuli kennen.

Visual presentation of the two female and two male posers, depicting each a happy and an angry facial expression.

Wenn Sie keine weiteren Fragen haben können Sie das Experiment mit einem Tastendruck starten. Viel Vergnügen!

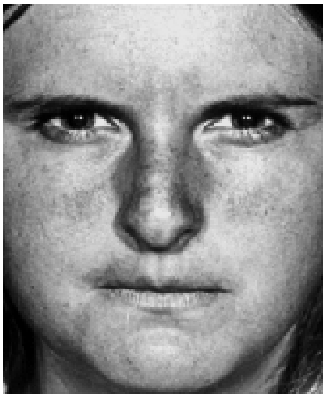
Visual Cues

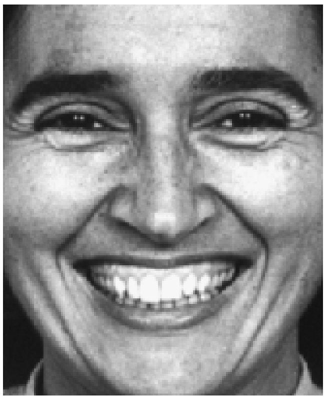


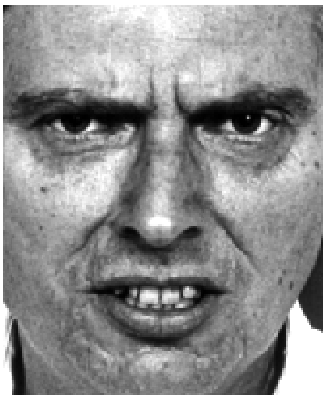
Bates, E., Federmeier, K., Herron, D., Iver, G., Jacobsen, T., Pechmann, et al. (2000). Introducing the CRL International Picture-Naming Project (CRL-JPNP). *Center for Research in Language Newsletter*, *12*, 1-14.

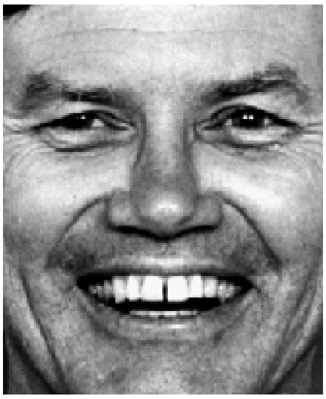
Facial Feedback Cues

(Two female and two male posers were used as feedback stimuli, each depicting a happy and an angry facial expression. Exemplarily, each poser is presented once).









Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Curriculum Vitae

Mag. rer. nat. Daniela M. Pfabigan

Date of birth:	03. 06. 1981
Contact information:	daniela.pfabigan@univie.ac.at

Education

11/2010	Autumn School “Cognitive, Affective, and Nociceptive Functioning of the Anterior Cingulate Cortex, University of Jena, Germany
07/2010	CINN Summer School on Computational Neuroscience, University of Reading, UK
03/2010	BION Spring School, Gießen, Germany 'Functional magnetic resonance imaging in psychological research'
10/2009	‘Introduction to time-frequency analysis’, pre-congress workshop at the 49th Annual SPR Meeting, Berlin, Germany
Since 10/2007	PhD student at the Brain Research Laboratory, Faculty of Psychology, University of Vienna; PhD-project ‘Decision making and feedback processing investigated with event-related potentials’
Since 03/2006	Advanced training in neuropsychology, Austria (GNPÖ)
01/2006 – 02/2007	Post-graduate courses in Clinical Psychology (University of Vienna)
10/1999 – 10/2005	University of Vienna, Faculty of Psychology, Master-thesis on ‘Neuronal correlates of depth perception – A DC-EEG study’
09/1991 – 06/1999	Highschool Waidhofen/Thaya, Austria
09/1987 – 06/1991	Primary school Karlstein/Thaya, Austria

Places of Employment

10/10 - current	Eggenburg Institute for Complex Systems, Health, and Neuroscience (EICoN), Eggenburg, Austria; Lecturer at the University of Vienna, Austria, for the Departments of Clinical and Biological Psychology
08/09 – 09/10	University of Vienna, Austria, Department of Clinical Psychology, Pre-doctoral Assistant
09/06 – 08/09	University of Vienna, Austria, Brain Research Laboratory, Student Assistant
12/06 – 09/08	Rehabilitation Centre Bad Pirawarth, Austria, Clinical Neuropsychologist
01/06 – 11/06	Hospital St. Pölten, Austria, Internship Clinical Neuropsychologist

Publications

10/2010	Pfabigan, D. M., Alexopoulos, J., Bauer, H. and Sailer, U. , Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. Psychophysiology, no. doi: 10.1111/j.1469-8986.2010.01136.x (access 15-10-2010).
09/2010	Pfabigan, D. M., Alexopoulos, J., Kryspin-Exner, I. & Sailer, U. (2010). Effects of antisocial personality traits on event-related potentials during face processing. *International Journal of Psychophysiology*, 77, 277-278.
09/2010	Schreiner, T., Alexopoulos, J., Pfabigan, D. M. & Sailer, U. (2010). Facial cues affect the feedback negativity to offers in the Ultimatum Game. An EEG investigation. *International Journal of Psychophysiology*, 77, 337.
04/2010	9th Scientific Meeting ÖGP, Salzburg, Talk 'Antisocial personality traits and feedback processing investigated with event-related potentials'
10/2009	Alexopoulos, J., Pfabigan, D., Fischmeister, F.P.S. & Bauer, H. Do we care about the powerless third? Psychophysiology, 46, S112.
10/2009	Pfabigan, D., Alexopoulos. J., Sailer, U. & Bauer, H. Antisocial personality traits and feedback processing – An ERP study. Psychophysiology, 46, S78.
04/2009	Int. Symposium on the Neural Basis of Decision Making, Groesbeek, Poster 'Antisocial Per-

	sonality Traits are Reflected in the Feedback Related Negativity'
07/2008	XXIXth International Congress of Psychology, Berlin, Poster ‘Better than expected – Decision Making with correct responses’
07/2008	6th FENS Forum Genf, Poster ‘Better than expected – Decision Making and correct responses’
07/2006	5th FENS Forum Wien, Poster ‘Frontal Lobe Contribution to Depth Perception’
04/2006	7th Scientific Meeting ÖGP, Klagenfurt, Poster ’Neuronale Korrelate der räumlichen Tiefenwahrnehmung’

Research Topics

- Neuronal correlates of feedback processing and its relation to personality constructs
- Emotional processing and physiological correlates

Scholarships

08/2008	Förderstipendium University of Vienna, StudFG 2008
02/2006	Leistungsstipendium University of Vienna, StudFG 2005