How reward value and probability drive human brain function and behavior

Iris Schutte

How reward value and probability drive human brain function and behavior

Iris Schutte

Printing and cover layout by Ipskamp Printing, Enschede

ISBN: 978-90-393-7192-3

How reward value and probability drive human brain function and behavior

Hoe beloningswaarde en waarschijnlijkheid menselijke hersenfunctie en gedrag bepalen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op

maandag 7 oktober 2019 des middags te 2.30 uur

door

Iris Schutte

geboren op 27 juni 1985
te Enschede

Promotoren:
Prof. dr. J.L. Kenemans
Prof. dr. J.M.P. Baas

Copromotor:
Dr. D.J.L.G. Schutter

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Table of contents

Chapter 1

General Introduction

Why do humans behave the way they do? Humans as well as most other species have an inherent tendency to seek out rewards and to avoid punishments. This tendency is crucially important for survival and reproduction, bodily homeostasis, and psychological wellbeing. Rewards can be defined as objects, activities, or situations that have the potential to: (1) induce positive emotion (e.g. pleasure), (2) increase the likelihood and strength of behavior that led to obtainment of the reward (reinforcement), (3) elicit approach behavior and "consummation" of the reward, and (4) lead to successful avoidance of a punishment (Kim, Shimojo, & O'Doherty, 2006; Schultz, 2015). The studies in this thesis are specifically devoted to reward pursuit and how the human brain anticipates rewards.

According to Schultz (2015), rewards can be categorized into primary rewards and non-primary rewards. He defines primary rewards as conditioned or unconditioned rewards that serve to directly promote survival and reproduction (e.g. nutrient intake and sex). In Maslow's Hierarchy of Needs theory (Maslow & Lewis, 1987) these primary rewards are referred to as "physiological needs", fundamental needs that must be fulfilled before one seeks to fulfill other needs such as compliments, social activities, friendship, and exploration. Schultz (2015) refers to this latter category of needs as non-primary rewards. These non-primary rewards may have acquired motivational value throughout evolution, because they increase the likelihood of survival and reproduction, for example by enhancing safety, cooperation, and the likelihood to find new food resources (Singh, Lewis, Barto, & Sorg, 2010). Secondary rewards (e.g. money) are a type of non-primary rewards that acquired motivational value via a learned association with primary rewards (Sescousse, Caldú, Segura, & Dreher, 2013).

The desire, "wanting" (Berridge, 2007) aspect of reward guides our behavioral actions towards obtainment of the reward. According to Panksepp's "SEEKING" framework (Wright & Panksepp, 2012) this is mediated by what he calls the "seeking system", which largely consists of brain structures connected with the dopaminergic midbrain and ventral striatum. The brain selects those behavioral actions that have high probabilities of leading to highly valuable outcomes. Therefore, the *value* as well as the *probability* of future rewards should be computed in order to make optimal decisions. Both these aspects are key elements of influential

theories of human decision-making, such as Blaise Pascal's expected utility theory and Kahnemann & Tversky's prospect theory (Kahneman & Tversky, 2013). One main question the present thesis aims to answer is how the human brain assesses the value and probability of potential future rewards.

The assessment of prospective reward value, as well as the prediction of its probability are subjective. That is, both differ strongly between individuals, and also vary within individuals dependent on particular states. For example, it may be very tempting to buy all kinds of high-calorie foods in the grocery store when being hungry. Hunger in this example is a motivational state factor that increases the subjective value of food (perhaps especially of high-calorie foods) and drives our behavior towards food acquisition. However, in this particular hunger state, individuals will differ in their preference for certain foods over others. These individual preferences may for example be determined by taste and by how easy it is to obtain the food. Furthermore, apart from state-dependent differences in reward valuation and individual preferences there are differences in "motivational style" that are more trait-like in nature. That is, some individuals are driven more by positive reward value (e.g., the good taste of strawberries), whereas others are driven more by negative value or punishment (e.g., the immediate occurrence of rashes by eating strawberries when being allergic to them) (Frank, Woroch, & Curran, 2005).

Clarification of factors that underlie individual differences in reward valuation is relevant for our understanding of various mental disorders associated with abnormal valuation of reward or motivational deficits. Depression is often characterized by anhedonia, a loss of pleasure in things that the patient used to enjoy, and hypersensitivity to punishment. Anhedonia may alternatively be described as a motivational deficit ("motivational anhedonia"), i.e., diminished reward pursuit (Treadway & Zald, 2011). Reward pursuit in humans has been operationalized in the laboratory for example by tasks in which subjects have to exert effort to obtain a reward. It has been shown that anhedonic subjects are less inclined to exert more effort for the opportunity to receive a larger or more probable reward compared to control subjects (Treadway, Bossaller, Shelton, & Zald, 2012). This tendency in patients has been hypothesized to relate to impairments in the coding of reward-cue information (Barch, Treadway, & Schoen, 2014).

Other examples of disorders associated with motivational deficits include Attention Deficit Hyperactivity Disorder (ADHD) (Sonuga-Barke, 2005), schizophrenia (Barch et al., 2014; Gold, Waltz, Prentice, Morris, & Heerey, 2008) and Parkinson's disease (Chong et al., 2015). Drugs that block or stimulate dopamine neurotransmission are one type of drugs often prescribed to treat motivational deficits. However, these as well as other type of drugs, e.g., affecting noradrenergic or serotonergic signaling, are not uniformly effective. This is due to the heterogeneity of these disorders with respect to symptoms and etiology. Thus, there is a need for a better understanding of the underlying mechanisms that may contribute to symptoms of impaired motivation and reward pursuit, as well as for predictors of treatment outcome.

The experiments described in the first part of this thesis (Chapter 2-4) focus on the assessment of (individual) differences in motivation and reward pursuit, using the methods of reflex modulation, learning-style assessment, and EEG (spontaneous oscillations). The second part of this thesis will focus on how the healthy human brain assesses information about future rewards.

1.1 The approach taken in this thesis

Overt approach behavior in order to obtain a reward involves a series of neural processes. Reward cues in the environment trigger neural processes associated with the prediction of reward value and probability. These prediction-related processes may interact with brain processes involved in e.g. attention and cognitive control functions (which may prioritize certain inputs over others). The chain of events ultimately leads to action selection. One of the aims of the current thesis is to analyze the neural substrates of these processes in healthy subjects. This is an excellent way to disentangle these processes and it provides an objective basis to differentiate these processes with respect to the effects of neurotransmitters. This approach will clarify the contributions of distinct processes involved in the anticipation of reward.

This research may also contribute to a better understanding of abnormal approach behavior. Breaking down approach behavior into different underlying processes (i.e., reward value versus probability

prediction) makes it possible to investigate which of these putatively separate processes, if any, is disturbed in patients with abnormal approach behavior.

As will be detailed later in paragraph 1.5 it is expected that the neural processes related to the prediction of reward value and probability are under control of separate neurotransmitters, i.e., dopamine and noradrenaline, respectively. These processes may therefore respond differentially to different treatments (this is tested in Chapter 6). Disentangling these processes and associated neurotransmitters is therefore relevant to treatment choice and the prediction of treatment outcome in psychiatric patients with motivational deficits.

The original theoretical basis for this approach was the theoretical framework provided by Glimcher (Glimcher, 2004). He argued that the function of neurons in monkey frontal and parietal regions associated with cognitive control is better understood in terms of "neuro-economics". In his earlier work (Platt & Glimcher, 1999) he observed that in monkeys activity of neurons in a parietal brain area associated with cognitive control (the lateral intra-parietal area) was modulated by the value and probability of prospective food rewards. It is possible that cortical areas of the fronto-parietal network associated with higher order cognitive control functions such as this lateral intra-parietal area (Corbetta & Shulman, 2002) may also serve the more fundamental assessment of prospective reward value and probability. However, we hypothesized that the actual predictions of reward value and probability are computed in sub-/meso-cortical regions (basal ganglia, ventral tegmental area, anterior- and posterior cingulate cortex) that have a more direct connection to the amygdala and the autonomic nervous system. The predictions computed in the abovementioned structures would then serve as an input for the fronto-parietal cortical network associated with higher-order cognitive control functions. The studies of the second part of this thesis aim at identifying reward- and probability-related activity of neurons in meso- and surrounding cortex, and aim at determining how this activity translates into behavioral action (Chapter 5). Chapter 6 describes an experiment on the neurotransmitters involved in these neural processes associated with reward value and probability anticipation. In the remainder of this

introduction I will provide a background on the concepts, techniques, and behavioral tasks involved in the different chapters.

1.2 Assessing the subjective value of reward

As mentioned earlier, one of the aims of this thesis is to investigate the cortical mechanisms associated with the prediction of reward value and probability. As will be explained in more detail later, this will be investigated using a behavioral task that involves manipulations of reward value and probability of (un)rewarded targets. In half of the task blocks subjects are rewarded for adequate responses to a target, whereas in the other half of the blocks they are not. The first study of this thesis project aims to titrate the specific monetary reward to be used during this behavioral task. The appreciation of monetary rewards of varying magnitude will be assessed by means of measuring startle eye blink reflex modulation during a passive gambling task.

The startle eye blink reflex is a defensive response which consists of the contraction of the orbicularis oculi muscle, a sphincter muscle located around the eye lids. The startle eye blink reflex is elicited by a sudden and intense stimulus (the startle probe), such as a loud noise. The intensity of the contraction can be measured by placing recording electrodes just below the eye. It has been consistently shown that the startle blink reflex increases in the context of an aversive stimulus (e.g. aversive pictures) or threat-of-shock (Baas, Kenemans, Böcker, & Verbaten, 2002; Baas et al., 2009; Grillon & Baas, 2003), whereas it tends to decrease in the context of a pleasant event (Lang, Bradley, & Cuthbert, 1990; Vrana, Spence, & Lang, 1988). Lang and colleagues (1990) theorized that this modulatory effect can be explained by a relative match or mismatch between the motivational state induced by the pleasant/unpleasant context and the defensive nature of the startle reflex (the motivational-priming theory). Unpleasant events or threats may induce preparation for defense, an action disposition that matches the response elicited by the aversive startle probe. The opposite may hold for pleasant stimuli that induce approach. The startle eye blink reflex may thus be seen as an objective index of motivational direction. As such, the startle eye blink will be used to assess which monetary reward magnitudes are large enough to have a significant impact on motivation. The

study described in Chapter 2 aims to examine potential differences in average subjective value between rewards of varying magnitude.

1.3 Individual differences in motivational style

Motivational style here refers to the extent to which subjects are sensitive to reward or punishment. An elegant way to assess individual differences in motivational style is to characterize subjects as positive or negative learners. Positive learners learn more from the positive outcomes of their actions, whereas negative learners learn more from the negative outcomes of their actions. The probabilistic selection task (Frank, Seeberger, & O'reilly, 2004) is used frequently in the literature to assess motivational style in healthy and patient populations, and has been used to test the effects of dopaminergic agonists on positive/negative learning biases. In the PST, individuals first learn to choose which stimulus characters of three different stimulus pairs are most likely to be rewarded. In a subsequent test phase it is tested whether choices predominantly reflect the influence of positive or the influence of negative feedback during the preceding learning phase. Individual differences in motivational style as assessed by the PST have been repeatedly related to neural and genetic measures associated with striatal and dopamine function (Cavanagh, Bismark, Frank, & Allen, 2011; Cockburn, Collins, & Frank, 2014; Cox et al., 2015; Doll, Hutchison, & Frank, 2011; Frank et al., 2004; Jocham, Klein, & Ullsperger, 2011; Waltz, Frank, Robinson, & Gold, 2007).

We originally aimed to classify subjects as positive or negative learner in order to investigate whether positive/negative learning style would predict the amplitude of an event-related potential (Box 3) associated with the assessment of reward value (see Paragraph 1.4). However, as described in Chapter 3, we observed that in seemingly straightforward implementations of the PST, categorization as negative or positive learner may not purely reflect motivational style, but also the salience or discriminability of the stimuli in the task.

A different way to assess individual differences in motivational style is to evaluate the relative contribution of spontaneous oscillations in the electro-encephalogram (See Box 2; Box 1 provides an overview of the basic principles of electro-encephalography). Prior studies have shown that

spontaneous oscillations in the theta (4-7 Hz) range and the ratio between theta and beta (13-30 Hz) oscillations are related to approach and avoidance-related motivational tendencies and reward-based decision making (Massar, Kenemans, & Schutter, 2014; Schutter & Van Honk, 2005; Schutter, de Weijer, Meuwese, Morgan, & van Honk, 2008). Schutter and van Honk (2005) observed a positive relationship between disadvantageous decision making during the Iowa gambling task and the ratio between theta and beta oscillations during resting-state. This relationship was driven by a positive association between theta oscillations and reward-seeking behavior (Massar et al., 2014). Recent studies have also shown that high EEG theta/beta ratio relates to poor executive control (Angelidis, Hagenaars, van Son, van der Does, & Putman, 2018; van Son et al., 2018; van Son, Angelidis, Hagenaars, van der Does, & Putman, 2018).

Resting-state oscillatory activity including the theta/beta ratio has a high test-retest reliability (Angelidis, van der Does, Schakel, & Putman, 2016; Corsi-Cabrera, Galindo-Vilchis, del-Río-Portilla, Arce, & Ramos-Loyo, 2007). In Chapter 4 the relationship between theta/beta EEG ratio and the ability to adapt behavioral choices in a reward task when reward and punishment contingencies change over time will be investigated.

Box 1 Electro-encephalography – basic principles

During the experiments described in chapter 4-6 of this thesis electro-encephalography (EEG) signals were recorded. EEG signals are indirect reflections of neuronal activity, which can be measured with electrodes placed on the subject's scalp. The recorded EEG signal is thought to mainly reflect post-synaptic potentials (PSPs). PSPs are changes in the membrane-potential of post-synaptic neurons, which are induced by current flow that is triggered by the binding of neurotransmitters to the cell receptors. The EEG signal can only be picked up by the electrodes on the scalp when PSPs occur at the same time in a large number of neurons that have a similar orientation and the same current flow pattern. Recorded EEG signals are thought to mainly reflect activity from cortical neurons (Luck, 2012). PSPs of cortical neurons summate and spread across the brain tissue and scalp, which is termed volume conduction.

EEG has a very high temporal resolution, it reflects neural activity with almost no delay. However, it has a relatively poor spatial resolution. It is difficult to define the intracranial source(s) of the recorded EEG signal. The distribution of the EEG signal across the different scalp electrode locations could virtually be explained by an infinite number of neural generators with certain current flow orientation.

Box 2 Resting-state EEG

During the experiment described in Chapter 4 resting-state EEG was recorded. Resting-state EEG refers to the measurement of EEG during periods without a specific task. It is usually recorded over a period of 2 to 5 minutes. The resting-state EEG signal consists of spontaneous oscillatory activity in different frequency bands: delta (0.5- 3.5 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (> 30 Hz). There is ample evidence showing that spontaneous oscillations play a critical role in various brain functions and in communication between groups of neurons (Fries, 2005; Knyazev, 2007). Chapter 4 focuses on the theta and beta frequency bands, and the ratio between both. Theta oscillations during rest have been associated with e.g. emotional processing (Knyazev, 2007). Theta activity has been source-localized to the medial frontal cortex (Scheeringa et al., 2008), which has strong connections with subcortical structures including the dopaminergic driven ventral tegmental area. Beta oscillations have been associated with top-down control, decision-making, and inhibition (Donner & Siegel, 2011; Engel & Fries, 2010). It has been argued that the ratio between theta and beta oscillations reflects the balance between subcortical motivational drive (represented by theta) and cortical regulatory activity (represented by beta) (Schutter & Van Honk, 2005).

Box 3 Event-related potentials

Event-related potentials (ERPs) in the EEG signal are induced by specific events such as task stimuli. ERPs are usually elicited during each trial of the task, but are often hardly visible at the single-trial level. This is because the signal is relatively small and hidden in the background EEG signal. The ERP can be made visible by averaging over dozens of trials. Trial-averaging cancels out the EEG signal and random noise without a temporal relationship with the stimulus, and isolates the neural activity specifically associated with presentation of the stimulus. The resulting ERP is said to be "time-locked" and "phase-locked" to the presented stimulus. Time-locked means that the change in neural activity occurs at (approximately) the same moment after stimulus onset during each trial. Phase-locked means that the phase angles of the oscillatory activity after stimulus presentation are consistent across trials. The event-related potential technique is a valuable tool for investigating how stimulus processing is affected by specific experimental manipulations. In the experiments in this thesis (Chapter 5 and 6) we investigated ERPs as elicited by stimuli that predict reward value and probability of (rewarded) targets.

1.4 Pinpointing anticipated reward value and probability

As outlined in section 1.1, both from a scientific and from a clinical perspective it is important to better understand the neural mechanisms underlying reward value and probability assessment during reward anticipation. One main aim of the current thesis was to pinpoint and disentangle the cortical substrates (putatively in the meso- and surrounding cortex) of reward value and probability predictions.

1.4.1 The cingulate cortex in reward value and probability anticipation

Prior neuroimaging studies (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Hahn et al., 2009; Howard, Gottfried, Tobler, & Kahnt, 2015; Kable & Glimcher, 2007; Kirsch et al., 2003; Knutson, Taylor, Kaufman, Peterson, & Glover, 2005; Padmala & Pessoa, 2011; Smith et al., 2009; Yacubian et al., 2006) have sought to identify brain regions involved in reward anticipation. With respect to cortical substrates, from these studies it follows that different portions of the cingulate cortex as well as different regions of the ventro-medial prefrontal cortex (which, according to Haber & Knutson (Haber & Knutson, 2010) includes Brodmann area (BA) 10, 11, 32) are affected by information about the value and/or probability of prospective rewards. See below for which of these regions are more tightly coupled to reward value and which more tightly to probability anticipation.

The cingulate cortex is part of the meso-cortex and can be divided in an anterior (anterior cingulate cortex; ACC; Brodmann area (BA) 24, 25, 32, 33) and a posterior part (posterior cingulate cortex; PCC; BA 23 and 29-31). The ACC can be further subdivided into a pregenual and ventral part (together referred to as rostral ACC (rACC)) and a dorsal part (dorsal ACC (dACC)), also referred to as mid-cingulate cortex (Stevens, Hurley, & Taber, 2011; Vogt, 2016). The rACC is strongly connected to subcortical structures (e.g., the amygdala and ventral striatum) and to the autonomic nervous system, consistent with its putative role in emotion-related processes. The dACC is strongly connected to motor areas and areas implicated in cognitive control (e.g., the dorso-lateral prefrontal cortex) (Stevens et al., 2011; Vogt, 2016).

Results of the previously mentioned neuroimaging studies (Breiter et al., 2001; Hahn et al., 2009; Howard et al., 2015; Kable & Glimcher, 2007;

Kirsch et al., 2003; Knutson et al., 2005; Padmala & Pessoa, 2011; Smith et al., 2009; Yacubian et al., 2006) point to expected reward value being represented in cortical regions surrounding the rACC, including the medial orbitofrontal cortex and dACC. However, activations in the PCC have also been observed. These regions act in concert with subcortical structures such as the ventral striatum and amygdala. Activations related to the anticipation of reward probability have been reported in the PCC (Knutson et al., 2005), the dACC (Smith et al., 2009), and in parts of the ventro-medial PFC (BA 10/11; Knutson et al., 2005; Yacubian et al., 2006).

In addition to the localisation of the relevant brain areas involved in reward anticipation, the temporal dynamics is another critical aspect for understanding temporal relations between activity associated with reward value and probability anticipation. The event-related potential (ERP) technique (Box 3) is well suited in temporally separating sequential neural activities that occur in brief time windows. In this thesis the ERP technique will be used to examine and disentangle cortical processes related to reward value and probability assessment.

1.4.2 Event-related potential studies of reward and probability processing

The ACC is one of the target structures of dopamine signals originating from the midbrain (see Box 4 for more background on the dopamine system). It has been hypothesized that transient midbrain dopamine signals (dips and bursts) reflect reward predictions (e.g., elicited by a cue that in itself is not a primary reward) and reward prediction errors (e.g., elicited by unexpected omission of a primary reward), which are presumed to play an important role in reinforcement learning (reinforcement learning theory; Holroyd & Coles, 2002; Holroyd & Yeung, 2012). This theory is based on electrophysiological recordings in primates demonstrating that unexpected reward delivery and cues predicting reward trigger dopamine burst firing in the midbrain. In contrast, transient pauses in dopamine firing were observed when expected rewards were not delivered (Schultz, Dayan, & Montague, 1997; Schultz, 2015). In humans, pauses in dopamine firing were originally hypothesized to be reflected in an ERP termed the feedback-related negativity (FRN) and a related component termed error-related negativity (ERN) (Holroyd & Coles, 2002). The FRN and ERN are negative-going brain potentials elicited around 200-300 ms after receiving negative

feedback or bad outcomes, and around 50 ms after making mistakes, respectively (Gehring & Willoughby, 2002; Holroyd & Coles, 2002). Later studies have shown that variance in the amplitude of FRN (resulting from a loss minus gain ERP subtraction) is explained by rewarding outcomes inducing a positive potential, rather than by negative outcomes inducing a negative potential. The apparent negativity following bad outcomes was found to be a rather common ("baseline") phenomenon that is also present following neutral feedback. In contrast, positive feedback seems to induce an additional positivity (Foti, Weinberg, Dien, & Hajcak, 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Proudfit, 2015).

More recently, studies have observed positive brain potentials following reward announcing cues (Doñamayor, Schoenfeld, & Münte, 2012; Flores, Münte, & Doñamayor, 2015; Holroyd, Krigolson, & Lee, 2011). All "reward positivities" observed in these studies were elicited around 200 ms after the reward-predicting stimulus and had a fronto-central scalp distribution.

With respect to (anticipated) probability, results of a prior study by Bekker and colleagues (Bekker, Kenemans, & Verbaten, 2004) demonstrated that cues predicting the probability of a behaviorally relevant target stimulus elicit a P300 ERP (or "P3b") with a parietal scalp distribution. The amplitude of this ERP was increased for cues indicating high compared to low probability of a subsequent target stimulus. This observation fits prior results indicating that the amplitude of parietal P300 varies as a function of the degree to which a stimulus provides information (Gratton et al., 1990). Both studies are consistent with the idea that the P300 reflects the continuous updating of probability (Duncan-Johnson & Donchin, 1982). Relatedly, a theoretical account by Friston and Stephan (2007) suggests that the P300 is part of a mechanism by which the brain minimizes surprise or prediction errors by the continuous updating of internal probability representations.

Consistent with neuroimaging studies investigating the representation of reward probability (as outlined in the previous paragraph), the most likely generators of the P300 ERP seem to be located in posterior cortical regions, including the PCC, and the temporo-parietal junction (Frodl-Bauch, Bottlender, & Hegerl, 1999; Neuhaus et al., 2006).

In sum, prior studies have shown that a frontally distributed positivity is sensitive to manipulations of reward value, and that a parietally distributed P300 ERP is sensitive to manipulations of probability. The aim of Chapter 5 is to reveal a dissociation between cortical processes related to the assessment of prospective reward value versus probability, thereby focusing on these ERPs. In this study a cued Go/NoGo (CGN) task will be used (see Figure 1 for an overview of the CGN task) in combination with ERP recording. The CGN task is adapted from the one used in our lab by Bekker and colleagues (Bekker et al., 2004). During the CGN task cue letters are presented that predict the appearance of a subsequent target letter. The probability of target letter appearance and the amount of money that could be won for adequate responding to the target stimulus will be orthogonally manipulated across different task blocks. It is expected that manipulations of reward value, and not manipulations of probability, will affect an early reward-related positivity (elicited around 200 ms) with a frontal scalp distribution. P300 amplitudes are expected to be larger in the high compared to low probability condition.

It should be noted that there is a relation between the P300 amplitude and the amount of noradrenergic signaling (see paragraph 1.5 and Box 5). This could imply that P300 reflects a kind of arousal response which is secondary to detection of a high-probability impending action or reward. P300 amplitudes may also be expected to be sensitive to manipulations of reward value, because of the known sensitivity of P300 ERPs to motivational significance of stimuli (Nieuwenhuis, Aston-Jones, & Cohen, 2005; Yeung & Sanfey, 2004). This latter effect may, however, be under control of dopamine (Pfabigan et al., 2014). Interactions between effects of varying probability and that of noradrenergic manipulation may shed light on this matter (as described in 1.5 and Ch.6). Thus, a complete dissociation between ERPs related to reward value and probability anticipation is not expected. However, we do expect specificity with respect to the manipulation of probability, which is hypothesized to only affect the P300 and not the frontal reward-related positivity. Furthermore, a pharmacological dissociation (section 1.5) of the effects of reward value and probability is expected.

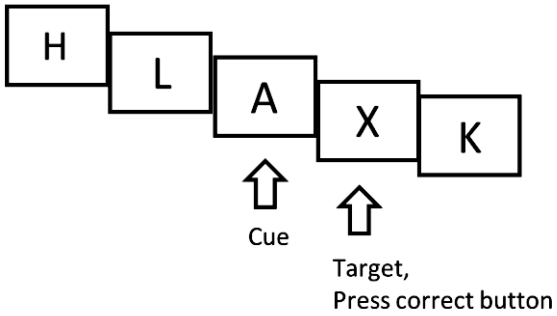


Figure 1. A typical trial sequence of the cued Go/NoGo task.
Subjects press a pre-specified button when letter X (target) follows letter A (cue), and when letter Y (target) follows letter A (cue). Reward value (amount of money that could be won for correct and fast responding to the target), as well as probability of target appearance will be orthogonally manipulated across four task blocks.

1.5 A pharmacological dissociation of reward value and probability anticipation

Another aim of this thesis is to dissociate the cortical processes related to the anticipation of reward value and probability in terms of underlying catecholaminergic neurotransmitters. With respect to reward value, prior studies have shown that drugs that block or stimulate dopaminergic signaling affect reward-value related ERPs (De Bruijn, Hulstijn, Verkes, Ruigt, & Sabbe, 2005; De Bruijn, Sabbe, Hulstijn, Ruigt, & Verkes, 2006; Forster et al., 2017; Zirnheld et al., 2004). The reward-value ERPs in these studies concern post-reward ERPs (i.e., the FRN and ERN). These findings are consistent with Holroyd and Coles' reinforcement learning theory (paragraph 1.4.2) which states that these ERPs reflect the impact of dopaminergic reinforcement learning signals on the anterior (cingulate) cortex when outcomes are better (or worse) than expected (Holroyd & Coles, 2002). This theory also posits that these dopaminergic signals "propagate back in time" to the earliest predictor of reward. Consistent with this theory, reward-anticipating cues have been shown to induce firing of midbrain dopaminergic neurons (Schultz et al., 1997) and increased activity in the (ventro)medial PFC (for the references see paragraph 1.4.1).

The firing of dopamine neurons in response to reward-announcing cues and the observation that dopamine antagonists reduce reward-value related ERPs are in agreement with the classical idea that dopaminergic signaling is positively associated with approach-related motivational tendencies (Berridge, 2007; Cools, 2008; Salamone & Correa, 2012). Based on the studies mentioned above it may be hypothesized that the frontal positivity elicited by predictive cues signaling reward value is under control of the dopamine system.

Box 4 Dopamine

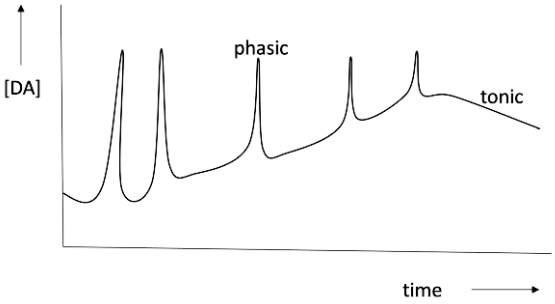
The most important dopaminergic projections in the human brain arise from the substantia nigra pars compacta and the ventral tegmental area (VTA) located in the midbrain. Dopaminergic neurons in these regions exhibit two modes of firing: a "tonic" or baseline level of dopaminergic signaling, and a transient burst-firing pattern ("phasic" signaling) (Grace, Floresco, Goto, & Lodge, 2007). See below for an example. Most of the studies on the functional role of dopamine focus on phasic signaling. The phasic responses have traditionally been viewed as reflecting positive reward-prediction errors, elicited by, e.g., occasional reward-predicting cues in a context of cues that do not predict reward (see also paragraph 1.4.2). The specific functional role of tonic signaling is less clear, but it has been hypothesized to code the average reward rate, and may modulate the amount of invested effort accordingly (Niv, Daw, Joel, & Dayan, 2007).

Multiple dopaminergic subcortical-cortical pathways have been described throughout the years. These include mesocortical projections from the VTA to prefrontal and cingulate cortex, respectively, either via the nucleus accumbens (NAcc) or directly, as well as nigrostriatal connections that modulate the signaling between different basal ganglia compartments and corresponding frontal cortex regions. Activity of dopaminergic neurons in the midbrain and part of their target structures (most notably the NAcc) is regulated by structures such as the amygdala and by specific nuclei in the brain stem and epithalamus.

The dopaminergic pathway connecting between the VTA and nucleus accumbens has been implicated in reinforcement learning and "incentive salience". The latter refers to the conditioned brain responses to reward-predicting cues (such as the reward-cues in the CGN task used in Chapter 5 and 6) that drive compulsive action to obtain the predicted reward (Berridge, 2007).

Mesocortical projections are implicated in a wide range of executive functions such as attention and inhibition (Martins, Mehta, & Prata, 2017). The nigrostriatal circuit plays an important role in voluntary movement, and has also been implicated in reward-seeking behavior. More specifically, phasic dopamine has been proposed to facilitate the execution of actions represented in the frontal cortex that lead to positive outcomes, whereas a dopaminergic dip may facilitate the suppression of actions that lead to negative outcomes (Frank & O'reilly, 2006).

Dopaminergic signaling in the abovementioned pathways occurs mainly via binding to two major dopamine receptor subtypes: the D1 and D2 receptors. Both receptor types are abundantly present in the subcortex, with the largest density in the basal ganglia. D1 receptors are, in contrast to D2 receptors, also highly expressed in (frontal-) cortical regions (Hall et al., 1994).



The (simplified) figure above displays the hypothesized modes of dopamine neuron firing: transient "phasic" bursts versus slow "tonic" background signaling.

With respect to probability, Nieuwenhuis and colleagues (2005) theorized that the P300 reflects phasic activity of the locus coeruleus (LC) noradrenaline system (Box 5). Evidence for this claim comes from studies demonstrating that lesions of the LC as well as noradrenergic antagonism result in an attenuation of the P300 (Pineda, Foote, & Neville, 1989; Svensson, Bunney, & Aghajanian, 1975; Nieuwenhuis et al., 2005). Moreover, a study in humans has shown sensitivity of the posterior cingulate cortex (which is, as mentioned before, a possible P300 generator) to enhancement of noradrenergic signaling by atomoxetine (Friedman et al., 2008). The link between the LC-noradrenaline system and P300 generation is furthermore supported by the observation that LC phasic responses and P300 ERPs are triggered by similar events (Nieuwenhuis et al., 2005). It may thus be hypothesized that the P300 elicited by predictive cues signaling probability is under control of the noradrenaline system.

In sum, we expect to observe a dissociation between reward value and probability anticipation in terms of underlying neurotransmitters. More specifically, it is speculated that manipulations of prospective reward value affect a frontal reward positivity and that this effect will be mediated by dopaminergic signaling. The hypothesis is that manipulations of probability affect a parietal P300 and that this effect will be mediated by noradrenergic signaling. This idea will be tested in a separate study described in Chapter 6. In this study we use the previously mentioned CGN task combined with ERP recording and pharmacological manipulation of the dopamine and noradrenaline system.

This pharmacological manipulation consists of administering drugs that block, respectively, dopaminergic and noradrenergic signaling in the brain. More specifically, haloperidol will be used to block dopaminergic signaling, and clonidine will be used to block noradrenergic signaling. These drugs are normally used in patients for the treatment of psychosis and high blood pressure, respectively. However, here these drugs will be used in healthy subjects to reveal the dissociation between reward value processing hypothesized to be dependent on dopamine on the one hand, and probability processing hypothesized to be dependent on noradrenaline on the other hand. For both drugs dosages will be used within the therapeutic dosage range (i.e., 2 mg haloperidol and 0.150 mg clonidine). Both haloperidol and clonidine may cause substantial side effects. The most common side effects of these drugs are dizziness (clonidine), extrapyramidal side effects (haloperidol), and fatigue (both drugs).

Haloperidol is a D2 receptor antagonist assumed to block dopaminergic signaling (Kapur, Zipursky, Jones, Remington, & Houle, 2000). This idea is consistent with increased prolactine levels after haloperidol 2 mg (dopamine inhibits prolactine release; Frank & O'reilly, 2006); reduced choice performance towards monetary gains even under haloperidol 1mg (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006); reduced negative reward-prediction errors under haloperidol (De Bruijn et al., 2006; Forster et al., 2017; Zirnheld et al., 2004); and, recently, impaired inhibitory control after 2 mg haloperidol (Logemann et al., 2017).

Clonidine is an alpha-2 agonist with a net inhibiting effect on noradrenergic signaling, via pre-sysnaptic mechanisms leading to the inhibition of release (Svensson et al., 1975). Recent as well as older studies

point to impairing effects of clonidine on aspects of human attention and inhibitory control (Clark, Geffen, & Geffen, 1989; Coull, Nobre, & Frith, 2001; Logemann, Böcker, Deschamps, Kemner, & Kenemans, 2013; Logemann, Böcker, Deschamps, Kemner, & Kenemans, 2014).

Box 5 Noradrenaline

The primary source of noradrenaline in the human brain is the locus coeruleus (LC), a nucleus located in the pons of the brain stem. The LC projects throughout almost the entire brain, except for the globus pallidus and the dorsal striatum (Berridge & Waterhouse, 2003). Noradrenaline acts at three receptor subtypes: alpha-1, alpha-2, and beta. Alpha-1 and beta receptors are thought to be primarily localized at post-synaptic locations, whereas alpha-2 receptors are also localized pre-synaptically (Berridge & Waterhouse, 2003); the presynaptic alpha-2 receptors have a higher affinity for certain substances, such as clonidine, than the postsynaptic alpha-receptors. Binding of noradrenaline or an alpha-2 agonist such as clonidine to the pre-synaptic alpha-2 receptor therefore primarily results in the suppression of noradrenergic transmission via presynaptic mechanisms (Svensson et al., 1975).

LC neurons exhibit tonic (background) and phasic patterns of firing. See Box 4 for an example. Note, however, that there are differences between the firing patterns of dopamine and noradrenaline. For example, noradrenergic neurons exhibit relatively short burst firing activity and these short bursts are followed by inhibition of firing (Berridge & Waterhouse, 2003).

Stimuli associated with P300 generation such as highly improbable stimuli elicit a phasic increase in noradrenergic firing (Nieuwenhuis et al., 2005). These increases in noradrenergic release in the cortex lead to increases of the sensitivity of these cortical areas to afferent input from other regions, which, in turn, may improve behavioral responding (Berridge & Waterhouse, 2003; Nieuwenhuis et al., 2005).

The level of phasic firing seems to be related to the level of tonic firing. Strong phasic responses are associated with moderate levels of background (tonic) activity. This "phasic mode" is characterized by good task performance. When tonic levels are either low or high, phasic noradrenaline responses are weak. The "tonic mode", characterized by high levels of tonic activity, has been associated with distractibility and poor task performance.

These different firing modes have been hypothesized to serve different functions. Whereas the phasic mode may support task performance and engagement within the current task, the tonic mode may support exploration of (reward) options outside the current task and disengagement from the current task (Aston-Jones & Cohen, 2005).

1.6 Outline of the thesis

The first part of this thesis (Chapter 2-4) focuses on the assessment of (individual) differences in motivation and reward-based decision-making. Chapter 2 describes two experiments implementing the startle modulation paradigm. These studies aim to assess the subjective value of monetary rewards of varying magnitude.

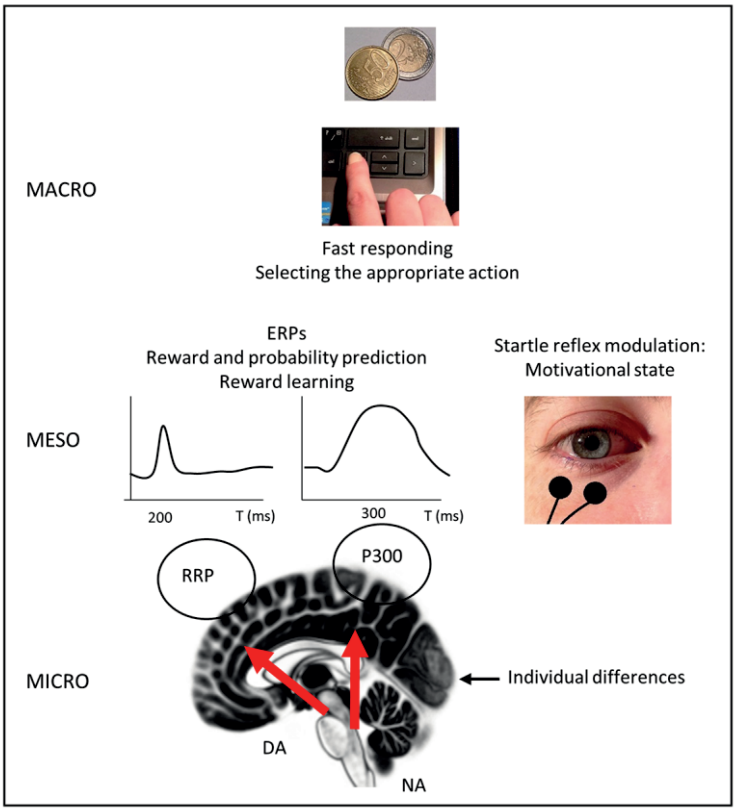


Figure 2. The impact of reward prospect will be examined at different levels in this thesis. At the micro level, cues predicting reward value and probability lead to neurotransmitter release. Dopamine (DA) release in the frontal cortex is hypothesized to be reflected in the reward-related positivity (RRP) event-related potential. Noradrenaline (NA) release in the posterior cortex is hypothesized to be reflected in the P300 ERP. The latter ERP is hypothesized to be sensitive to probability. Furthermore, at the meso level reward obtainment/loss is hypothesized to affect the startle reflex. At the macro level, prospective reward value and probability will affect behavior (e.g., reaction times and making choices). The brain image was obtained using sLORETA software (Pascual-Marqui, 2002).

In this chapter we furthermore investigated whether startle modulation in a reward context is predominantly characterized by startle inhibition after winning or by startle potentiating after missing out on an anticipated reward. Chapter 3 focuses on the probabilistic selection task. In three separate studies we show that the probabilistic learning task does not always only assess positive and negative learning style.

In Chapter 4 we test the relationship between theta/beta ratio and reversal learning during a reversal learning gambling task. The second part of this thesis (Chapter 5-6) focuses on the neural processes associated with the assessment of prospective reward value and probability. In Chapter 5 we aim to disentangle the neural processes associated with reward value processing from those associated with probability anticipation by investigating which ERPs are specifically affected by cued reward value and probability, respectively. In Chapter 6 we further elaborate on this matter and focus on the role of dopamine and noradrenaline in reward value and probability anticipation. Chapter 7 provides a general discussion of the different chapters. Figure 2 summarizes the different levels at which reward anticipation will be examined.

Chapter 2

Startle-reflex modulation due to anticipation, withholding and obtainment of reward – a non-replication

Iris Schutte
Johanna M.P. Baas
Ivo Heitland
J. Leon Kenemans

Submitted

Author contributions

LK, JB, and IS were involved in study conceptualization; IS, JB, and IH designed the task; IS programmed the task; IS and a master student collected the data, IS performed the startle analyses with input from IH and JB, and wrote the initial draft of the manuscript; all authors contributed to the final manuscript.

Abstract

Previous studies have not clearly demonstrated whether motivational tendencies during reward feedback are mainly characterized by appetitive responses to a gain or mainly by aversive consequences of reward omission. In the current study this issue was addressed employing a passive head or tails game and using the startle reflex as an index of the appetitive-aversive continuum. A second aim of the current study was to use startle-reflex modulation as a means to compare the subjective value of monetary rewards of varying magnitude. Startle responses after receiving feedback that a potential reward was won or not won were compared with a baseline condition without a potential gain. Furthermore, startle responses during anticipation of no versus potential gain were compared. Consistent with previous studies, startle-reflex magnitudes were significantly potentiated when subjects anticipated a reward compared to no reward, which may reflect anticipatory arousal. Startle-reflex magnitudes after reward feedback were significantly larger when money was at stake and the actual outcome of the trial was "not won" compared to "won". This pattern was apparent across three different reward-magnitude conditions. Specifically for the largest reward (20-cents) the difference in startle-reflex magnitude between winning and not winning appeared to be mostly driven by potentiation when a potential reward does not materialize, rather than by inhibition as a result of a positive outcome. However, neither of these effects were replicated in a more targeted second experiment. The discrepancy between these experiments may be due to differences in motivation to obtain rewards or differences in task engagement. A future experiment may shed more light on the question whether startle-reflex modulation after feedback is mainly characterized by appetitive responses to a gain or mainly by aversive consequences of reward omission.

Introduction

Appetitive and aversive motivations are fundamental for human functioning as they promote automatic behavioral tendencies and adaptive responses in the context of reward and punishment (Elliot, 2006; Krieglmeyer, Deutsch, de Houwer & de Raedt, 2010). A large body of research has shown that the startle eye blink reflex captures motivational direction (i.e., appetitive or aversive oriented motivation) within the context of emotional picture-viewing (Lang & Bradley, 2013). More specifically, the startle eyeblink reflex is typically potentiated in the context of an aversive stimulus and inhibited in the context of an appetitive stimulus (e.g., Vrana, Spence & Lang, 1988; Bradley, Cuthbert, & Lang, 1999; Wannemüller et al., 2015). These effects may be explained by a relative match or a mismatch between the subject's motivational state induced by the emotional context and the defensive nature of the startle reflex (Lang, Bradley & Cuthbert, 1990). In line with this theory, it has been shown that startle-reflex modulation is associated with self-reported sensitivity to reward. Startle magnitudes during positive picture-viewing were more strongly reduced in high- compared to low reward sensitive individuals (Hawk & Kowmas, 2003). Furthermore, it has been found that manipulation of approach motivation by leaning forward causally influences startle modulation during appetitive events (Price, Dieckman & Harmon-Jones, 2012). In the current study we used startle-reflex modulation as a means to assess appetitive and aversive motivational drives during the obtainment and loss of monetary rewards of varying magnitude.

A number of prior studies investigated startle-reflex modulation after obtainment and loss of monetary (Skolnick and Davidson, 2002) or food rewards (Hackley, Muñoz, Hebert, Valle-Inclán & Vila, 2009). Consistent with the motivational priming theory described above (Lang, Bradley & Cuthbert, 1990), these studies observed relative startle inhibition after obtainment of reward compared to loss of reward or punishment. These reports claim that emotional modulation of the startle reflex (i.e., a difference in startle magnitude between the reward obtainment and loss condition) is explained by obtainment of rewards causing attenuation of the blink reflex. However, because of lacking appropriate neutral conditions, or of comparative analysis of startle magnitudes in neutral versus reward

conditions, this claim is hard to evaluate. For example, in Skolnick and Davidson (2002) the difference between a neutral condition without any potential reward, and a condition with potential reward which did eventually not materialize was not statistically evaluated. Hence, previous studies have not clearly demonstrated whether motivational tendencies after reward feedback, as indexed by the startle reflex, are mainly characterized by appetitive responses to a gain or mainly by aversive consequences of reward omission, or by both. The present study addressed this issue using a design in which gain and reward omission can be compared separately to a neutral baseline, and in which reward expectation and actual reward obtained were orthogonally manipulated.

A wide variety of reward types and magnitudes are used across studies employing reward tasks. It is, however, unclear whether these different rewards have comparable subjective value on average. A second aim of the current study was to assess the subjective value of monetary rewards of varying magnitude. Startle-reflex modulation was used as a readout measure of the motivational state during the obtainment and loss of rewards of varying magnitude. Evidence from event-related potential (ERP) (e.g., Goldstein et al., 2006; Yeung & Sanfey, 2004) and imaging studies (e.g., Knutson, Adams, Fong & Hommer, 2001; Smith et al., 2010) suggests that at least some motivation-related processes operate in a graded (i.e., large reward > small reward > no reward) rather than all-or-none (i.e. any reward > no reward) fashion. However, reward magnitude was manipulated within-subjects in these studies, leaving the possibility that valuation of rewards was scaled by the context of other rewards (Nieuwenhuis et al., 2005). We, therefore, assessed motivational state during the obtainment and omission of rewards with varying magnitude in a between-subjects design.

A secondary issue in the present work concerns the phase in which subjects anticipate a possible reward. A large body of studies investigating the anticipation of emotional pictures suggests that the startle response is predominantly potentiated during anticipation of an emotional event. Specifically, the startle reflex is potentiated during anticipation of both negative and positive picture slides compared to neutral picture slides (Dichter, Tomarken, & Baucom, 2002; Lipp, Cox, & Siddle, 2001; Nitschke et al., 2002; Sabatinelli, Bradley, & Lang, 2001). This has been interpreted as

indicating that startle reflexes during anticipation reflect global arousal, rather than motivational direction (Sabatinelli, Bradley, & Lang, 2001). Startle facilitation has also been observed when participants anticipate a potential reward. During the anticipation phase of the Skolnick and Davidson (2002) lottery paradigm startle responses were potentiated when there was a chance to win money compared to when there was no chance to win money. So, in general, anticipating positive emotional information enhances startle magnitude, although startle *inhibition* has also been reported after cues signaling a potential reward compared to cues signaling a potential loss and neutral cues (Low, Lang, Smith & Bradley, 2008). A secondary aim of the present study was therefore to replicate the anticipatory startle potentiation in a passive reward feedback-anticipation condition. Startle-reflex modulation was also evaluated during the presentation of an initial information screen indicating how much money could potentially be won during the current trial. This was done because there is a possibility that information about potential rewards in itself modulates appetitive-aversive motivation or arousal.

To address the above questions, a "head or tails game" was used, which was based on the task as implemented by Skolnick and Davidson (2002). Within this context a completely orthogonal design including the factors win potential (no money, money) and outcome (not won, won) was implemented. This design allows systematic investigation of the effect of money at stake and actual outcome on startle magnitude after feedback. Finally, we included a behavioral measure of appetitive motivation or arousal by analyzing the time needed to initiate the trial as a function of win potential and magnitude. During the head or tails game participants only had to press a button to spin a coin which could either land on heads or tails. This stands in contrast to paradigms in which participants may be led to believe that their actions influence reward probability. Examples are the lottery paradigm during which participants choose their 'winning numbers' (Skolnick and Davidson, 2002), or a weather prediction task based on reinforcement learning (Hackley et al., 2009). Such paradigms may induce substantial variability in individual strategies driven by suspected associations between performance on the one hand and reward outcome on the other.

In sum, we set out to investigate startle modulation during anticipation and materialization (or not) of monetary rewards, varying reward magnitude between subjects (Experiment 1). For reasons explained later, we attempted to replicate the major results in a within-subjects design (Experiment 2).

Experiment 1

Methods

Participants

Fifty-seven volunteers were recruited using posters at the campus of Utrecht University. None of the subjects used psychoactive medication and none of them had a history of psychiatric or neurological disorders. Participants were asked to refrain from consuming caffeine and smoking on the day of the experiment. Written informed consent was obtained and this study was approved by the medical ethical committee of the University Medical Centre Utrecht. All subjects declared to have normal or corrected-to-normal vision. Subjects were given a financial compensation of 6 Euro per hour and an additional monetary bonus dependent on the experimental condition they were assigned to. Participants were assigned to one of the following experimental conditions: 1 cent, 5 cent, 10 cent or 20 cent (reward magnitude; see Task an procedure). For the 1 cent, 5 cent, 10 cent and 20 condition the total amount of money "won" was fixed at € 0,28, € 1,40, € 2,80 and € 5,60, respectively. Reward magnitude groups were matched for gender, Behavioral Activation System (BAS) and Behavioral Inhibition System (BIS) scores and age. The mean age was 22.71(±2.48), 21.59(±1.85), 22.26(±3.30) and 22.76(±3.50) years for the 1 cent, 5 cent, 10 cent and 20 cent group, respectively. All groups consisted of 10 females and 2 males. None of the participants were aware of the aim of the experiment. One participant was excluded because of impaired hearing and 7 participants were excluded during analysis because they had too few valid startle responses in one or more conditions (see Data processing). Data of one participant were discarded because of a technical issue during the experiment. Therefore, the final sample consisted of 48 participants, 12 participants in each reward magnitude group.

Stimuli and psychophysiological recordings

The task was presented on a 16 inch Dell CRT monitor. Presentation of task and startle probes was controlled by Presentation® software (version 16.0, www.neurobs.com). Startle probes consisted of 50 ms, 102 dB white noise bursts presented binaurally through headphones (Sennheiser electronic GmbH, HD201, Wennebostel, Germany). During the anticipation phase of the task (see Task and procedure) a movie clip of a yellow spinning coin on a white background was presented (width: 4.4°, height: 3.2°).

The Active-Two system (BioSemi, Amsterdam, The Netherlands) with matching Ag/Ag-Cl FLAT type electrodes with a diameter of 11 mm was used to record the EMG data. Recording electrodes were placed over the orbicularis oculi muscle below the right eye, according to the guidelines given by Blumenthal et al. (2005). Common mode and ground electrodes were placed on the forehead. The EMG signal was sampled at 2048 Hz and filtered online with a 400 Hz low pass filter (DC).

Task and procedure

Upon arrival at the laboratory, participants were informed about the experiment and consent was obtained. Participants were seated in a dimly lit room, one meter in front of a monitor on which the task was presented. Because this experiment was part of a larger study, subjects first filled in four questionnaires (including BIS/BAS to assess behavioral inhibition and behavioral activation, respectively) and were subjected to a different task for about 20 minutes. The BIS/BAS questionnaires (Carver & White, 1994; Dutch translation by Franken, Muris & Rassin, 2005) were used to test for between-group differences in aversive and appetitive tendencies (behavioral inhibition and behavioral activation, respectively), as BIS and BAS potentially interact with startle modulation during a reward task (see Hawk & Kowmas, 2003). The last eight participants filled out the questionnaires in a prior session for matching purposes. Next, headphones were applied and subjects received instructions for the experiment. Subjects were told that they were going to play a head or tails game. Written instructions (on-screen) were as follows: 'In this game you can win money during some trials. The amount of money you can win will be indicated each time. When indicated on the screen, you can press the spacebar to spin a coin. In case the coin lands on "tails" you win the indicated amount of money

and when the coin lands on "heads" you will not win this amount of money. During this experiment you will hear short, loud noises.'

The experiment started with a habituation phase during which subjects received nine startle probes while they looked at a fixation cross on the screen. The inter-startle interval during the habituation phase ranged between 20 and 23 seconds. The purpose of this habituation phase was to reduce the initial large startle responses. There was a small break of approximately 2 minutes at one-third and two-third of the task. Both rest breaks were followed by two startle habituation trials spaced apart by 10-13 seconds before the task continued.

Figure 1 presents an overview of the task. A trial consisted of different steps to assess responses to anticipation of reward and response to reward feedback. The trial started with a screen presented for 4 seconds showing whether participants could win money or not that trial (information phase). The amount of money that could be won during a trial was either € 0,- or a fixed amount of either 1 cent, 5 cent, 10 cent or 20 cent, depending on the reward magnitude group a participant was assigned to. Next, a picture of a coin was presented together with an instruction to press the spacebar in order to spin the coin. This screen was present until the spacebar was pressed. Upon pressing the spacebar, a movie clip of a spinning coin was presented for 4 seconds (anticipation phase). The anticipation phase was followed by a 7 seconds feedback screen indicating whether the coin landed on heads or tails and how much money was won accordingly (feedback phase). The feedback phase was followed by a screen displaying the amount of money won so far for 3 seconds.

The task consisted of 112 trials of which 96 with a startle probe. On a given trial, startle probes were either presented during the information phase (24 trials in total; probes were delivered between 3.5 and 4 seconds after the start of the trial), or during the anticipation phase (24 trials in total; probes were delivered between 2.5 and 3 seconds after the coin began to spin), or during the feedback phase (48 trials in total, probes were delivered between 5 and 7 seconds following feedback; 24 probes were delivered after a positive outcome ('Tails!') and 24 after a negative outcome ('Heads!)) of a trial in a counterbalanced fashion. The 16 trials without a startle probe were inserted to increase unpredictability of startle probe delivery. These trials were not considered in any of the analyses. Money could potentially

be won in half of the trials, and the outcome was "tails" in half of the trials. Outcome in terms of heads (not won) and tails (won) was evenly distributed between trials in which money and no money was at stake.

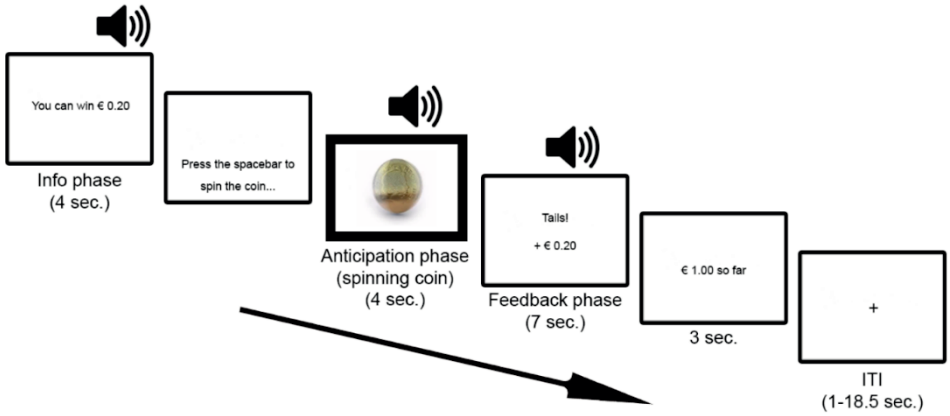


Figure 1. Overview of the task. The "head or tails paradigm" started off with an information screen indicating whether money could potentially be won or not (information phase). Next, participants pressed the spacebar to start a movie clip of a spinning coin (anticipation phase). This phase was followed by a feedback screen indicating whether the coin landed on heads or tails and how much money was won accordingly (feedback phase). Finally, an overview was given of the money won so far. Trials were separated by an intertrial interval (ITI) ranging between 1 and 18.5 seconds, to allow spacing of subsequent startle probes within the range of 25.25-33.25 seconds. Startle probes were delivered either during the information- (24 probes), anticipation- (24 probes) or feedback phase (48 probes; 12 per win potential and outcome condition). Note that unlike during the information- and anticipation phase, participants do know both win potential and trial outcome during the feedback phase. Therefore, the number of conditions/startle probes in each phase (before feedback (information and anticipation phase) and after feedback (feedback phase) is unequal. The amount of money at stake was either 0 Euro or a fixed amount of 1 cent, 5 cent, 10 cent or 20 cent, and was dependent on the reward magnitude group a participant was assigned to.

Inter-trial intervals (ITI) consisted of a white fixation cross presented on a black screen. ITIs were adapted from trial -to -trial and ranged between 1 second and 18.5 seconds in such a way that inter-startle intervals were kept within the range between 25.25 seconds and 33.25 seconds, and were on average similar across conditions (29.25 seconds).

Participants were subjected to one of two task versions which differed with respect to the pseudo-random sequence of conditions. The

conditions were evenly distributed during the experiment to avoid a differential effect of habituation on startle magnitudes for the different conditions. To summarize the number of trials and conditions: conditions were based on win potential (money [56 trials in total] or no money [56 trials in total] at stake), actual outcome (heads [56 trials in total: 28 no money and 28 money trials] or tails [56 trials in total: 28 no money and 28 money trials]) and phase (24 probes during the information phase, 24 probes during the anticipation phase, 48 probes during the feedback phase and 16 non-probed trials).

Data processing

Offline the EMG data were processed using Brain Vision Analyzer software (Brain products, Gilching, Germany), according to the published guidelines (Blumenthal et al., 2005) and studies previously published by our group (e.g. Klumpers et al., 2009). Trials were segmented, filtered with a 28 Hz high pass (24 dB/oct) and 500 Hz low pass (24 dB/oct) filter, rectified, smoothed with an additional 14 Hz (24 dB/oct) low pass filter and baseline-corrected (baseline from 30 ms before startle probe onset until 20 ms after startle probe onset). Peak magnitudes were scored within an interval between 25-120 ms after startle probe onset. Trials with a startle peak latency outside this 25-120 ms window or with EMG activity in the baseline period 2 SD greater than for the other trials in that subject were automatically detected and removed from the data. Trials in which there was less than 55% increase in activity within the above mentioned interval with respect to the 50 ms peri-startle probe baseline period were marked as being a null response (Klumpers et al., 2009), but not set to zero (i.e., the actual values were included in the analyses). Data of seven participants were discarded during the analyses, because more than half of the startle responses were invalid (null response or artifact) in one or more conditions.

Startle magnitudes were subsequently transformed to individual z-scores in order to remove variance due to large between-subject differences in overall startle magnitude. Consistent with previously published procedures (Klumpers et al., 2009; Klumpers, Heitland, Oosting, Kenemans & Baas, 2012), z-scores were based on all individual habituation and experimental startle magnitudes recorded during the experiment. All

further analyses were based on z-transformed data. Z-scores were transformed to T-scores for illustrational purposes (Z*10+50).

Statistical analysis

To explore the effect of reward magnitude and win potential on the time taken to initiate a trial, mean reaction times measured from the onset of the screen indicating that the spacebar can be pressed to the onset of the spacebar press were entered into a mixed-model ANOVA. This ANOVA (SPSS version 20) included win potential (no money, money) as within-subjects factor and reward magnitude (1 cent, 5 cent, 10 cent, 20 cent) as between-subjects factor.

Regarding EMG, for each participant z-scores were averaged across trials for each condition and phase. These values were entered in separate mixed-model ANOVAs (GLM, SPSS version 20) for the trials in which startle probes were delivered before receiving feedback (information phase and anticipation phase) and trials in which the startle probes were delivered after receiving feedback (feedback phase). The ANOVA for the feedback phase (main analysis) included win potential (no money, money) and outcome (not won, won) as within-subjects factors and reward magnitude (1 cent, 5 cent, 10 cent, 20 cent) as between-subjects factors. The ANOVA for the pre-feedback phases (secondary analysis) included phase (information phase, anticipation phase) and win potential (no money, money) as within-subjects factors and reward magnitude (1 cent, 5 cent, 10 cent, 20 cent) as between-subjects factors. The factor task version was initially included in these tests, but, as expected, no main or interaction effects with respect to task version were found for any of the ANOVAs. Therefore, data were collapsed across task versions.

Table 1 shows raw startle magnitudes in microvolts for each condition and each reward magnitude group separately. Z-scores were computed from all individual raw startle values including habituation trials (see Data processing). The across-subject averages for the means and standard deviations upon which z-scores were based amounted to 51.37 μV (sd=40.35) and 32.41 μV (sd=16.87), respectively.

Additional analyses were conducted using standardized startle magnitudes for which habituation trials were excluded from the Z-

transformation. The results of these analyses largely replicated the reported findings[1].

Table 1
Mean raw startle magnitudes (and standard deviations) in microvolts for each condition and habituation phase (before task, after first break, after second break) for each reward magnitude group.

	Reward magnitude group			
	1 cent	5 cent	10 cent	20 cent
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Condition				
Info phase				
No win potential	28.00 (21.84)	55.30 (42.21)	49.54 (54.10)	38.94 (20.51)
Win potential	30.86 (20.47)	56.46 (43.03)	49.00 (51.88)	39.72 (26.66)
Anticipation phase				
No win potential	35.15 (24.28)	63.61 (52.85)	50.72 (49.66)	45.23 (25.58)
Win potential	38.92 (27.27)	76.66 (65.89)	52.30 (47.02)	45.50 (24.99)
Feedback phase				
No win potential - not won	30.07 (23.17)	64.69 (50.95)	47.74 (50.11)	37.92 (26.31)
No win potential - won	27.89 (18.41)	62.64 (49.62)	45.94 (43.22)	39.26 (26.14)
Win potential - not won	30.19 (20.49)	60.65 (45.10)	46.26 (41.69)	44.93 (29.80)
Win potential - won	27.84 (21.21)	64.74 (48.01)	45.19 (50.34)	39.39 (28.22)
Habituation				
Before the task (9 probes)	85.89 (47.58)	124.30 (73.30)	107.46 (80.46)	120.17 (64.03)
After first break (2 probes)	51.19 (42.27)	65.42 (60.09)	62.26 (69.82)	48.69 (36.47)
After second break (2 probes)	34.21 (31.74)	72.50 (64.07)	51.90 (50.32)	36.01 (19.65)

[1] Additional analyses for Experiment 1 with standardized scores omitting habituation trials show that the results only differ (not crucially) for the analysis of the pre-feedback probes. The phase x win potential interaction for the pre-feedback probes was now marginally significant ($p = .062$). The significant main effects of win potential and phase were still present.

Results

Behavioral data

A win potential (no money, money) by reward magnitude (1 cent, 5 cent, 10 cent, 20 cent) mixed-model ANOVA revealed that the amount of time it took participants to press the spacebar was significantly affected by win potential, $F(1,44) = 18.2$, $p < .001$, $\eta_p^2 = .293$. Participants were significantly faster on trials during which money could be won (RT = 1068 ms, SD = 373.7 ms) compared to trials during which no money could be won (RT = 1200 ms, SD = 434.3 ms). There was no significant main or interaction effect of reward magnitude, $F < 1$, all $\eta_p^2 < .047$.

Startle reflex data - Information and anticipation phase

A phase (information phase, anticipation phase) x win potential (no money, money) by reward magnitude (1 cent, 5 cent, 10 cent, 20 cent) ANOVA revealed significant main effects of phase, $F(1,44) = 34.4$, $p < .001$, $\eta_p^2 = .439$ and win potential, $F(1,44) = 10.0$, $p = .003$, $\eta_p^2 = .186$ (see Figure 2, note that for illustrative purposes startle magnitudes are displayed as T-scores). Moreover, there was a significant interaction between phase (information phase, anticipation phase) and win potential, $F(1,44) = 4.3$, $p = .045$, $\eta_p^2 = .088$. Separate T-tests for the information and anticipation phase comparing startle magnitude when money could be won and could not be won revealed only a significant difference between win potential conditions during the anticipation phase, $t(47) = -3.1$, $p = .004$, but not during the information phase, $t(47) = -0.6$, n.s. No significant interactions with the between-subjects factor reward magnitude were found, win potential x reward magnitude: $p = .061$, $\eta_p^2 = .153$; all other interactions with reward magnitude: $p > .289$, $\eta_p^2 < .081$. There was also no correlation between the effect of win potential on startle magnitude during the anticipation phase and the effect of win potential on time to initiate the trial, $r = .172$, $p = .244$.

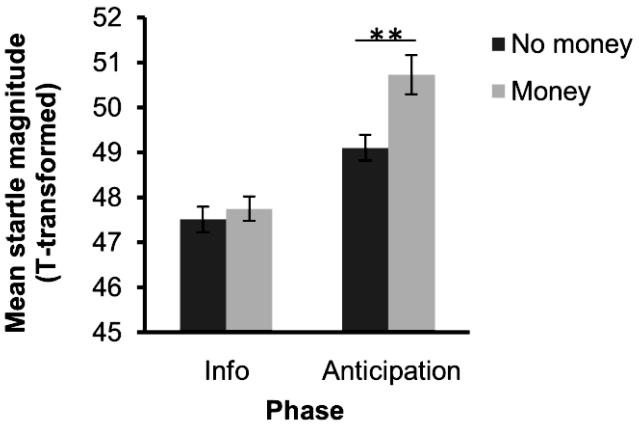


Figure 2. Mean startle magnitudes during the information phase and anticipation phase. Dark grey bars represent startle magnitudes in trials where no money could be won. Light grey bars represent startle magnitudes when there was a potential to win money. Startle magnitude was significantly increased when money could be won, compared to when no money could be won during the anticipation phase. The interaction between phase and win potential was significant ($p < .05$). Error bars represent ± standard error of the mean. ** $p < .01$. Note that z- scores are transformed to T-scores for illustrative purposes.

Startle reflex data - Feedback phase

In the feedback phase, there was a significant three-way interaction between win potential (no money, money), outcome (not won, won) and reward magnitude, $F(3,44) = 3.0$, $p = .041$, $\eta_p^2 = .169$. Interactions between win potential and outcome for all reward magnitude groups are presented in Figure 3. Note that for illustrative purposes startle magnitudes are displayed as T-scores. Close inspection of Figure 3 suggests that the three-way interaction was driven by an aberrant pattern in the 5-cents condition. The two-way interaction pattern in the 5-cents condition differed much more from the other three conditions than any of the other three differed from the rest, even to the extent that the sign of the interaction was opposite for the 5-cents group relative to the other three groups. The win potential x reward outcome x reward magnitude interaction was, therefore, again tested without the inclusion of the 5-cents group. In this ANOVA with only three groups (1 cent, 10 cent, 20 cent) the three-way interaction was no longer significant, $F(2,33) < 1$. The win potential x outcome interaction (i.e.,

independent of reward magnitude) was, however, significant, $F(1,33) = 5.24$, $p = .029$, $\eta_p^2 = .137$. Breaking down the win potential x outcome interaction across the three groups, in the trials in which money could be won (the 'money' condition), startle magnitude was significantly larger when the actual outcome was 'not won' than when the outcome was 'won', $t(35) = 3.241$, $p = .003$. In contrast, when no money could be won anyway (the 'no money' condition), there was no significant difference in startle eye blink reflex between both outcomes, $t(35) = .271$, $p = .788$. Furthermore, there was no significant difference in startle magnitude between the two win potential conditions when the outcome was "not won", $t(35) = -1.567$, $p = .126$, nor was there such a difference when the outcome was "won", $t(35) = 1.533$, $p = .134$.

Although there was no significant interaction with reward magnitude in the test with three groups included, Figure 3 suggests a particularly outspoken win potential by outcome interaction for the 20-cents group compared to the other two groups. Furthermore, particularly in the 20-cents condition startle modulation seems driven by potentiation after not winning rather than by inhibition after winning. This pattern was explored in a post-hoc analysis testing the win potential x outcome interaction for the 20-cents group only.

The post-hoc analysis revealed a significant win potential x outcome interaction for the 20-cents group, $F(1,11) = 10.0$, $p = .009$, $\eta_p^2 = .477$. Breaking down this interaction in the 20-cents group, in the trials in which money could be won, startle magnitude was significantly larger when the actual outcome was 'not won' than when the outcome was 'won', $t(11) = 2.5$, $p = .031$. In contrast, when no money could be won, there was no significant difference in startle eye blink reflex between both outcomes, $t(11) = -1.2$, n.s. Moreover, when comparing the difference in startle-reflex magnitude between the money and no money potential condition for each outcome, there was a significant difference in startle magnitude only when the outcome was negative (not won), $t(11) = -3.0$, $p = .012$ and there was no significant difference when the trial was won, $t(11) = 0.4$, n.s.

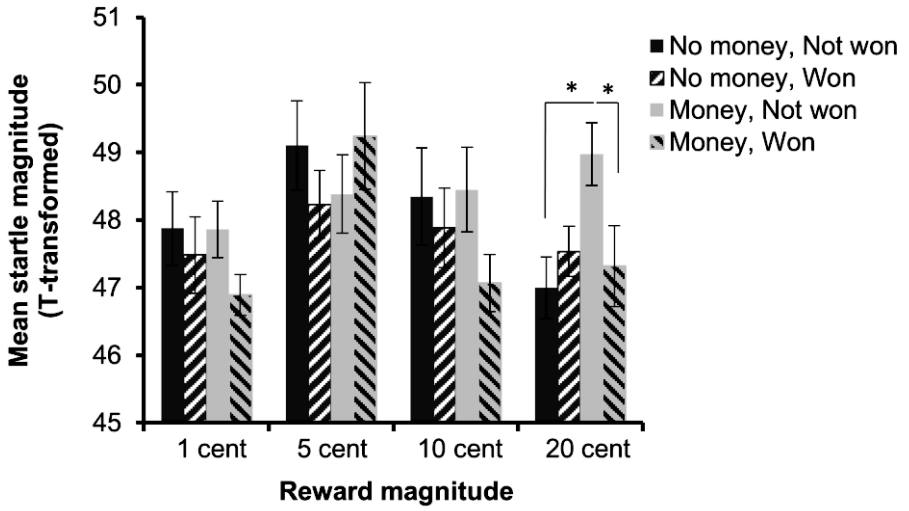


Figure 3. Bar graph showing the interaction between win potential and outcome on startle magnitudes for each reward magnitude group in the feedback phase of the task. Dark grey bars represent no money potential and light grey bars represent potential to win money. Striped bars represent the outcome "won". Across the 1 cent, 10 cent and 20 cent condition, there was a significant interaction between potential to win money and the actual outcome of the trial. Startle magnitude was significantly increased when money was at stake and participants did not win compared to when they did win. In contrast, there was no such difference when no money was at stake. Note that the 5-cents group was omitted from further analysis, because of the aberrant interaction pattern for this condition that likely drove the initial interaction between reward magnitude, win potential and outcome. Error bars represent ± standard error of the mean. Note that z-scores are transformed to T-scores for illustrative purposes.

Discussion Experiment 1

The main goal of the first experiment was to investigate appetitive and aversive motivation after feedback about whether a potential to win money had indeed materialized, relative to a baseline without potential gain. Startle blinks during the feedback phase were significantly modulated by feedback about whether the coin landed on heads (not won) or tails (won), as expected. In trials in which money was at stake, the startle reflex was potentiated when the monetary reward was not won compared to when it was won. This effect was observed across three reward-magnitude conditions (1, 10, and 20-cents per trial), but not, however, in the 5-cents

condition. The aberrant pattern in the 5-cents condition may reflect a sampling error or the effect of some unknown characteristic of this condition.

The observation that the startle reflex was relatively potentiated for not winning compared to winning a reward across the 1, 10, and 20-cents groups is consistent with findings of prior studies (Hackley et al., 2009; Skolnick & Davidson, 2002). In contrast to these studies, the present study used an orthogonal design including neutral trials in which subjects knew they would not win money anyway (no money/ won and no money/ not won). This allowed us to explicitly test the difference between the money/ not won condition versus the no money/ not won condition, and the difference between the money/ won versus the no money/ won condition. We found that startles, and hence withdrawal motivation, were enhanced when money was anticipated but not obtained, relative to when anticipated and obtained, and, in the 20-cents condition only, also relative to when not anticipated and not obtained (i.e., 0-cents not obtained). However, one of our conditions (5-cents) did not at all reveal this pattern. To find out whether this could be due to a sampling error was one of the aims of a follow-up experiment.

In all, these findings support the idea that particularly for the 20-cents condition a startle-reflex difference between winning and not winning may be better explained in terms of potentiation when a potential reward does not materialize, rather than in terms of inhibition as a result of the positive state. To the extent that missing out on a potential appealing reward can be considered as experiencing a loss, this outcome may be related to Kahneman and Tversky's (1979) prospect theory, which states that the pain of a loss is experienced more intensely than the pleasure of a gain. In addition, the stronger modulation in the 20-cents relative to the other reward-magnitude conditions, indicates that small rewards typically used in reward experiments may not be subjectively equivalent in value, which is relevant to future studies using monetary reinforcement. More specifically, a 20-cents reward may have higher subjective value on average than the lower-magnitude rewards, and hence, the "pain" of missing out on a 20-cents reward may be experienced more intensely than that of a lower-magnitude reward. This may also apply to other and larger rewards. However, it should be noted that stronger startle modulation for the 20-

cents condition compared to the other conditions could not be confirmed statistically.

A secondary aim was to replicate the results of prior studies showing potentiation of the startle reflex during reward anticipation. As expected, startle-reflex magnitudes in the anticipation phase were enhanced when a reward was at stake compared to when no reward was at stake. The observation that the startle reflex was enhanced when subjects anticipated a reward compared to the anticipation of no reward is consistent with a large body of literature showing that the anticipation of both appetitive and aversive pictures (Dichter et al., 2002; Lipp et al., 2001; Nitschke et al., 2002; Sabatinelli et al., 2001), as well as potential rewards (Skolnick & Davidson, 2002) enhances the startle reflex. These results and the results of the current study are in keeping with the assumption that the startle reflex during anticipation of an emotional event mainly (but not exclusively, Nitschke et al., 2002) reflects anticipatory arousal (Sabatinelli et al., 2001).

Startle response magnitudes did not differ between the money and no money condition during the information phase, when participants just saw whether they could win money or not. This result may fit with an observation by Sege and colleagues (Sege, Bradley & Lang, 2014) who showed that startle potentiation during anticipation of positive slides was only present late in the anticipation phase. Alternatively, the video of the spinning coin with a predictable duration and anticipation of feedback regarding the reward that comes right after the end of the video may have increased physiological reactivity during the anticipation phase.

Experiment 2

Introduction

In Experiment 1 we aimed at comparing startle-reflex modulation between rewards of varying magnitude and found that this was most pronounced for the highest reward level (20 cents). However, this could only be confirmed statistically in an a-posteriori analysis for the reward magnitude conditions separately. The direct comparison between startle-reflex modulation for large compared to small rewards was not significant. One interpretation is

that, with the presently used between-subjects manipulation of reward magnitude, rewards of different magnitudes all induced the same state because each reward was the maximum attainable reward (relative to zero reward) for that subject; this could be an instance of adaptive scaling (Walsh & Anderson, 2012). Other evidence also suggests that reward value is scaled by the context of other reward options in an experiment (Nieuwenhuis et al., 2005). The absence of a context of varying reward magnitudes for individuals in the different reward groups may have led to the absence of robust differences in valuation across the reward conditions. Therefore, in the follow-up experiment, we directly compared the 20-cents condition with one of the lower reward conditions (5-cents) within-subjects.

In the follow-up experiment with new participants we aimed to confirm the results of the a-posteriori analysis on reward feedback separately for the different reward magnitude conditions, and furthermore we aimed to re-investigate the pattern of startle-reflex modulation for the 5-cents condition. In this follow-up experiment we used a very similar task in a larger number of subjects. The within-subjects design of Experiment 2 allowed us to investigate whether startle modulation is affected by relative reward magnitude. Based on the results of Experiment 1, we expected to find larger startle magnitudes after not winning a reward compared to neutral trials and trials during which a reward was won. We expected this effect to be stronger for the 20-cents compared to the 5-cents condition, based on Experiment 1 and because of the hypothesized scaling of reward value by the context of the other rewards. More specifically, the 5-cents reward was expected to have less subjective value, because even a larger reward could have been won or lost (Goldstein et al., 2006; Yeung & Sanfey, 2004; Knutson, Adams, Fong & Hommer, 2001; Nieuwenhuis et al., 2005; Smith et al., 2010). We also expected larger startle magnitudes during the anticipation of 5 and 20 cents compared to the anticipation of no reward, replicating our first experiment.

Methods

Participants

Twenty-six new and naive subjects were recruited via advertisement at the campus of Utrecht University and via advertisement on social media. The

in/exclusion criteria were the same as in Experiment 1. Participants were asked to refrain from caffeine and smoking on the day of the experiment. Written informed consent was obtained and this study was approved by the local ethical committee of the Faculty of Social Sciences of Utrecht University. Subjects received a financial compensation of 14.50 Euros or study credits. Subjects additionally received the money won during the task (fixed at 5.25 Euros, which was unknown to the subjects during the task). Two subjects were excluded because they had too few valid startle responses in one or more conditions. Therefore, the final sample consisted of 24 subjects (4 males, mean age ± sd: 22.6 ± 4 years).

Stimuli and psychophysiological recordings

The startle stimuli and the psychophysiological recording procedures were the same as in Experiment 1.

Task and procedure

Subject eligible for participation signed the informed consent form and subsequently filled in the BIS/BAS questionnaire. The procedure was the same as in Experiment 1, unless stated otherwise. The head or tails paradigm was largely similar to the one used in Experiment 1. As in Experiment 1 a trial consisted of three phases: an information phase, anticipation phase (coin spinning), and a feedback phase. However, unlike in Experiment 1, startle probes were only presented during the anticipation and feedback phase, as startle modulation was only observed during these phases in Experiment 1. The task consisted of 126 trials, of which 108 with a startle probe. Thirty-six startle probes were delivered during the anticipation phase and 72 were delivered during the feedback phase, of which 36 following a positive outcome (tails) and 36 following a negative outcome (heads). Unlike in Experiment 1, reward magnitude was manipulated within-subjects. The amount of money at stake (either 0, 5 cents, or 20 cents) as well as the outcome in terms of heads (not won) and tails (won) were counterbalanced across trials. Outcome was evenly distributed between trials in which 0 cents, 5 cents, or 20 cents was at stake.

The timing of startle probe delivery, the duration of the events on the screen, and the average duration of the ITIs were the same as in

Experiment 1. Three rest-breaks (self-paced) were provided after, respectively, 25, 50, and 75 percent of the trials.

Data processing and statistical analyses

The data processing steps in Brainvision analyzer were as described before. The raw startle magnitudes in microvolts were transformed to Z-scores based on all individual experimental startle magnitudes excluding the habituation trials. These standardized values were averaged across trials for each condition and phase.

With respect to the EMG data, a repeated-measures ANOVA with the factor win potential (0, 5, 20 cents) was run for the anticipation phase. A win potential (0, 5, 20 cents) x outcome (not won, won) repeated-measures ANOVA was run for the feedback phase. Reaction time data were entered into a repeated-measures ANOVA with the within-subjects factor win potential (0, 5, 20 cents).

Results

Behavioral data

Reaction times were marginally significantly affected by the amount of money at stake, $F(1.5,34.8) = 3.26$, $p = .063$ (Greenhouse-Geisser corrected). Reaction times tended to become shorter with increasing amounts of money at stake (average RTs were 836, 798, and 784 ms, for the 0, 5, and 20-cents conditions, respectively).

Startle reflex data - anticipation phase

Startle-reflex magnitudes during the anticipation phase were not significantly affected by the amount of money at stake, $F(2,46) = 1.2$, $p = .309$. The pattern of results was, however, in the expected direction. That is, magnitudes tended to be larger with increasing amounts of reward at stake. See Figure 4.

Startle reflex data - feedback phase

The win potential x outcome ANOVA for startle magnitudes during the feedback phase revealed a significant main effect of win potential, $F(2,46) = 4.27$, $p = .02$, $\eta_p^2 = .156$. See Figure 5. Follow-up T-tests indicated that the

startle magnitudes across both outcomes were larger for the 5 compared to 0-cents condition, t(23) = -3.20, p = .004, and marginally larger for the 20 compared to 0-cents condition, t(23) = -1.91, p = .068. There was no significant interaction with outcome.

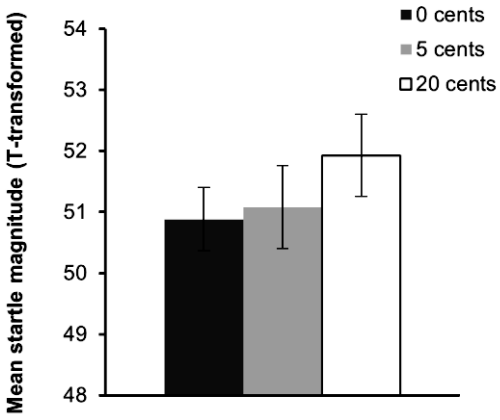


Figure 4. Mean startle magnitudes during the anticipation phase of the follow-up study. Startle magnitudes tended to be increased with increasing reward at stake. This effect was, however, not significant. Error bars represent ± standard error of the mean. Note that z-scores are transformed to T-scores for illustrative purposes.

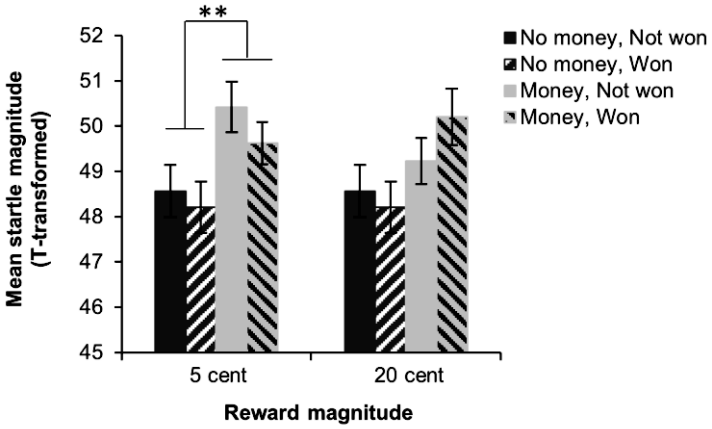


Figure 5. Mean startle magnitudes during the feedback phase of Experiment 2. Startle magnitudes during the feedback phase were differentially affected by the potential to win a reward, irrespective of positive/negative outcome (main effect of win potential, p=.02). This effect was driven by startle magnitudes across the won/ not won outcome being larger in the 5-cents compared to the 0-cents condition. The win potential conditions (0 cent, 5 cent, and 20 cent) were manipulated within-subjects. Error bars represent ± standard error of the mean. Note that z-scores are transformed to T-scores for illustrative purposes. Note also that the outcomes in the 0-cents condition (no win potential) are in the Figure twice, visualized as "No money, Not won" and "No money, won". This was done in order to make this Figure easier to compare with Figure 3.

Discussion Experiment 2

The main aim of this additional experiment was to replicate the pattern of startle modulation that we observed for the 20-cents condition in Experiment 1. Specifically, we investigated whether startle modulation after outcome feedback is mainly driven by startle potentiation when a relatively large reward does not materialize. The results for the feedback phase of Experiment 2 were, however, inconclusive with respect to this question.

Startle magnitudes during the feedback phase of Experiment 2 were larger for trials in which a 5-cents reward (and to a lesser degree a 20-cents reward) was at stake compared to trials without win potential, irrespective of trial outcome. This finding is puzzling and is clearly different from what we observed for the feedback phase in Experiment 1, in which the effect of win potential was dependent on the outcome of the trial. The pattern of startle modulation in the feedback phase of Experiment 2 seems more comparable to the pattern of startle modulation during the anticipation phase of Experiment 1. It is not clear why the findings of Experiment 2 are so different from those of Experiment 1. Though the task is very similar to the task we used in Experiment 1, a few differences may perhaps have contributed to these discrepancies. A first difference is that the rewards at stake were manipulated between subjects in Experiment 1, and manipulated within subjects in Experiment 2. In Experiment 1 the only reward options that an individual subject encountered were 0 cents or a reward of fixed magnitude, whereas in Experiment 2 the trial types were more diverse (0 cents, 5 cents or 20 cents). Anecdotally, subjects reported after the experiment that they did not always remember what the reward condition was by the time the feedback came. The inclusion of multiple reward conditions in Experiment 2 also led to differences in the reward potential : no reward potential ratio, which was 50 : 50 in Experiment 1 and 66 : 33 in Experiment 2. It may be speculated that the relatively high reward potential : no reward potential ratio in Experiment 2 caused less news-value for trials with reward potential, possibly leading to a loss of motivation or reduced attention to the reward information screens. This possibility is supported by the reaction time data showing that in Experiment 2 subjects were only marginally faster on potential reward compared to no reward trials, and with the startle data during the anticipation phase indicating that

startle magnitudes were not significantly increased when a reward was at stake. To the extent that startle during the anticipation phase reflects arousal, this suggests that the anticipation of reward did not lead to increased arousal relative to the anticipation of no reward.

The possible loss of motivation to obtain rewards, reduced attention to reward information, and the increased complexity of the task may explain why there was no significant interaction between reward potential and trial outcome for startle responses during the feedback phase in Experiment 2. These factors do, however, not explain why startle responses in the feedback phase were significantly affected by reward potential, irrespective of trial outcome, in a manner similar to the pattern observed for the anticipation phase of Experiment 1.

A second difference between the experiments was the moment of startle probe delivery. Startle probes in Experiment 2 were delivered either during the anticipation or the feedback phase, whereas in Experiment 1 they could also be delivered during the information phase. It is, however, unclear whether, and if so how, this difference between the two tasks contributes to the discrepancies between the results.

Finally, the results of Experiment 2 do not support our hypothesis based on the results of Experiment 1 that the difference in the startle-reflex magnitude between winning and not winning is driven by potentiation when a potential reward does not materialize, rather than by startle inhibition as a result of the positive state.

General discussion

The main goal of this study was to investigate whether motivational tendencies after reward feedback, as indexed by the startle reflex, are mainly characterized by appetitive responses to a (potential) gain or mainly by aversive consequences of reward omission, or by both. Another main aim was to use the startle reflex as a readout measure of motivational state during the obtainment and loss of rewards with varying magnitudes in order to compare the average subjective value of these rewards. This was studied by comparing startle modulation during positive and negative outcome feedback relative to baseline conditions without potential gain,

and by comparing startle-reflex modulation between different monetary reward conditions.

In Experiment 1 we found that startle during the feedback stage was sensitive to loss, in that not winning money relative to winning potentiated the startle reflex across the 1, 10, and 20-cents conditions. Specifically for the largest reward (20-cents) the difference in startle-reflex magnitude between winning and not winning appeared to be mostly driven by potentiation when a potential reward does not materialize, rather than by inhibition as a result of a positive outcome. Additionally, as expected based on previous studies that show enhanced startle in conditions of emotional arousal, we observed startle potentiation during the anticipation phase when a reward was at stake.

However, neither of these results were replicated in our follow-up study. In Experiment 2, the main effect of reward potential that we expected during the anticipation phase was significant for the feedback phase. The non-replication of the startle potentiation for not won trials when a reward was at stake in Experiment 2 precludes drawing a conclusion regarding the main question about whether startle is mostly affected by the not won or by the won trials.

It almost seems that in the second experiment the feedback phase was the first time that subjects actually noticed what was at stake. Why startle during feedback shows this effect of win potential irrespective of trial outcome rather than the previously published pattern (Hackley et al., 2009; Skolnick & Davidson, 2002, partly replicated in Experiment 1) is subject of speculation. Several changes in the experimental set-up (discussed above) cannot satisfactorily explain the discrepancies in results between the two experiments. Apparently, even though Experiment 2 was set-up as a replication of Experiment 1, these differences in the context in which the different conditions were experienced made the set-up less sensitive to reward effects at the feedback stage. The possibility that differences in motivation to obtain rewards contributed to the observed discrepancies between the results is supported by the reaction time data showing that subjects in Experiment 1 were initially faster to spin the virtual coin as soon as they knew that money was at stake. This reaction time effect was, however, not significant in Experiment 2, suggesting that the appetitive tendency was relatively reduced in Experiment 2.

Some changes based on the apparent reduction in appetitive motivation in Experiment 2 can be recommended to increase subjects' motivation and task engagement. We recommend leaving the reward value of the current trial on-screen during the course of the trial. This prevents that subjects forget the reward magnitude that is at stake. We also recommend emphasizing before the experiment that subjects can take home the money that they win during the task. Furthermore, future studies may want to equalize the reward : no reward potential ratio, as this may have been a factor contributing to the discrepant results obtained in Experiment 1 and 2.

It should be noted that there is a possibility that the non-replication was due to the findings of Experiment 1 being chance effects rather than genuine effects. This alternative scenario is however less likely, because the pattern of results obtained in Experiment 1 (i.e., increased startle responses when anticipating a reward and increased startle responses for not winning compared to winning) was also observed in other studies (Hackley et al., 2009; Skolnick & Davidson, 2002).

In conclusion, the finding that the startle reflex is potentiated during the anticipation of a reward compared to no reward seems relatively robust. The difference in startle-reflex magnitudes between winning and not winning that we observed in Experiment 1 has been observed before in prior studies. The failure to replicate this effect in Experiment 2 may have been due to differences in motivation or task engagement. It could therefore be useful to repeat Experiment 2 with the proposed changes to the setup and procedure aimed at tackling the potential issue of reduced motivation. Perhaps this new experiment sheds more light on the question whether startle-reflex modulation after feedback is mainly characterized by appetitive responses to a gain or mainly by aversive consequences of reward omission. The data of the two experiments do not provide evidence for systematic increases in startle modulation with increased reward magnitude at stake. We therefore do not recommend using startle-reflex modulation as a means to assess the subjective value of rewards.

Chapter 3

Stimulus discriminability may bias value-based probabilistic learning

Iris Schutte
Heleen A. Slagter
Anne G.E. Collins
Michael J. Frank
J. Leon Kenemans

PloS one (2017)

Author contributions

LK, MF, IS, and HS were involved in study conceptualization; IS and LK set up Experiment 1 and 2; MF and AC set up Experiment 3; IS programmed the tasks and collected the data of Experiments 1a, 1b and 2; AC collected the data of Experiment 3; IS, AC, and LK were involved in analyzing the data; IS wrote the initial draft of the manuscript (Experiment 1 and 2); AC wrote the initial draft of Experiment 3; all authors contributed to the final manuscript; LK supervised the study.

Abstract

Reinforcement learning tasks are often used to assess participants' tendency to learn more from the positive or more from the negative consequences of one's action. However, this assessment often requires comparison in learning performance across different task conditions, which may differ in the relative salience or discriminability of the stimuli associated with more and less rewarding outcomes, respectively. To address this issue, in a first set of studies, participants were subjected to two versions of a common probabilistic learning task. The two versions differed with respect to the stimulus (Hiragana) characters associated with reward probability. The assignment of character to reward probability was fixed within version but reversed between versions. We found that performance was highly influenced by task version, which could be explained by the relative perceptual discriminability of characters assigned to high or low reward probabilities, as assessed by a separate discrimination experiment. Participants were more reliable in selecting rewarding characters that were more discriminable, leading to differences in learning curves and their sensitivity to reward probability. This difference in experienced reinforcement history was accompanied by performance biases in a test phase assessing ability to learn from positive vs. negative outcomes. In a subsequent large-scale web-based experiment, this impact of task version on learning and test measures was replicated and extended. Collectively, these findings imply a key role for perceptual factors in guiding reward learning and underscore the need to control stimulus discriminability when making inferences about individual differences in reinforcement learning.

Introduction

Reinforcement learning refers to the ability of humans and other animals to learn from the outcome of their actions. Actions that lead to a positive outcome are likely to occur more frequently in the future than actions that yield punishment (Thorndike, 1927). Motivated or reward-driven behavior is thought to reflect the balance between contributions of different brain systems involved in reward processing and punishment processing, respectively (Ernst, Pine, & Hardin, 2006). There is great individual variation in reward sensitivity as well as in sensitivity to signals of punishment. In other words, humans differ in motivational style.

Reinforcement learning tasks are used to assess motivational style often by testing individual differences in reinforcement learning overall (Schonberg, Daw, Joel, & O'Doherty, 2007) or whether one is better at choosing stimuli that likely lead to reward or better at avoiding stimuli that likely lead to punishment (e.g., Bodi et al., 2009; Frank, Seeberger, & O'reilly, 2004; Maril, Hassin-Baer, Cohen, & Tomer, 2013; Palminteri et al., 2009; Piray et al., 2014). The probabilistic selection task (PST; Frank et al., 2004) has been used many times to assess effects of patient populations and dopamine medications on learning from positive vs. negative feedback, and individual differences in these measures have been repeatedly related to neural and genetic measures associated with striatal and dopamine function (e.g., Cavanagh, Bismark, Frank, & Allen, 2011; Cockburn, Collins, & Frank, 2014; Cox et al., 2015; Doll, Hutchison, & Frank, 2011; Frank et al., 2004; Frank, Samanta, Moustafa, & Sherman, 2007; Frank, Woroch, & Curran, 2005; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Gründler, Cavanagh, Figueroa, Frank, & Allen, 2009; Jocham, Klein, & Ullsperger, 2011; Klein et al., 2007; Waltz, Frank, Robinson, & Gold, 2007).

In the PST, individuals first learn to choose which stimulus characters of three different stimulus pairs are most likely to be rewarded. Stimulus A is associated with positive feedback on 80% of AB pair trials whereas stimulus B is associated with negative feedback on 80% of AB trials. Outcome contingencies within the other pairs (CD and EF) are less consistent. Choosing C and D is followed by positive and negative feedback, respectively, in 70% of the trials. E and F are paired with positive and negative outcomes, respectively, in 60% of the trials. During the test phase

all possible combinations of training stimuli are presented in pairs. Subjects who more often avoid stimulus B in novel test pairs during the test phase, relative to choosing A, are classified as negative learners. In contrast, subjects who more often choose stimulus A in novel test pairs during the test phase, relative to avoiding B, are classified as positive learners. These learning biases have been related to individual differences in event-related potential (ERP) and neuroimaging components of feedback processing, dopaminergic medication status and genetic polymorphisms associated with dopamine function (see references above). These observations are in line with the notion that performance on the PST and other RL tasks captures effects of manipulations (e.g. drugs) and individual differences in motivational style related to the reward valuation system itself.

However, thus far the modern reinforcement learning literature has not investigated whether overall learning – and especially positive vs. negative learning – may depend on perceptual aspects of the stimuli themselves. An important aspect is discriminability, defined here as the extent to which a stimulus stands out from the background and from other stimuli due to its distinctive global and/or local features (e.g. shape or color). For example, stimuli that are perceptually more discriminable than others may be easier for participants to remember to select once they have been rewarded. Salience or discriminability may also affect preference (ratings of trustworthiness of faces; Kiss et al., 2007). Furthermore, it has been shown that discriminability and preference interact to affect free choice in a classification task (Imai & Garner, 1965). In addition, there is a collection of older literature describing how learning of the affective value of a stimulus is modulated by its salience. This holds in particular for salience as conveyed by the relative novelty or unfamiliarity of the stimulus. The relation between salience and the extent to which a stimulus affords learning has become manifest in classic phenomena such as blocking (e.g., Schmajuk & DiCarlo, 1992) and latent inhibition (Lubow, 1973). It is reflected in incidental learning about salient relative to non-salient items as described already in 1933 by von Restorff (Von Restorff, 1933); see also Fabiani, Karis, & Donchin, 1986, and in relatively strong physiological orienting reactions to novel stimuli predicting subsequent affective learning about these stimuli (Grillon, 2002; Maltzman, 1979).

Also recent research indicates that our choice between alternatives is not only affected by reward value but also by stimulus salience, and that perceptual and reward properties may influence decision making via common neural mechanisms (Gold, Waltz, Prentice, Morris, & Heerey, 2008; Lou, Hsu, & Sajda, 2015; Navalpakkam, Koch, Rangel, & Perona, 2010; Polanía, Krajbich, Grueschow, & Ruff, 2014; Smith & Ratcliff, 2004). Indeed, Cavanagh et al. (Cavanagh, Wiecki, Kochar, & Frank, 2014) have shown in the test phase of the PST that eye gaze time toward a stimulus increases the chances of selecting it, regardless of its value. This body of research mainly pertains to conditions in which reward contingencies have already been learned but salience might also influence learning about reward itself. The latter is not only of profound theoretical significance by itself but also a potential concern as assessment of motivational style often requires comparison of learning bias across different task conditions, which may differ in the relative salience or discriminability of the stimuli associated with more and less rewarding outcomes, respectively.

The aim of the current study was therefore to investigate if classification as a positive or negative learner in the PST task solely reflects variation in motivational style, or may also reflect variation in the extent to which aspects of the stimulus configuration afford learning from the stimuli in general. To this end, in two experiments (Experiment 1a and 1b), which only differed in the time available to respond, participants performed one of two versions of the PST. These versions differed in assignment of physical characters to the functional A-F categories as discussed above. We used Hiragana letter stimuli (see Figure 1) as typically used in the PST as described by Frank et al. (Frank et al., 2004; Frank et al., 2005). For both task versions we used the same stimuli and a fixed Hiragana- A-F mapping. However, critically, in the second version, compared to the first, Hiragana letters were switched within the AB, CD and EF pairs. Given prior work suggesting that the PST task is sensitive to individual differences in learning bias that reliably relates to genetic and neural measures when randomizing or counterbalancing stimulus assignments, we expected to find no effect of specific fixed assignments on the proportion of positive and negative learners across versions. However, to foreshadow our results, in both experiments, we found that stimulus-to-feedback mapping strongly affected the learning curves of the individual stimuli and subsequent positive/

negative classification in the test phase. Moreover, a subsequent experiment showed that some Hiragana stimuli were more discriminable (here referring to the extent to which these stimuli stood out from the environment specifically because of their shape) than others. This indicates that for any arbitrary configuration of stimuli and stimulus-contingency mappings, classification of reinforcement learning does not depend only on individual differences in positive and negative learning mechanisms themselves, but can also be affected by variation in the extent to which stimuli afford learning in general. This issue concerns not only the specific probabilistic learning task addressed presently, but also other varieties typically using a single set of stimuli for each condition such as those showing similar effects of dopaminergic manipulations on learning from reward or punishment (Bodi et al., 2009; Palminteri et al., 2009; Piray et al., 2014). Furthermore, this issue may be especially problematic when making inferences about positive/negative learning on the individual level. It should, however, be emphasized that the studies cited above using the PST have all randomized and counterbalanced the stimulus-to-feedback mappings and hence these reported findings cannot be attributed to differences in discriminability between the stimuli.

Indeed, a key prerequisite for interpreting test phase choices in the PST is that feedback experiences during the training phase are similar across subjects (Frank et al., 2004). That is, assessment of a given subject/group's ability to generalize positive /negative learning about A/B in the test phase requires that this subject/group had experienced comparable positive and negative associations for these stimuli during learning compared to other subjects/groups. We therefore additionally compared performance during the training phase between both task version groups. We found that task version had strong effects on learning curves, such that subjects were more reliable in selecting rewarding characters that were more discriminable, leading to differences in experienced reward probability for the critical items across versions. Furthermore, to replicate the findings from Experiment 1, the effect of stimulus-to-feedback mapping on feedback learning was additionally investigated in a large-scale web-based experiment (Experiment 3). Experiment 3 furthermore addressed the possibility that feedback learning primarily depends on the extent to which salient stimuli uniquely stand out

relative to non-salient stimuli in conveying either positive or negative feedback.

Methods – Experiments 1a and 1b

Participants

Thirty-one subjects participated in Experiment 1a and 25 new subjects participated in Experiment 1b. All subjects were recruited through posters at the campus of Utrecht University and were given a financial compensation of 3 Euro per half an hour or study credits. All subjects declared to have normal or corrected-to-normal vision and all subjects were unaware of the aim of the study. Participants were asked to refrain from caffeine use and smoking on the day of the experiment. Exclusion criteria were a history of psychological or neurological disorders, age below 18 years, and caffeine use or smoking at the day of and before the experiment. This experiment was approved by the local ethics advisory board of the Faculty of Social Sciences of Utrecht University. Written informed consent was obtained and participants were treated according to the Declaration of Helsinki.

Nine participants in Experiment 1a were discarded during analysis because they did not satisfy performance criteria during the learning phase (see section task and procedure). The final sample consisted of 22 participants (Experiment 1a). Half of them was assigned to version 1 of the probabilistic learning task (mean age: 22.6, SD: 3.0, range: 19.1-27.8 years, 9 females) and the other half to version 2 of the probabilistic learning task (mean age: 23.0, SD: 2.3, range: 19.7-26.6 years, 9 females). The two task version groups were matched for age and gender.

Of the 25 participants in Experiment 1b, five were discarded during analysis because they did not satisfy performance criteria during the learning phase and/or test phase (see section task and procedure). The final sample of Experiment 1b consisted of 20 participants. Half of them finished version 1 of the probabilistic learning task (mean age: 20.7, SD: 1.3, range: 18.5-23.0, 8 females) and the other half finished version 2 (mean age: 21.8, SD: 2.0, range: 19.2-24.9, 8 females). The two task version groups were matched for age and gender.

Task and procedure

Experiment 1a

Upon arrival at the lab, subjects received information on the procedure of the experiment and written informed consent was obtained. Participants were seated in a chair 85 cm in front of a computer screen. Presentation of instructions and stimuli was controlled by Presentation® software (version 16.0, www.neurobs.com). Participants were subjected to a slightly adapted version of the probabilistic learning task as described in the article by (Frank et al., 2005). For a subset of participants EEG was recorded during task performance (not discussed in this article).

Participants were assigned to one of two task versions (1 and 2) which differed in assignment of Hiragana characters to elements A to F. Figure 1 represents the complete mapping of specific Hiragana characters on A to F for the two task versions. Note that in version 2, relative to version 1, Hiragana-A-F mapping was only switched within pairs (the pairs being AB, CD, and EF). Letter stimuli within a frame of 3° by 3° were white on a black background. One character of each pair was presented on the left side of the screen and the other character on the right side of the screen (center-to-fixation distances were 5 degrees of arc). The left or right character was selected by pressing the "z" or "m" key, respectively, on a qwerty keyboard. The position of each Hiragana stimulus was pseudo-randomized across trials.

The task consisted of two phases, a learning phase and a test phase. During the learning phase, subjects were exposed to a maximum of seven training blocks each consisting of a random sequence of 10 repetitions of each of the six Hiragana combinations (60 trials in total). Each trial consisted of a fixation cross (a white plus sign) presented for a random duration between 250 and 750 ms, a pair of Hiragana stimuli presented for 750 ms, followed by a blank screen for 250 ms, followed by visual feedback presented for 600 ms. On each trial subjects chose one of the stimuli and feedback was presented regarding their choice. If participants did not respond within 1000 ms, the message "no response made" was presented. The feedback stimulus was either a circle to signify that the choice had been 'correct' (positive feedback), or a triangle for a choice that had been 'incorrect.

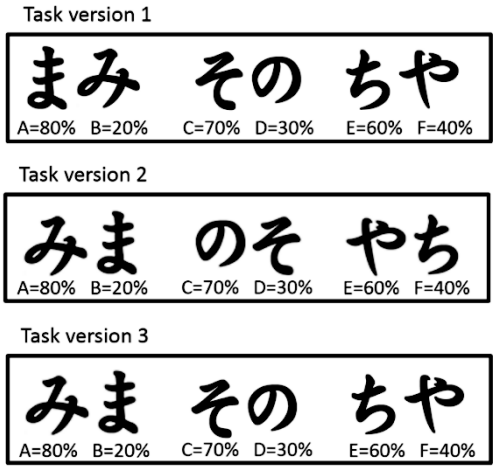


Figure 1. Hiragana stimuli used in the probabilistic learning task in Experiments 1a and 1b and Experiment 3. Each pair of stimuli was randomly presented in separate trials. During each trial participants chose one of the stimuli of the pair. Feedback following participant's choice was determined probabilistically. Reward probability (indicated below each stimulus) differed between characters. Task version 1 and 2 differed with respect to the characters associated with more probable positive and more probable negative feedback, respectively. Specifically, in task version 2, Hiragana stimuli were switched within the AB, CD and EF pair. Half of the participants were subjected to version 1 and the other half to task version 2. In Experiment 3, a task version was used for which only the Hiragana stimuli within the AB pair were switched.

Choosing the A stimulus resulted in positive feedback on 80% of the AB trials, and in negative feedback on 20% of the trials. For choosing the B stimulus these percentages were 20 % for positive and 80 % for negative feedback. The percentages for positive and negative were 70 and 30 for choosing C, 30 and 70 for choosing D, 60 and 40 for choosing E, and 40 and 60 for choosing F, respectively. Participants were told in advance that feedback would be probabilistic (the literal wording being 'possibly incorrect') and that they should select the stimuli that were most likely to be rewarded.

Before advancing to the test phase, subjects needed to meet a performance criterion. During a training block participants needed to choose at least 65% A, 60% C and 50% E stimuli (percentage of the total number of AB, CD and EF trials in a block, respectively). After each training block it was evaluated whether the performance criterion was met. If the performance criterion was not met, the next training block was presented. A minimum of two and a maximum of seven training blocks were presented. Participants who did not meet the performance criterion at the end of the

seventh block (this held for 9 participants) were excluded from further participation, to make sure that every subject was at the minimum required performance level at the start of the test phase.

Preceding the training phase, participants were told that in addition to the stimulus pairs from the training phase, new combinations of the same Hiragana characters would be presented. They were instructed to again choose the character that would most likely be rewarded, but that feedback would no longer be given. The test phase consisted of three blocks of 90 trials, resulting in 270 trials. During each block all 30 (6 letters, with 2 possible positions for each letter) stimulus combinations including novel and training pairs were randomly presented three times. Each trial consisted of a fixation cross (same as during the learning phase), the stimulus pair presented for 1000 ms, and a blank screen presented for 950 ms. No feedback was given on participant's choice. If participants did not respond within 1000 ms after stimulus onset, the message "no response made" was presented. After completion of the test phase (or the seventh training round in case the performance criterion was not met) participants were subjected to an awareness questionnaire to assess whether participants were aware of the reward-punishment contingencies of each character. Finally, subjects were paid and dismissed.

Experiment 1b

The task in Experiment 1a was adapted from a task used in an EEG study by (Frank et al., 2005), in which a response window of 1000 ms was implemented, which is short compared to other studies. To address the possibility that the relatively short response window induces a stronger influence of stimulus-related factors, an additional experiment was run in which the maximum amount of time participants had to respond was changed to 4000 ms. Thus, the task and procedure were identical to the task and procedure of Experiment 1a, except that participants had a maximum of 4000 ms after stimulus onset to respond to the Hiragana stimuli in both the training phase and the test phase, as comparable to (Frank et al., 2004).

Data reduction and analysis

Participants who did not satisfy the performance criterion at the end of the seventh training round were excluded (9 participants in Experiment 1a (5 in version 1 and 4 in version 2) and 2 participants in Experiment 1b). Furthermore, participants who did not choose more than fifty percent A on the AB trials of the test phase (AB) were not analyzed further (0 participants in Experiment 1a and 3 participants in Experiment 1b). The overall percentage (of response trials, excluding no-response and double-response trials) of choosing stimulus A and avoiding stimulus B in novel test pairs (A or B paired with either C, D, E or F) during the test phase was calculated for each subject. Participants were categorized as positive learner if percentage choosing A was higher than percentage avoiding B. Participants were categorized as negative learner if percentage avoiding B was higher than percentage choosing A. A t-test was conducted for each task version group to assess whether participants were more inclined to choose stimulus A or to avoid stimulus B (in other words, a within-subjects test comparing % choose A in novel test pairs to % avoid B in novel test pairs). A Pearson Chi-square test (or Fisher's exact test if appropriate) was conducted to investigate whether there was a relationship between task version (version 1 or version 2) and categorization as positive or negative learner.

To analyze performance during the training phase, for each participant the percentage choices for the most rewarded stimulus (of response trials, excluding no-response and double-response trials) within each of the training pairs was computed. This was done for the first training round and for the last training round (i.e., the round before advancing to the test-phase, which was individually determined). These percentages were entered in a mixed MANOVA with stimulus pair (AB, CD, EF) and round (first, last) as within-subject factors and task version (1, 2) as between-subjects factor.

Results – Experiments 1a and 1b

Experiment 1a

Ten out of 11 participants who finished version 1 of the probabilistic learning task were better at choosing A than avoiding stimulus B in novel

test pairs and were consequently classified as positive learner (within-group t-test for the difference between choosing A and avoiding B: $t(10) = 4.31$, $p = .002$, $d = 1.3$). Ten out of 11 participants who finished version 2 of the task were categorized as negative learner because of their greater accuracy in avoiding stimulus B compared to choosing A (within-group t-test: $t(10) = -5.39$, $p < .001$, $d = -1.62$). Mean percentage A and B accuracy of the two task versions is shown in Figure 2A. A Pearson Chi-square test confirmed the relationship between task version and categorization as positive or negative learner (χ^2 (1) = 14.73, $p < .001$). The corresponding phi coefficient was 0.82 ($p < .001$), which represents a high association between task version and categorization as positive or negative learner.

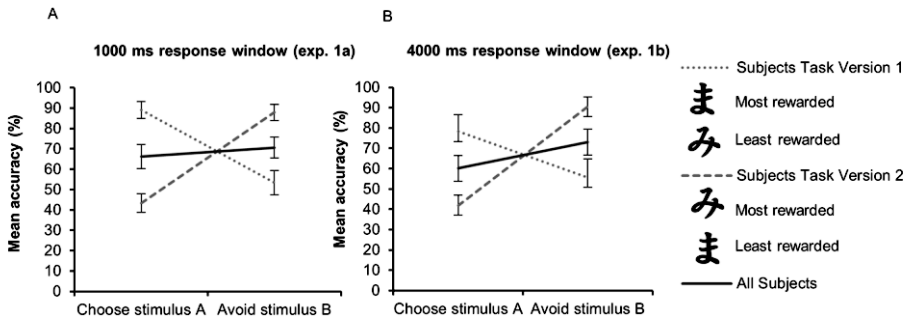


Figure 2. Stimulus properties affect extent of learning from positive vs. negative feedback. This figure shows stimulus A and B accuracy (choosing the most rewarded stimulus) in the test phase for both task versions in Experiment 1a (panel A) and Experiment 1b (panel B). In both experiments (i.e., regardless of the length of the response window), subjects who completed task version 1 were overall better at choosing A, whereas in version 2 subjects were better at avoiding B. These findings indicate that stimulus salience may strongly bias value-based probabilistic learning. Error bars represent ± 1 SE.

To determine whether these differences were accompanied by differences in experienced reward outcomes during learning (e.g., if better choose-A performance in version 1 was associated with more instances of positive feedback for choosing A compared to other stimuli), we analyzed performance in the training phase. Figure 3A displays training performance for each of the Hiragana pairs for task version 1 and task version 2, respectively. The repeated-measures MANOVA comparing training performance between task versions revealed a main effect of block, indicating that individuals were better at choosing the most rewarded

character of each pair during the last compared to the first block of the training phase ($F(1,20) = 31.4$, $p < .001$, $\eta_p^2 = .611$). Importantly, there was also an interaction between task version and stimulus pair ($F(2,19) = 4.39$, $p = .027$, $\eta_p^2 = .316$). Follow-up MANOVAs for each task version separately showed that there was a significant main effect of pair for task version 1 ($F(2,9) = 4.36$, $p = .047$, $\eta_p^2 = .492$) and for version 2 ($F(2,9) = 21.93$, $p < .001$, $\eta_p^2 = .830$). Post-hoc tests revealed that in version 1, AB performance was greater than CD performance at near significance ($p = .052$); whereas no such differences were observed in version 2 ($p = .946$), although AB and CD performance were both significantly better than EF performance (p's $< .001$). These findings suggest that enhanced choose-A vs. avoid-B performance in version 1 might relate to relatively greater experience of positive feedback for A and hence more learning about the positive value of A compared to e.g., stimulus C, whereas better avoid-B performance in version 2 might relate to relatively worse performance in AB and hence more negative feedback for B. The larger sample sizes in Experiments 3 (and replication in Experiment 1b) below further support this intuition.

Experiment 1b

Of the ten participants subjected to task version 1, eight participants were classified as positive learner, because of their greater accuracy in choosing stimulus A in novel test pairs, compared to avoiding B. All ten participants subjected to task version 2 were classified as negative learner, because of their greater accuracy in avoiding stimulus B in novel test pairs, compared to choosing A. A t-test conducted for each task version group showed that participants subjected to task version 2 were significantly better at avoiding stimulus B than choosing A in novel test pairs ($t(9) = -6.07$, $p < .001$, $d = -1.92$). Participants subjected to task version 1 tended to be better at choosing stimulus A in novel test pairs ($t(9) = 1.52$, $p = .164$, $d = 0.48$). Mean percentage A and B accuracy are shown in Figure 2B for each task version separately. Results of the Fisher's exact test indicated a strong relationship between task version and categorization as positive or negative learner ($p = .001$). The corresponding phi coefficient was 0.82 ($p = .001$).

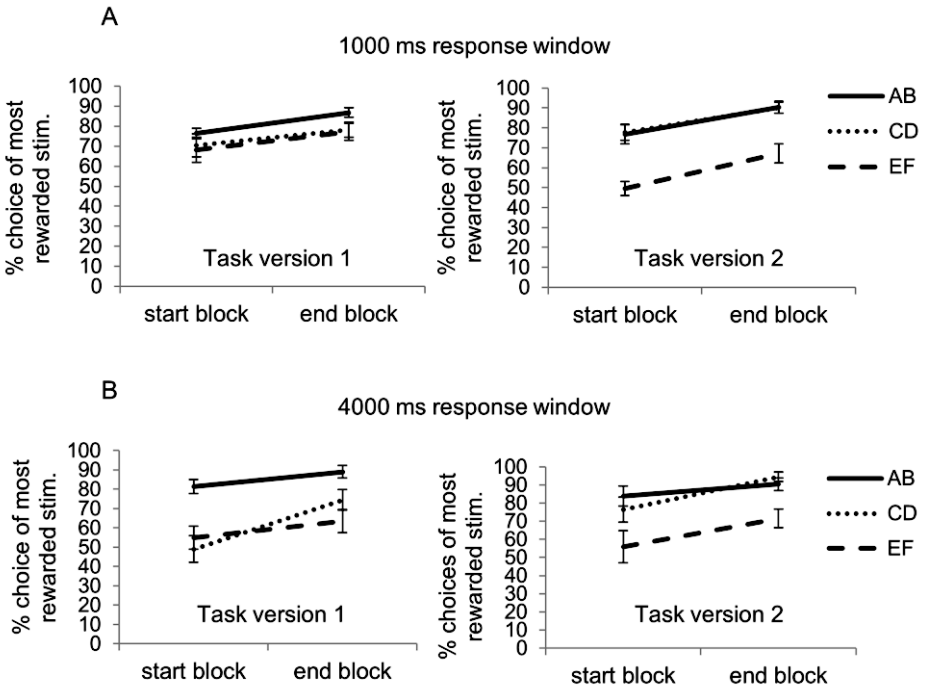


Figure 3. Stimulus properties affect choices during the training phase of the probabilistic learning task. The figure displays accuracy (choosing the most rewarded stimulus) for each of the three Hiragana pairs for both task versions in Experiment 1a (panel A) and 1b (panel B). In task version 2, subjects performed better at both the AB and CD pair relative to the EF pair, compared to task version 1. This pattern was consistent across both experiments. Furthermore, subjects in task version 1 of Experiment 1b performed better at the AB pair relative to the other pairs, compared to task version 2.

Results in the training phase were also similar to Experiment 1a. Figure 3B displays training performance for each of the Hiragana pairs for task version 1 and task version 2, respectively. As expected, the repeated-measures MANOVA comparing training performance between task versions indicated overall better performance during the last compared to the first round of the training phase ($F(1,18) = 37.24$, $p < .001$, $\eta_p^2 = .674$). The interaction between task version and stimulus pair was again significant ($F(2,17) = 4.37$, $p = .029$, $\eta_p^2 = .339$). Follow-up MANOVAs for each task version separately showed a significant main effect of stimulus pair for both versions (task version 1: $F(2,8) = 22.58$, $p = .001$, $\eta_p^2 = .850$; task version 2:

$F(2,8) = 9.23$, $p = .008$, $\eta_p^2 = .698$). Replicating the results of Experiment 1a, performance in the AB pair was significantly better compared to both of the other pairs in task version 1 (p's ≤ .003), whereas in version 2 AB performance did not differ from CD (p =. 71) while performance for both the AB and CD pair was enhanced relative to the EF pair (p's ≤ .004). In version 1 there was no difference in performance between the CD and EF pair (p = .673).

Discussion – Experiments 1a and 1b

The results of Experiment 1a show that classification as positive or negative learner was strongly affected by the stimulus-to-feedback mappings. We replicated this finding in Experiment 1b in which subjects were given more time to respond. Importantly, this biased classification pattern during the test phase was preceded by differences in performance between the task version groups during training. Participants subjected to version 1 of the task chose the rewarded A over B more reliably than they chose C over D or E over F. In contrast, participants subjected to version 2 consistently chose the more rewarded stimulus C over D equally often as they chose A over B. Because of that, the difference in amount of experienced positive feedback for A relative to C was reduced in version 2 relative to version 1 (in which participants experienced more positive feedback for A than for C). In contrast, the difference in amount of experienced negative feedback for B relative to D was increased in version 2 relative to version 1. These differences in feedback learning or experience may have contributed to the biased classification as positive/negative learner. For example, in version 2, the reduced amount of positive feedback for A relative to C may have resulted in more choices for C over A during the test phase, compared to version 1. Similarly, version 2 enhanced the degree to which subjects experienced more negative feedback for stimulus B than D. Together, these effects would contribute to differential ability to choose A and avoid B across versions. The training results will be more extensively interpreted in the general discussion.

Hence, we hypothesized that in the pair-wise setup of the probabilistic learning task some of the characters may have been more easily discriminable than others. This would contribute to relatively

enhanced learning for easily discriminable letters about contingent mainly positive or mainly negative feedback. In this scenario, choices during learning and test phases would not reflect 'positive' or negative' learning styles exclusively, but also the extent to which an easily discriminable stimulus was mainly associated with positive or negative feedback. We therefore explicitly investigated whether some stimuli of the set used in Experiment 1 are more discriminable than others, in a follow-up Experiment 2. Each combination of two stimuli in a random sequence was framed in the context of a standard choice-reaction-time decision-making experiment. We assume that on each trial the evidence for either stimulus identity drifts until a criterion is reached of sufficient evidence to trigger a response. We further assume that the more discriminable the two stimuli are, the more quickly the evidence criterion is reached for either stimulus. Therefore, for pairs consisting of one or two highly discriminable stimuli, reaction times for either of the stimuli will be relatively short, and so will the average reaction time across all trials in that sequence. A systematic difference in discriminability among the six characters would then be revealed by comparing reaction times averaged across all trials within each combination of two characters, subsequently averaged across all combinations containing that specific stimulus.

Method – Experiment 2

Participants

We included 10 new participants in the second experiment (mean age: 23, range: 19.5-30.9 years, 6 females, 9 right-handed). The sample consisted of trainees and subjects recruited through posters. All participants declared not to have consumed caffeine or nicotine on the day of the experiment. None of the subjects was receiving mental health treatment and none of them was aware of the aim of the experiment. All subjects had normal or corrected-to-normal vision and gave written informed consent. A financial compensation of 3 Euro per half an hour or study credits was given to externally recruited subjects. Participants were treated according to the Declaration of Helsinki. This experiment was approved by the local ethics advisory board of the faculty of social sciences of Utrecht University.

Task and procedure

Upon arrival at the lab, subjects received information on the procedure of the experiment and informed consent was obtained. As in Experiment 1, participants were seated 85 cm in front of a computer screen. Presentation of instructions and stimuli was controlled by Presentation® software (version 16.0, www.neurobs.com). One non-Dutch participant received task instructions in English, the others in Dutch. The same Hiragana characters (white on a black background within a frame of 3° by 3°) were used as in Experiment 1.

An example of the task is shown in Figure 4. Each block of the discrimination task started with a visual instruction of 5 seconds showing which button to press in response to which specific Hiragana stimulus. During each block the "z" key had to be pressed when one of the stimuli of the current pair was presented and the "m" key had to be pressed when the other stimulus was displayed. A random Hiragana character of the current pair could either be presented on the left or right side of the screen and was accompanied by a square on the opposite side. Participants were instructed to respond to the stimuli as fast and accurately as possible. It was emphasized that they had to react to the stimulus character independent of its location on the screen.

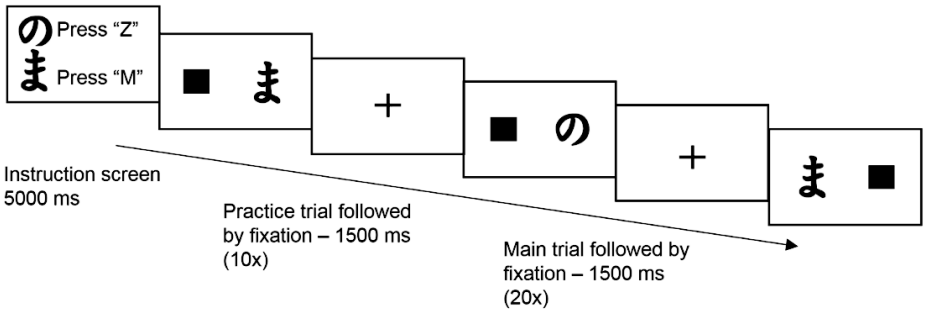


Figure 4. The discrimination task used in Experiment 2. The six Hiragana characters were exhaustively combined in separate pairs. Participants were instructed to respond as fast and accurate as possible to each of the stimuli of the current pair according to the instructions presented at the beginning of each block. Note that stimuli and text are presented in black for illustrational purposes, in contrast to the actual task.

Presentation and position of stimuli was counterbalanced across trials. A block started with 10 training trials to practice responses to the Hiragana stimuli of the current block and was followed by 20 experimental trials. A message was presented between the last practice trial and the first experimental trial stating "This was a practice round. The main round will follow now". The task consisted of 30 blocks presented in random order comprising all combinations of the aforementioned Hiragana characters and corresponding button presses. Subjects pressed the spacebar to continue with the next block. There was opportunity for a small break after the 10th and 20th block. After completion of the task participants were paid (if applicable), thanked for their participation and dismissed.

Data reduction and analysis

The total number of trials for each subject amounted to 600. Individual trials with RTs > 1000 ms, incorrect responses and multiple responses were removed from the analysis. This pertained to 8.6% of the trials across subjects (range: 3.5 - 21.2% of the trials per subject). Out of a total of 1200 cells (10 subjects x 30 blocks x 2 stimuli x 2 positions), 17 cells (1.4%) did not contain data due to zero correct and valid responses for those cells (range: 0 – 5.8% empty cells per subject). Empty cells were replaced by values corresponding to the same block and stimulus (e.g. A versus B block, presentation of stimulus A), but with reversed hand mapping. The number of empty cells did not differ across the 20 conditions (5 combinations x 2 hand mappings x 2 positions). For each subject RTs were averaged (across at maximum 5 trials) for each character within each stimulus combination, for left and right hand mapping and incongruent and congruent (i.e., screen position with respect to response hand) presentations separately. Next, for each subject all reaction times were averaged within each stimulus combination. Data were subsequently averaged across all stimulus combinations containing one specific stimulus (i.e., all combinations containing stimulus A, all combinations containing B, etcetera). Finally, six average RTs for each subject (one for each Hiragana character) were entered into a MANOVA. Post-hoc pairwise comparisons were conducted in order of RT difference between two Hiragana characters, starting with the pair with the largest difference in RT, followed by the pair with the next largest

difference, etcetera. This procedure was continued until a non-significant difference was obtained.

The data were also analyzed in a different manner, namely by only including trials with the presentation of a specific stimulus and by subsequently averaging these data across all conditions with that specific stimulus, see the Supplementary Materials (Alternative analysis).

Results – Experiment 2

A repeated-measures MANOVA revealed a significant main effect of stimulus character on RT ($F(5,5) = 29.91$, $p = .001$, $\eta_p^2 = .968$). Post-hoc pair-wise comparisons in order of RT difference between two Hiragana characters showed significantly shorter RTs for D compared to B, C, E and F (all p values ≤ .001) and for A compared to B, C, E and F (all p values ≤ .008). Results are shown in Figure 5.

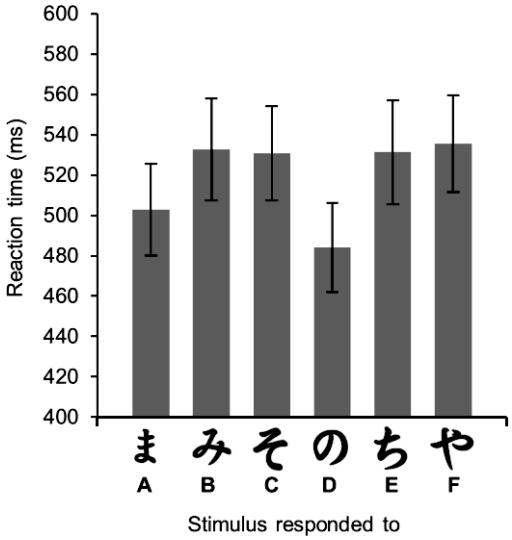


Figure 5. Reaction times to each of the Hiragana stimuli. Subjects reacted significantly quicker to pairs containing stimulus A and D compared to pairs containing the other characters. Error bars represent ± 1 SE. Note that error bars depict variation per condition, not the difference between conditions (as used in the statistical error term).

Discussion – Experiment 2

Results of the discrimination experiment clearly show that RTs differ between Hiragana stimuli. Stimulus A, which was the most likely to be rewarded in version 1 of our reinforcement learning experiment and least

rewarded in version 2 of that experiment seems more easily discriminable, as evidenced by faster responses for pairs containing this stimulus compared to pairs containing other stimuli except stimulus D. These findings suggest that the relatively high discriminability of stimulus A led to the observed differences in feedback-based learning between the two task version groups during the training phase of the PST in Experiment 1. These differences in learning, in turn, may explain the observed strong relationship between stimulus-to-feedback mapping and whether or not an individual is classified as a positive or negative learner. Specifically, high discriminability of stimulus A in task version 1 in Experiments 1a and 1b led to more choices for A and therefore to a relatively high amount of positive feedback for stimulus A compared to the other training stimuli, especially compared to C. This, in turn, resulted in a positive learning bias (i.e., more accurate choices of A over other stimuli during test). In contrast, low discriminability of stimulus A combined with salience of C in task version 2 may have led to equal training performance for AB as for CD. This in turn resulted in less experienced positive feedback for A (relative to C) in version 2, and to more experienced negative feedback for B (relative to D). This, in turn, may have resulted in relatively few choices for A during the test phase as well as increased avoidance of B, and therefore in a negative learning bias (i.e., more avoidance of B compared to choosing A during test).

Moreover, results of the discrimination experiment show that stimulus character D is also relatively easy to discriminate compared to the other stimuli. This fits fairly well with the performance data of the training phase in Experiments 1a and 1b. Participants more readily selected this character when it was the more rewarding C stimulus during training in task-version 2, where it was rewarded in 70% of the trials. This explains the relatively high amount of positive feedback during the training phase for the CD pair in task version 2. In contrast, participants were less inclined to select the more rewarded stimulus C from the CD pair in version 1, presumably because it was less discriminable.

It is thus possible that the relatively weak inclination to choose A during the test phase in task version 2 reflects a relatively strong inclination to choose stimulus C in AC pairs. Thus in version 1, subjects might have learned both greater positive value for A and simultaneously, less positive value for C, facilitating their ability to choose A over C; the reverse bias in

version 2 might effectively yield a more positive learned value for C than for A.

To test this idea, and to confirm the findings of the first study, we conducted a third experiment with a large sample size using a web-based implementation in samples of American rather than Dutch subjects. Along with the first two task versions, a third task version was implemented. Here, instead of swapping all of the characters (as in version 1 vs 2), we maintained the same Hiragana stimuli for the CD and EF pair as in version 1, but only swapped the A and B stimuli (i.e., to the mappings used for version 2). The logic here was that the *asymmetry* in discriminability between A and C was the main factor driving the large differences in versions 1 and 2 with opposite learning biases, and hence if these were matched such that A and C were both less discriminable, this should produce an intermediate effect with no clear bias one way or the other. Indeed, this would result in low-salient A and C stimuli that are mostly rewarded, and high-salient B and D stimuli being mostly punished. In turn this would prevent a salience-driven bias across subjects to systematically choose A or avoid B during testing.

Method – Experiment 3

Participants

A sample of 300 participants were recruited via Amazon's Mechanical Turk (AMT). Participants were paid $3 for their time. Based on self-report, subjects were required to be between 18-40 years of old, fluent in English, to have no history of brain injury, no medical history of mental/psychiatric disorders or drug/alcohol abuse, and were only able to complete the task once. In Experiment 3 five versions of the PST were compared: version 1 and 2 with a short response window (1000 ms), version 1 and 2 with a long response window (4000 ms) and version 3. Participants completed one of these five task versions. Participants (n = 70) in version 1 long included 42% female and mean age was 31.3 (SD=5.7). Version 1 short (n = 56) included 48% female, mean age was 31.8 (SD=5.6). Version 2 long (n = 62) included 37% female and mean age was 30.4 (SD=5.7). Version 2 short (n = 59) included 44% females and mean age was 29.4 (SD=5.1). Version 3 (n = 53) included 42% female and mean age=29.7 (SD=5.2). Experiment 3 was

approved by the Brown University institutional review board. Participant identities were protected. Participants provided their consent online by clicking 'I Agree' after reading the study information and consent language in accordance with procedures approved by Brown University.

Task and procedure

Data were collected using the Amazon Mechanical Turk platform. Subjects completed a practice phase with verbal instructions with example stimuli (not used in the actual task) using deterministic feedback. Based on previous studies using AMT (Crump, McDonnell, & Gureckis, 2013; Gillan & Daw, 2016), participants were given a 5-item basic true-false comprehension test regarding the rules of the task. Failure to answer the questions correctly resulted in subjects repeating the practice phase and instructions.

Task versions 1 and 2 were identical to the ones described under Experiment 1a and 1b. The only exception was that in the long-response-window versions (comparable to those in Experiment 1b), maximal stimulus duration was not 750 ms but 4000 ms, effectively covering the complete response window, as is typical for behavioral studies with this task (although note that response times are typically < 1.5s and hence stimulus duration is rarely greater than 2000 ms). For task version 3 the CD and EF pairs were as in version 1, but the AB pair was used from version 2 (see logic for this manipulation above). A response window (and maximal stimulus duration) of 4000 ms was implemented in task version 3.

Data reduction and analysis

Participants were excluded from the data analysis if they missed or responded faster than 200 ms to more than one-third of learning or test phase trials, and if they performed at or worse than chance in test phase training pairs. The final analysis included 58 out of 70 participants in version 1-long, 46 out of 56 in version 1-short, 46 out of 62 in version 2-long, 48 out of 59 in version 2-short, and 42 out of 53 in version 3.

Compared to the relatively extensive data analysis presented for Experiments 1a and 1b, the data from Experiment 3 were subjected to a concise and maximally surveyable test procedure addressing the essential hypotheses. To corroborate the findings of Experiment 1a and 1b we first

examined the effect of stimulus mapping and presentation duration on learning bias. The overall percentages of choosing stimulus A and avoiding stimulus B in novel test pairs (A or B paired with either C, D, E or F) during the test phase were computed for each subject. Bias scores were subsequently computed by subtracting the percentage avoid-B from the percentage choose-A. These scores were entered in an ANOVA with task version (version 1 and 2) and response window (short, long) as between-subject variables.

Next, we sought to assess whether the effects of stimulus-to-feedback mapping on test phase performance as observed in Experiment 1a and 1b could be explained more proximally by differences in the reinforcement history experienced by the different groups, as suggested above. Specifically, we examined the difference in rewards experienced for A compared to the next most rewarding item C (and conversely the number of punishments experienced for B compared to the next most punishing item D). We, therefore, subtracted the percentage choose-C in CD pairs from the percentage choose-A in AB pairs in the last training block. These percentages were entered in an ANOVA with task version (version 1 and 2) and response window (short, long) as between-subject variables.

To directly test whether the critical factor in test phase bias was the relative experience of feedback for A compared to other stimuli an ANOVA was run comparing AB-CD performance for the last training block and test phase bias scores between the three task versions (now also including task version 3).

Results – Experiment 3

First, in order to determine whether the results of Experiment 1a and 1b were replicated in a larger sample we compared test phase bias between the two task versions presented previously (version 1 vs. version 2), with two different response windows (long; up to 4000 ms vs. short; up to 1000 ms). Replicating the results of Experiment 1, we found a main effect of task version on choose A-avoid B bias ($F(1,194) = 127.2$; $p < .001$, $\eta_p^2 = .4$). As can be seen in Figure 6, left panel, percentage choose A compared to avoid B was higher for version 1 but lower for version 2. Furthermore, stimulus presentation duration had no effect on choose A-avoid B bias ($F(1,194) <$

0.04; p = 0.86) and did not interact with task version group (F(1, 194) = 0.1; p = 0.75), confirming that the previous results are not dependent on the specific timings used in the PST.

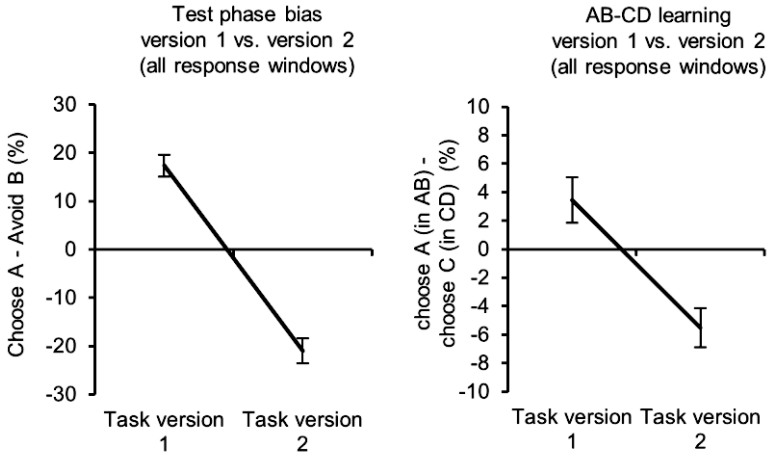


Figure 6. Stimulus properties affect positive and negative learning bias and choices during the training phase (Experiment 3: large international sample of participants, web-based study). Left: replicating the results of Experiment 1a and 1b, subjects in version 1 of the PST were better at choosing A during test phase whereas subjects in task version 2 were better at avoiding stimulus B. Right: replicating the results of Experiment 1a and 1b, test phase bias was accompanied by group differences during training. In task version 1 participants performed better on the AB pair relative to the CD pair, whereas the opposite pattern was observed for task version 2. There were no significant interactions with the time to respond (1000 ms or 4000 ms). Data is therefore collapsed across both time windows. Error bars represent ± 1 SE.

The ANOVA comparing AB-CD performance during training between task version 1 and 2 showed that there was a strong main effect of task version ($F(1,194) = 16.5$; $p < .001$, $\eta_p^2 = .08$). Again, participants in the version 1 group (most of them with a positive test bias, choose A > avoid B) also chose A over B more (and hence received comparatively more positive feedback) than C over D. Conversely, participants in the version 2 group (most of them with a negative test bias, choose A < avoid B) selected B over A more than D over C during training and experienced more negative feedback for doing so (Figure 6, right).

For the comparisons between the three groups we only used the long (4000 ms) response window/ presentation duration as the previous analysis showed that presentation did not impact test-phase bias and hence version 3 was only run with this standard presentation. Results from an ANOVA showed a main effect of task version on AB vs. CD choices during the last learning block ($F(2,143) = 4.78$; $p < 0.01$, $\eta_p^2 = .06$; Figure 7, right) and test-phase bias ($F(2,143) = 23.62$; $p < .001$, $\eta_p^2 = .25$; Figure 7, left panel). The effect of task version on AB-CD choices during training was driven by a significant difference between version 1 and 2 ($p = .002$). There was no significant difference between task version 3 and the other versions (v1 vs 3: $p = .13$; v2 vs 3: $p = .19$).

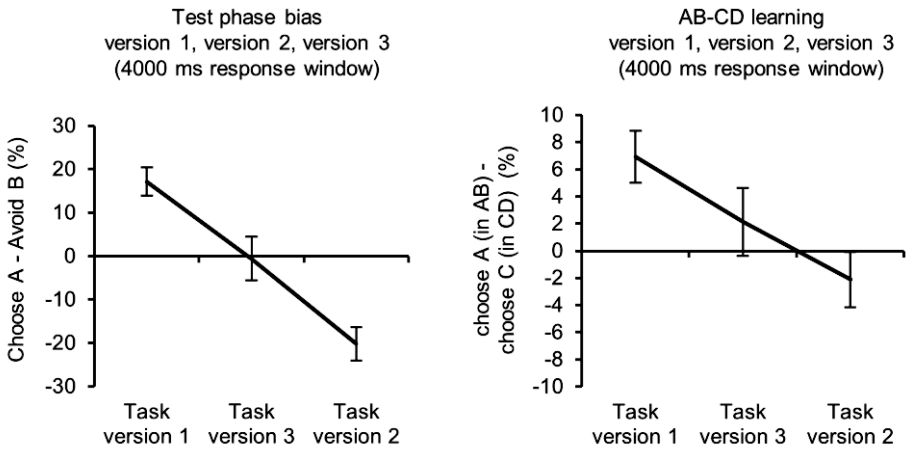


Figure 7. Relative experience of feedback for A compared to all other stimuli critically determined test phase bias (Experiment 3). Left: participants in task version 3 in which the A and C stimulus were more comparable in discriminability showed no overall test phase bias, and it was significantly lower than in version 1 and higher than version 2. Right: there was a main effect of task version group in last learning block AB vs. CD performance. Participants in task version 3 performed equally well on the AB and CD pair in contrast to participants in task version 1 and 2. The task versions presented had a response window of 4000 ms. Error bars represent ± 1 SE.

Critically, group 3 showed no overall test- phase bias, and this bias was significantly lower than that for group 1 (v1 vs. v3: $t(98) = 3.11$; $p = .002$, $d = .63$) and higher than that for group 2 (v3 vs. v2: $t(86) = 3.14$; $p = .002$, $d = .67$). This ranking of version1>version3>version2 in the test phase bias confirms our hypothesis that the critical factor driving test phase biases is relative discriminability and hence reinforcement history for A/B compared to C/D.

Discussion – Experiment 3

The results of Experiment 3 replicate the findings of Experiment 1a and 1b in a large scale web-based sample and provide further evidence that test-phase bias is strongly predicted by stimulus-to-feedback-mapping. These experiments also show similar between-group differences in learning curves for each of the training pairs, indicating that the positive/negative bias is likely explained by differences in feedback learning or experience during the training phase.

In an additional task version 3, we examined performance in a case in which only the stimuli of the AB pair were switched (i.e., CD and EF pair as in version 1, AB as in version 2). The aim was to investigate whether the absolute amount of reinforcement for stimulus A and B or the relative amount of reinforcement for A and B compared to all other stimuli was the critical factor determining positive/negative learning bias. The results indicate that the latter was the case. No clear positive or negative learning bias was observed for this particular task version, suggesting that when discriminability is matched, individual differences are more likely related to veridical differences in positive and negative learning per se.

Training performance for version 3 was in between that for version 1 and version 2 (as assessed in Experiment 3; Figure 7 right panel). Specifically, there was a large difference between AB and CD performance in version 1; a slightly reversed difference in version 2 (which means that C was chosen over D slightly more than A over B); and an intermediate (slightly positive) difference in version 3. For version 3 this resulted in an amount of experienced feedback for A (relative to C), as well as an amount of experienced negative feedback for B (relative to D) that was intermediate between that for version 1 and version 2.

This was associated in the test phase with an about equal amount of choices to avoid B and to avoid D in BD pairs. By the same token, there was an equal amount of choices for A as there were for C in AC pairs. Together this effectively annihilated the contribution of relative salience to negative or positive-learner classification.

General discussion

The aim of the current study was to investigate if the extent to which individuals learn better from positive or negative outcomes is influenced by the extent to which aspects of the stimulus configuration afford learning from the stimuli in general. In two experiments (Experiment 1a and 1b), participants performed one of two versions of a common probabilistic selection task described by Frank et al. (Frank et al., 2004; Frank et al., 2005). These task versions were identical except instead of randomizing or counterbalancing the stimulus assignments as is typically done, here we explicitly fixed them in each version but manipulated them across groups to study their impact. Specifically, the assignment of positive and negative feedback to the stimuli was switched, e.g., the A stimulus (rewarded 80%) in the first version became the B stimulus (punished 80%) in the second version, and vice versa. Critically, we found that during the test phase subjects almost exclusively chose A (rather than avoid B) in version 1 but avoided B (rather than choose A) in version 2. Similar results were obtained regardless of presentation duration and response windows across all experiments. When task-version-1 stimulus A was associated with mostly positive feedback, all except one (Experiment 1a) or two (Experiment 1b) participants were characterized as positive learners, whereas when this same stimulus became B in task version 2 and was associated with negative feedback, all except one participant (Experiment 1a) or all participants (Experiment 1b) were categorized as negative learners. Furthermore, these results were replicated in a large-scale web-based experiment (Experiment 3).

Notably, a follow-up discrimination experiment revealed that two of the six Hiragana characters (A and D in version 1; B and C in version 2) were more discriminable than the others, as assessed by response times. Together these findings indicate that the relative perceptual

discriminability of characters assigned to high or low reward probabilities strongly influenced performance.

Importantly, biased performance during the test phase of the PST was preceded by group differences in learning curves during training. These differences were consistent across all experiments. In task version 1, participants most readily selected the highly salient A over B and hence received a high amount of positive feedback for the A stimulus relative to the other training characters. This pattern was most pronounced for the task with a response window of 4000 ms. In contrast, in task version 2, participants showed comparable accuracy in AB and CD pairs, despite the lower reward probability of the latter. This implies that, relative to version 1, the amount of experienced positive reinforcement for A relative to C was reduced, and the amount of experienced negative reinforcement for B relative to D was increased. This may have resulted in a relative prevalence of C over A choices during the test phase, as well as a stronger avoidance of B over D, which in turn resulted in the inclination to more avoid B than to choose A. The latter pattern was observed for both response windows. The pattern of differences during training performance was also visible already in the start training blocks in Experiments 1a and 1b, indicating that it emerged rapidly within the very first ten repetitions of each stimulus and associated feedback.

Experiment 3 was conducted in order to test whether these findings were robust by replicating the experiment in a larger sample. Indeed, Experiment 3 confirmed the interpretation in terms of relative discriminability and feedback. Moreover, task version 3 was developed to more specifically engineer a stimulus mapping that would produce less bias under this interpretation. In this version, relative to version 1, only the stimuli within the A-B pair were switched, thus better matching the discriminability of A vs C, and indeed no learning biases were observed in the test phase. The results indicate that the relative amount of reinforcement for A and B compared to all other training stimuli influenced learning bias. Specifically, in version 3 the relative AB – CD performance difference during learning – and hence the degree to which A was reinforced more than C and B punished more than D – was intermediate between version 1 and version 2. In the test phase this resulted in an about equal

amount of choices to avoid B and to avoid D, and to an equal amount of choices for A as there were for C.

Together, our findings indicate that stimulus salience/ discriminability strongly affects preference behavior during reinforcement learning. A highly salient stimulus that is most frequently rewarded rapidly attains a preferred status (perhaps due to being more memorable), is chosen much more often, and as such obtains a higher learned value, than a less salient stimulus that is less frequently rewarded. In contrast, a relatively less salient stimulus that is most frequently rewarded attains a level of preference that is comparable to or even lower than that of a highly salient stimulus that is less frequently rewarded.

On a more general level our results are consistent with prior studies indicating that stimulus salience and value information interact to bias learning and decision making (Gold et al., 2008; Lou et al., 2015; Navalpakkam et al., 2010; Polanía et al., 2014; Smith & Ratcliff, 2004). Our analysis of discriminability in relation to learning also fits earlier accounts that stress the importance of similarity among less discriminable objects rather than the distinctiveness of more discriminable objects (e.g., Hunt, 1995). As argued in the discussion of Experiment 2, the four less discriminable stimuli in our setup can be considered mutually less discriminable and therefore relatively similar or homogenous, compared to the two stimuli that were more discriminable relative to the less discriminable ones. To the extent that superior learning for salient items actually depends on the homogeneity of the non-salient items (as, e.g., in the von Restorff effect), this homogeneity was actually realized in the present training sessions.

On a more operational level our results imply that one needs to exert caution in interpreting data from individual subjects in the PST and comparable reinforcement learning tasks (e.g., Bodi et al., 2009; Frank et al., 2004; Palminteri et al., 2009; Piray et al., 2014). Note that the cited studies using the PST counterbalanced or randomized the stimulus mappings across subjects, and hence previous replicated findings of individual differences due to genetics (Cockburn et al., 2014; Doll et al., 2011; Frank et al., 2007; Klein et al., 2007), striatal D1 and D2 dopamine receptor binding (Cox et al., 2015), and neural responses to feedback (Cavanagh et al., 2011; Frank et al., 2005; Jocham et al., 2011; Klein et al., 2007) cannot be

attributed to discriminability, given that these observed effects held despite stimulus counterbalancing, and largely were present without differences in learning curves (in contrast to effects of discriminability seen here). However, the present results suggest that first, such counterbalancing is paramount; second, that test phase performance differences should always be accompanied by analysis of learning curves during training; and third, that it is difficult to interpret the results of any one individual (as opposed to a group of individuals with e.g. same genotype) unless care is taken to match the discriminability of the stimuli.

On the other hand, the results of the third version in effect point the way to the implementation of the PST in a manner that reduces the contribution of stimulus salience to classification as either negative or positive learner, at least at the level of group averages, by matching discriminability. Indeed an initial analysis of our other online datasets using everyday geometric shapes that are all highly discriminable (e.g. blue square, red triangle etc.), revealed no effect of stimulus assignment on learning bias (unpublished data). The very possibility of such a contribution, as well as the solution proposed here, have ramifications for any task variety in which per subject a fixed relation between physically different stimuli on the one hand, and feedback conditions on the other is maintained (see, e.g., the references to other PST varieties provided in the Introduction).

It may also be noted that some tasks use the very same stimuli that are assigned to both reward and punishment, where those values alternate across blocks (as in reversal learning). In this case there are no differences in discriminability across conditions. Notably, much like the PST, individual differences in sensitivity to positive and negative outcomes in reversal tasks are related to individual difference in striatal dopamine levels assessed with PET and these biases are similarly affected by dopaminergic medications (Cools, Altamirano, & D'Esposito, 2006; Cools et al., 2009).

Nevertheless, the current data suggest that differences in the discriminability of stimuli can add significant noise to the measurement of individual differences in learning style. In previous studies using the PST or a comparable reinforcement-learning task, participants likely also integrated sensory evidence and value to form their own estimates of expected reward (or punishment). It is notable in this respect that a recent

ERP study found that the ERN is modulated by both stimulus salience and reward level (Lou et al., 2015), in line with the notion that both the perceptual properties of stimuli and their value influence probabilistic learning and future decision making.

Although there was a strong overall relationship between salience-reward contingencies and outcome as positive/negative learner, not all participants subjected to version 1 of the reinforcement task (in which a highly salient stimulus was associated with positive feedback) were classified as positive learner in our study. In the same vein, not all participants subjected to version 2 (in which a highly salient stimulus was associated with negative feedback) were classified as negative learner. Thus, while our results indicate that differences in stimulus discriminability bias the outcome of the reinforcement learning task in terms of 'positive' or 'negative' learning style, stimulus discriminability is not the only factor that determines learning style. This is again in line with recent work suggesting that stimulus salience and stimulus value both affect learning about expected reward (or punishment) (Gold et al., 2008; Lou et al., 2015; Navalpakkam et al., 2010; Polanía et al., 2014; Smith & Ratcliff, 2004).

In sum, the current study shows that differences in stimulus salience /discriminability are a potential confounding factor in reinforcement learning tasks during both learning and transfer phases. The extent of selecting rewarding stimuli or avoiding punished ones depends on the relative salience of these stimuli, both during and after learning. To avoid these potential confounds in future research, task stimuli should be matched at the individual level in terms of relative discriminability; learning curves should always be assessed before interpreting test phase data, and/or a larger number of stimuli should be used within subject having similar values. On a more theoretical note, the present study elucidates how reinforcement and punishment learning are influenced by the relative salience tied to choice alternatives. This extends a growing body of both older (e.g., Lubow, 1973; Von Restorff, 1933) and recent literature on the relation between salience and learning.

Supplementary materials

Alternative analysis Experiment 2

The data were also analyzed in a different manner, namely by only including trials with the presentation of a specific stimulus and by subsequently averaging these data across all conditions with that specific stimulus. Differences as revealed by this analysis may reflect not only discriminability but also differences in preference. This may hold to the extent that certain single specific stimuli are responded to more quickly not only because they are more discriminable, but also because of their inherent likeability, inducing a more vehement approach tendency. Six average RTs for each subject (one for each Hiragana character) were entered into a MANOVA. The results of this analysis yielded very similar results compared to the analysis reported in the Results section of Experiment 2, which indicates that preference or approach does not have an additional impact on reaction times and that discriminability is the exclusive driving factor for the observed differences.

Chapter 4

Resting-state theta/beta EEG ratio is associated with reward- and punishment-related reversal learning

Iris Schutte
J. Leon Kenemans
Dennis J. L. G. Schutter

Cognitive, Affective & Behavioral Neuroscience (2017)

Author contributions

All authors were involved in study conceptualization; DS designed the paradigm; IS and DS collected the data; IS and DS performed the analyses; IS wrote the initial draft of the manuscript; all authors contributed to the final manuscript.

Abstract

Prior research has shown that the ratio between resting-state theta (4-7 Hz)-beta (13-30 Hz) oscillations in the electroencephalogram (EEG) is associated with reward- and punishment-related feedback learning and risky decision making. However, it remains unclear whether the theta/beta EEG ratio is also an electrophysiological index for poorer behavioral adaptation when reward and punishment contingencies change over time. The aim of the present study was to investigate whether resting-state theta (4-7 Hz)-beta (13-30 Hz) EEG ratio correlated to reversal learning. A four minute resting-state EEG was recorded and a gambling task with changing reward-punishment contingencies was administered in one hundred and twenty-eight healthy volunteers. Results showed an inverse relationship between theta/beta EEG ratio and reversal learning. Our findings replicate and extend previous findings by showing that higher midfrontal theta/beta EEG ratios are associated with poorer reversal learning and behavioral adaptive responses under changing environmental demands.

Introduction

The sensitivity to reward and punishment signals guides decision-making by exploiting acquired knowledge to shape a long-term adaptive strategy (Bechara, Damasio, Damasio, & Anderson, 1994). Prior research has demonstrated that low punishment sensitivity together with a strong reward dependency predicts risky disadvantageous decision making, whereas high punishment sensitivity and weak reward dependency predicts advantageous decision making (van Honk, Hermans, Putman, Montagne, & Schutter, 2002).

There is now ample evidence that spontaneous oscillations play a critical role in brain functions (Fries, 2005; Knyazev, 2007) and electrophysiological studies have demonstrated that the sensitivity to reward and punishment is reflected in spontaneous oscillatory activity (Schutter & van Honk, 2005; Schutter, de Weijer, Meuwese, Morgan & van Honk, 2008). Specifically, we previously showed a positive association between the ratio of relatively slow theta oscillations (4-7 Hz) to beta oscillations (13-30 Hz) and disadvantageous decision making during the (Iowa) gambling task (Schutter & Van Honk, 2005; Massar, Kenemans & Schutter, 2014). During the Iowa Gambing Task (IGT) disadvantageous decisions are associated with large immediate rewards, but in the long run these decisions result in even larger punishments, whereas advantageous decisions are linked to moderate immediate rewards but smaller punishment. It was proposed that the theta/beta EEG ratio is the manifestation of a brain state that promotes reward-drive. However, as indicated, another feature of the IGT as traditionally implemented is that involves a clear reversal aspect, that is, choices that are initially advantageous suddenly become mainly disadvantageous. It is possible that theta/beta EEG ratio reflects a relative inability to adapt to such reversals, rather than reward sensitivity. In the present work we explicitly address this possibility. In the following, we first provide an overview of functional connotations of theta and beta activity separately; after that we provide an integrated perspective featuring the theta/beta EEG ratio.

Spontaneous slow oscillations in the theta range (4-7 Hz) have been characterized neuro-anatomically and –physiologically. Scheeringa et al. (2008) reported a negative correlation between theta power and BOLD

responses in medial-frontal cortex, as well as in a number of other cortical regions. One interpretation suggested by the authors is a general increase in low-frequency EEG power (including theta power) with decreasing BOLD signals across the cortex. While this relation between theta activity and BOLD may reflect biophysical rather than functional aspects, it is strongest in medial-frontal regions. An MFC-based generator was also confirmed by source-localization analysis (Scheeringa et al., 2008).

A large body of work has shown that mid-frontally generated theta activity is linked to activity of the anterior cingulate cortex (ACC), in line with the findings by Scheeringa et al. (2008). Task-related theta activity has been assessed mainly in response to or during stimuli that induce conflict or uncertainty (e.g., Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015; Cavanagh, Zambrano-Vazquez, & Allen, 2012; van Driel, Swart, Egner, Ridderinkhof, & Cohen, 2015; Van de Vijver, Ridderinkhof, & Cohen, 2011). A recent study applying transcranial alternating current stimulation (tACS) suggests a causal relationship, as tACS at theta frequency reduced the performance manifestation of conflict (van Driel, Sligte, Linders, Elport, & Cohen, 2015). Conflict-induced theta activity is generally thought to reflect an error signal carried by disinhibition of medial-frontal-cortex neurons (Cohen, 2014). Such error signals would especially occur when things get more difficult than average or than expected.

Spontaneous theta activity is thought to have a similar error-like MFC disinhibition origin. Another, related perspective is that theta activity is evoked by feedback signals. It scales proportionally to the valence (i.e., stronger for negative than for positive feedback), to the size of the prediction error (i.e., stronger with larger differences between expected and obtained reward or punishment), and also with the learning rate within a task (Mas-Herrero & Marco-Pallarés, 2014).

As noted by Cohen (2014), conflict-related theta activity is a temporary non-phase-locked increase ('burst') in endogenous-theta power, triggered by conflict-detecting neural units in deeper layers of the MFC. Endogenous theta activity is spontaneous and thought to be generated in more superficial layers that also receive inputs from subcortical reward and punishment structures. From this perspective, enhanced spontaneous theta activity would reflect a condition of less reward than expected, even in the absence of discrete signals (such as feedback stimuli) of obtained reward

magnitude. As such it could reflect a relatively enhanced continuous striving to obtain rewards, even in the absence of task-related stimuli that signal the possibility and the subsequent obtainment (or not) of reward. Note that while scalp-recorded theta activity may be driven by subcortical inputs, it is mainly or exclusively a direct reflection of cortical activity.

Beta (13-30 Hz) activity is a type of fast oscillatory activity associated with top down control and decision-making processes (Donner & Siegel, 2011). Increasing evidence indicates that beta oscillatory activity reflects active inhibitory processes involved in maintenance of the current motor and cognitive state (Engel & Fries, 2010). A recent review (Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015) emphasizes the beta response to rewarding events which is "in charge of transmitting a fast motivational signal to downstream brain structures" (p. 4; see also Van den Vijver et al, 2011).

Given these presumed complementary associations between theta versus beta oscillations on the one hand, and reward sensitivity versus reward processing on the other, it seems natural to evaluate theta power relative to beta power. A conceptual underpinning is that theta activity is in part driven by subcortical signals, whereas beta activity represents endogenous cortical activity (Schutter and van Honk, 2005; Schutter, Leitner, Kenemans & van Honk, 2006). Hence, the theta/beta EEG ratio reflects the inverse of cortical regulating signals relative to subcortical to-be-regulated activity. Given that spontaneous theta activity reflects uncertain anticipation of reward, the spontaneous beta rhythm could represent a quenching signal towards subcortical structures that drive the theta activity, signaling that reward anticipation can be toned down in average everyday-life or laboratory conditions. Beta oscillations in this perspective are most prominent across anterior midline sites. They should be distinguished from more lateral beta oscillations that have been associated with error-related sensory-motor adjustments (Luft, Takase, & Bhattacharya, 2014).

In the natural world reward-punishment contingencies are subject to change and as a result individuals may encounter situations with different pay-off schedules. It is critical for the individual to react to such changes in reward-punishment contingencies by rapidly shifting to situationally appropriate decision-making strategies (Clark, Cools, &

Robbins, 2004). Such an adaptation in the situation in which reward-punishment contingencies reverse has been termed reversal learning (Bechara, Damasio, Tranel, & Damasio, 2005). Reversal learning requires that an individual makes an optimal trade-off between exploiting acquired knowledge and exploring other response options that may lead to more profit (Daw, O'Doherty, Dayan, Seymour & Dolan, 2006). Functional neuroimaging research has found that the subcortical reward circuit mediates exploitation in concert with the medial frontal cortex, whereas exploratory decision-making in uncertain environments involves activation of the frontopolar cortex (Daw et al., 2006).

In sum, task-related theta activity might signal the need to adjust the level of cognitive control to optimize behavior during uncertainty, conflict or punishment (anxiety provoking situations) (Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015). This is especially manifest in the association between theta activity and aspects of reversal learning (Mas-Herrero & Marco-Pallarés, 2014), whereas beta activity may represent a quenching signal towards subcortical structures that drive the theta activity. Building on our earlier work (Schutter & Van Honk, 2005; Massar et al., 2014; see above), we ask whether an association exists between endogenous theta/beta EEG ratio and aspects of reversal learning. We previously reported a positive association between theta/beta EEG ratio and risky decisions. The latter could be related to a negative association with the ability to adapt to changing choice-reward/ punishment contingencies. Therefore the present hypothesis is that higher theta/beta EEG ratios are associated with poorer reversal learning. In addition, we explored the relations between self-reported sensitivity to reward and punishment and reversal learning. As prior research has shown that low punishment sensitivity and strong reward dependency are associated with disadvantageous decision-making (van Honk et al., 2002), we anticipated that high reward sensitivity and low punishment sensitivity would be associated with poorer reversal learning.

Methods

Participants

One hundred and thirty-three volunteers participated in the study. Participants were recruited through advertisement at the campus of Utrecht University. Five subjects were excluded because of prior experience with the task. The final sample consisted of 128 participants (mean age: 22.3 years (sd: 3.3 years); 87 females; 122 right-handed). All were unaware of the aim of the experiment and had no prior experience with the task. All subjects were healthy and none of them had a history of psychiatric or neurological conditions and none of them used psychoactive medication. All subjects had normal or corrected-to-normal vision. Participants were requested to abstain from caffeine and smoking on the day of testing. Subjects gave written informed consent and were paid for participation or received study credits instead. The study was approved by the local ethical committee of the faculty of Social and Behavioral Sciences of Utrecht University.

The reversal learning gambling task

The task is based on the Iowa gambling task (Bechara et al., 1994) and the affective reversal learning task by Fellows and Farah (2003). On each trial participants chose one of two squares that contained an amount of money that could either be won or lost. The aim was to win as much fictitious money as possible. On each trial a combination of a high and low value was presented. Participants could either engage in high-risk decision making by choosing the high monetary value or in low-risk decision making by choosing the low monetary value. Eight possible stimulus combinations, namely [5-25], [25-5], [10-30], [30-10], [15-35], [35-15], [20-40], and [40-20] were used and presented vertically and in random order. Participants made a high risk choice by pressing the right mouse button and a low risk choice by pressing the left mouse button. For instance, when presented with [25-5] or [5-25], participants would take high risk by choosing the numeral 25 (i.e., press the right button) over 5 (i.e., pressing the left button). Feedback was provided 500 ms after the subject's decision by coloring the squares either green or red. The amount of fictitious money displayed in the square was either won or lost depending on whether the chosen square

turned green (won) or red (lost). The non-chosen square also colored green or red to provide additional feedback on what would have been the outcome of their alternative choice. Each trial was ended by providing a balance update (i.e., score) that was displayed 2000 ms after feedback onset for a duration of 1500 ms. The inter-trial onset time varied between 800-1200 ms. Figure 1 displays the events sequence of a typical trial.

Figure 1. Typical trial sequence. During the reversal learning task participants either choose a high or low amount. After a 500-ms delay feedback is shown whether the amount has been won (the chosen value turns green) or lost (the chosen value turns red). The amount that has not been chosen also turns red or green to provide additional feedback. Finally, the total score so far is shown. The yellow rectangles indicate the participant's choice. Note that these rectangles are only displayed in this figure for illustrational purposes and are not displayed during the actual task. In this particular example, the participant has chosen the low amount (left button-press) and has lost.

The gambling task consisted of two practice trials and six rounds of twenty trials each and was divided in three phases with a different reward-punishment (R-P) schedule for high-risk decision making. During phase 1 (i.e. round 1 and 2, trial 1-40) choosing the high amount (risk-taking) was rewarded in 80% of the trials. During phase 2 (i.e. round 3 and 4, trial 41-80) the reward-punishment schedule was reversed and choosing the high

amount was only rewarded in 20% of the trials (choosing the low amount was rewarded in 80% of the trials). During phase 3 (i.e. round 5 and 6, trial 81-120) choosing the high amount (risk-taking) was again rewarded in 80% of the trials. Participants were not informed about the reward-punishment schedule.

Resting-state EEG

The Active-Two system (BioSemi, Amsterdam, The Netherlands) was used for recording the resting-state EEG. Thirty-two electrodes were placed and EEG data was sampled at 2048 Hz and a default online low pass filter (DC to 400 Hz) was applied. Four minutes of resting-state EEG was recorded (2 minutes with the eyes open and 2 minutes with the eyes closed).

Reward and punishment sensitivity

Carver and White's (1994) orthogonally-dimensioned behavioral inhibition system (BIS) and behavioral activation system (BAS) self-report questionnaire was used to index the punishment and reward sensitivity of the subjects (van Honk et al., 2002). This questionnaire is derived from Gray's framework of human personality (Gray, 1987), wherein BAS mediates approach behavior in response to cues of reward and BIS is sensitive to cues of punishment and activates avoidance.

Procedure

Upon arrival at the laboratory, participants were informed about the experiment and written informed consent was obtained. Participants were seated in a dimly lit room and first filled in the BIS-BAS questionnaire. A subset (71) of the participants also filled in two other questionnaires that were part of another study. The cap and electrodes were placed and the resting-state EEG was recorded subsequently. Next, participants were subjected to the reversal learning gambling task after they received on-screen instructions. Participants were encouraged to win as much fictitious money as possible. Subjects chose the low or high value by pressing the left mouse button with their left thumb or by pressing the right mouse button with their right thumb, respectively. The duration of the task was approximately 20 minutes. Seventy-one participants were subjected to another task as part of larger study after completion of the reversal learning

gambling task. Resting-state EEG was recorded in a separate session for this group of participants. Twenty subjects of this group were lost to follow-up after the first session. Therefore, no resting-state EEG was recorded for these subjects. Note that for all participants resting-state EEG was recorded before any task performance.

Data reduction and statistical analyses

Percentage high risk-taking for each round was calculated and for each subject polynomial quadratic trend scores were computed for the percentage high risk-taking across the six rounds by using the program DAR (Kenemans, 1991). These quadratic trend scores capture the U-shaped pattern of risk-taking which would be expected for individuals who successfully adapt behavior on the basis of shifts in reward-punishment contingencies (i.e., taking high risk during round 1 and 2, followed by low risk-taking during round 3 and 4, followed by high risk-taking during round 5 and 6). High positive quadratic trend scores represent good reversal learning. Note that this is essentially a regression procedure, as the individual readout measure is a coefficient representing the fit of the individual data to a quadratic (parabolic) model.

Next, we computed difference scores for the percentage high risk-taking representing the adjustment in behavior following a contingency reversal. This was computed for the reward (phase 1) to punishment (phase 2) transition (R-P), as follows: $high\ risk\ phase\ 1\ -\ \%\ high\ risk\ phase\ 2$. Difference scores for the punishment (phase 2) to reward (phase 3) transition (P-R) were computed as follows: $\%\ high\ risk\ phase\ 3\ -\ \%\ high\ risk\ phase\ 2$

Subsequently, these difference scores were normalized for the individual's total percentage risk-taking during the respective phases. These "reversal learning ratios" were calculated for the reward (phase 1) to punishment (phase 2) transition (R-P) and for the punishment (phase 2) to reward (phase 3) transition (P-R).

R-P reversal learning ratio was calculated as follows:

$$\frac{\%\ high\ risk\ phase\ 1\ -\ \%\ high\ risk\ phase\ 2}{\%\ high\ risk\ phase\ 1\ +\ \%\ high\ risk\ phase\ 2}$$

The P-R reversal learning ratio was computed as follows:

$$\frac{\%\ high\ risk\ phase\ 3\ -\ \%\ high\ risk\ phase\ 2}{\%\ high\ risk\ phase\ 3\ +\ \%\ high\ risk\ phase\ 2}$$

Task performance (percentage high risk-taking) on the group level was investigated by testing the average of the individual quadratic trend scores against zero. In case of significance, follow-up paired-samples t-tests between successive rounds were conducted. A final paired-samples t-test was performed to test whether there was a difference between the first and the second reversal learning ratio on the group level.

Raw EEG signals were analyzed offline using Brain Vision Analyzer 2.0 (Brain Products GmbH). EEG data was resampled to 256 Hz and re-referenced to the average reference. Data were divided into two-second segments which were baseline corrected in order to suppress potential DC drifts. An automatic artifact rejection procedure subsequently removed segments containing (ocular/muscle) activity exceeding 50 or -50 μV. On average, out of 120 segments, 114.2 (SD = 1.5) segments remained for Fz, 118.1 (SD = 3.7) remained for Cz and 112.4 (SD = 11.4) remained for Pz. Spectral power in the theta (4-7 Hz), and beta (13-30 Hz) band was estimated by using a fast Fourier transformation (Hanning window: 10%). Spectral power estimates were averaged across segments and theta/beta EEG ratios were calculated for the midline electrodes Fz, Cz and Pz (Schutter & van Honk, 2005; Schutter et al., 2006: Massar et al., 2014). Analyses for the eyes-open and eyes-closed condition separately demonstrated that theta/beta EEG ratio power values for the eyes-open and eyes-closed condition were highly correlated (rho = .821, $p < .001$). Data were therefore collapsed across both conditions.

Topographical plots were obtained for data as analyzed by the steps described above and for the same data re-referenced to the average

mastoids instead of the average reference. The average of electrode P7/P8 was used as an approximation of the signal from the mastoids for 51 subjects for which we did not record from the mastoids. A 1-40 Hz band-pass filter and a subsequent eye blink correction (Gratton et al. method; Gratton, Coles, & Donchin, 1983) were applied to the data re-referenced to the mastoids (P7/P8) in order to ensure sufficient remaining data for the frontal channels. For both reference schemes we excluded individual electrode-channels in case less than half of the data segments (< 60) were left for that channel after artifact rejection. The grand average of the spectral power for each channel was based on data of 75-108 subjects (mean ± SD, 100 ± 8.3) and 102-108 subjects (mean ± SD, 106 ± 1.9) for the average reference and the mastoids reference scheme, respectively.

Table 1 displays average raw theta and beta power values for electrode Fz, Cz and Pz. Figure 2 displays the topographical distribution of theta/beta EEG ratio, theta-, and beta power across the scalp. The top row represents data re-referenced to the average reference and shows that the distribution of the theta/beta EEG ratio was maximal over the mid-frontal cortex, as expected. Distributions of the average referenced theta and beta power separately were maximal over the parieto-occipital cortex. This pattern is a feature of the average reference scheme and has been observed before (Cavanagh et al., 2012). As expected, theta/beta EEG ratio and theta and beta power were all maximally distributed over the frontal cortex, surrounding electrode Fz, when the data were re-referenced to the average mastoids (Figure 2, bottom row). Note that the theta/beta EEG ratio cancelled out the effects of the reference procedure on the scalp distribution.

Table 1

Mean theta and beta power in μV^2

Electrode location	Theta	Beta
	Mean (SD)	Mean (SD)
Fz	0.51 (0.23)	0.08 (0.05)
Cz	0.43 (0.22)	0.07 (0.05)
Pz	0.43 (0.26)	0.07 (0.04)

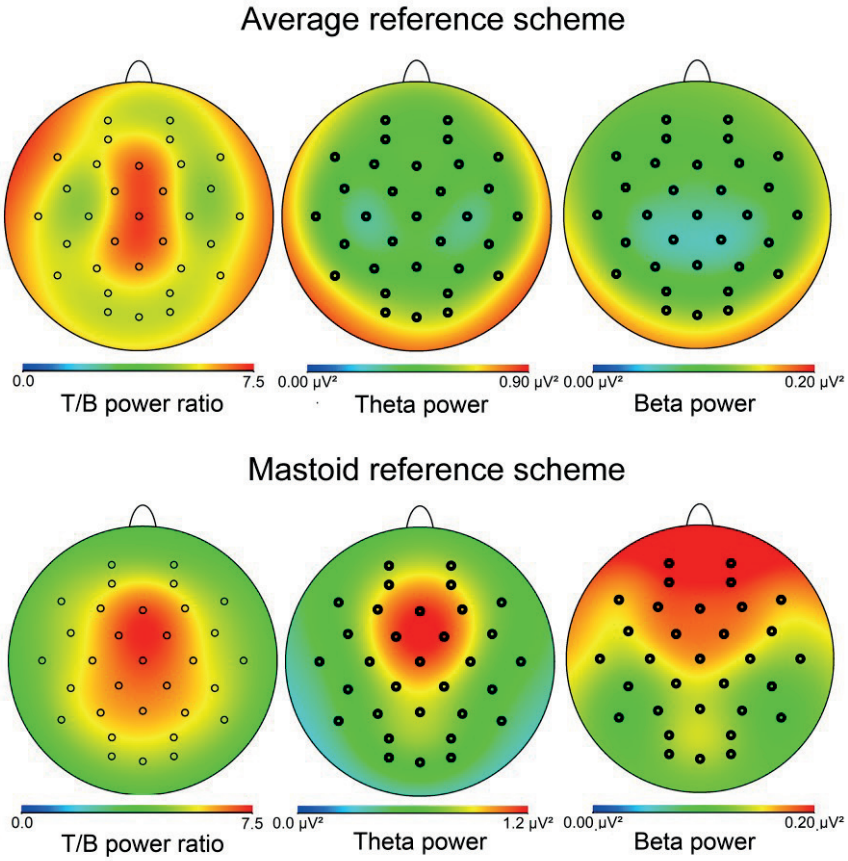


Figure 2. The topographical distribution of theta/beta EEG ratio, theta and beta power across the scalp. Theta/beta EEG ratio was maximal over midline frontal sites (Fz) (above, left). The other figures in the top row represent the topographical distribution of theta (middle) and beta power (right). Theta and beta power plotted separately were maximal over the parieto-occipital cortex when using the average reference. This pattern has also been found in a study by Cavanagh et al. (2012) and is a feature of the average reference. In contrast, the theta and beta power bands (as well as the theta/beta EEG ratio) showed a frontal maximum when the averaged mastoids were used as a reference (bottom row, middle and right figure, respectively). Note that the ratio between theta and beta power cancelled out the effects of the reference method on the scalp distribution of the power values.

Significant Kolmogorov-Smirnov tests indicated that the distributions of the theta/beta EEG ratios at Fz, Cz and Pz deviated from normality. Therefore, Spearman's correlations were used for the correlations involving theta/beta EEG ratios. A series of correlational

analyses were conducted to test our main hypothesis that a high theta/beta EEG ratio measured at Fz is associated with poor reversal learning. We first examined the relationship between theta/beta EEG ratio at Fz and individual quadratic trend scores for the percentage risk-taking across the six rounds. Next, correlational analyses were conducted to test whether the difference scores (behavioral adaptations during the R-P and P-R transitions) were related to theta/beta EEG ratios. We subsequently investigated whether these latter correlations were contaminated by the subject's overall percentage risk-taking by testing the correlation between theta/beta EEG ratios and the R-P and P-R reversal learning ratios. We additionally investigated the relationship between individual quadratic trend scores for the percentage risk-taking across the six rounds and theta/beta EEG ratios measured at Cz and Pz (for which we expected less strong associations in line with our prior studies). We also additionally explored the relationship between theta and beta power separately and reversal learning (i.e., quadratic trend scores).

Finally, Pearson's correlations were conducted to test the relationship between reversal learning ratios and self-reported reward-punishment sensitivity (our second aim).

Alpha level was set to .05 for all analyses, unless stated otherwise. Bonferroni corrections were applied for follow-up paired samples t-tests. Greenhouse-Geisser adjustments were made when appropriate.

Results

A significant quadratic trend for the percentage high risk-taking across the six rounds was found, $F(1,127) = 152.7$, $p < .001$, which confirmed that participants learned to successfully adapt their decision making. Follow-up paired-samples t-tests showed significant increases in high risk-taking in the first phase (round 1 and 2) of the task, $t(127) = -10.49$, $p < .001$. Furthermore, a significant decrease in high risk-taking was observed between round 2 and round 3, $t(127) = 12.35$, $p < .001$. During phase 2 (round 3 and 4) a further decline in high risk-taking was found, $t(127) = 6.66$, $p < .001$, while risk-taking increased again between round 4 and round 5, $t(127) = -14.59$, $p < .001$, and during phase 3 (round 5 and 6), $t(127) = -4.08$, $p < .001$. The significant changes in high risk decision-making between

round 2 and 3 and between round 4 and 5 further demonstrate that participants learned to adapt behavior on the basis of shifts in reward-punishment contingencies. Figure 3 shows the percentage risk-taking during each round and reversal learning across the task. Note that all p-values were below the Bonferroni-corrected threshold for significance of p = .01.

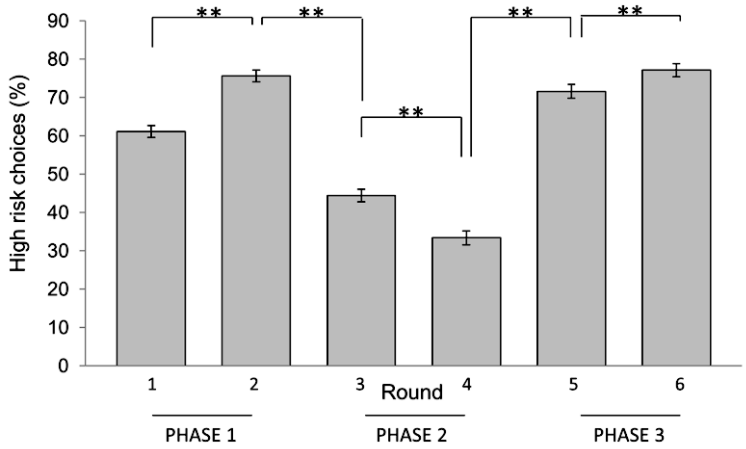


Figure 3. Percentage of high-risk choices during each round. Error bars represent ±1 SE. The quadratic trend for the percentage high risk-taking across the six rounds was highly significant ($p < .001$), indicating that, on average, participants learned to adapt their behavior.

A significant negative correlation was observed between theta/beta EEG ratio measured at Fz and the individual quadratic trend scores for the percentage high risk-taking across the rounds (a high positive score means good learning), rho = -.306, p = .001. These results were confirmed by significant negative correlations between theta/beta EEG ratio measured at Fz and percentage high risk difference scores (behavioral adaptations after a contingency reversal), rho = -.295, p = .002 (phase 1-2) and rho = -.286, p = .003 (phase 2-3). These correlations remained significant when the normalized difference scores (i.e., reversal learning ratios) were used instead, rho = -.285, p = .003 (phase 1-2 transition) and rho = -.277, p = .004 (phase 2-3 transition). Together these results indicate that participants with high theta/beta EEG ratio were less inclined to change to more

adaptive decision making when reward-punishment schedules were reversed. See figure 4A (phase 1-2 transition) and 4B (phase 2-3 transition).

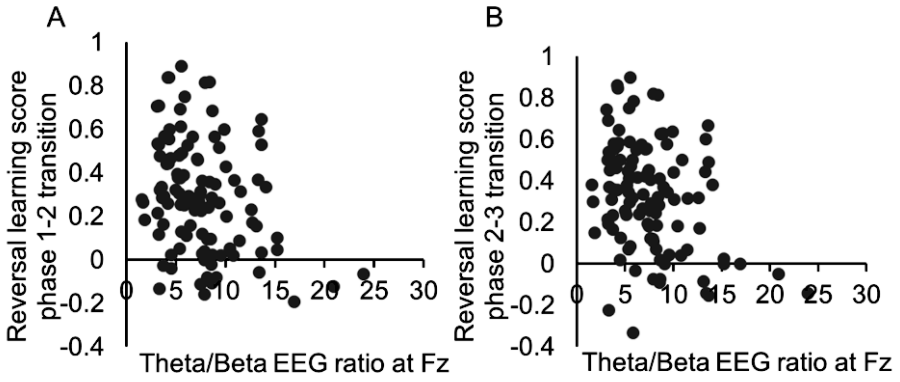


Figure 4. Reversal learning scores correlate negatively with theta/beta EEG ratio. A reversal learning score of 1 represents perfect learning. Scatterplots are shown for reversal learning during the phase 1-2 transition (panel A) and 2-3 transition (panel B).

In addition, a paired samples t-test comparing reversal learning ratios between the first and the second reward-punishment transition revealed significantly larger reversal learning ratios for the second transition, mean ± SD = 0.32 ± 0.27, compared to the first, mean ± SD= 0.29 ± 0.25, $t(127) = -3.27$, $p = .001$.

We also tested whether theta/beta EEG ratio measured at electrode Cz and Pz similarly predict reversal learning. Significant negative correlations between theta/beta EEG ratio and individual quadratic trend scores for the percentage risk-taking were also observed for the electrodes Cz and Pz, rho = -.243, p = .011, and rho = -.316, p = .001, respectively.

Next, we examined whether the relationship between theta/beta EEG ratio at Fz and reversal learning was explained by risky decision making during phase 2 or by low risk-taking during phases 1 and 3. Therefore, correlations were computed between theta/beta EEG ratio at Fz and the percentage risk-taking during each of the three phases. Interestingly, this analysis revealed that individuals with a high theta/beta EEG ratio made less risky decisions during both reward phases, rho = - .253, p = .008 (phase 1), rho = - .213, p = .027 (phase 3). There was also a

significant positive correlation between theta/beta EEG ratio and percentage risk-taking during phase 2, rho = .216, p = .025. Figure 5 displays the percentage risk-taking during each of the six rounds for participants with a low (below the median) and high (above the median) theta/beta EEG ratio separately.

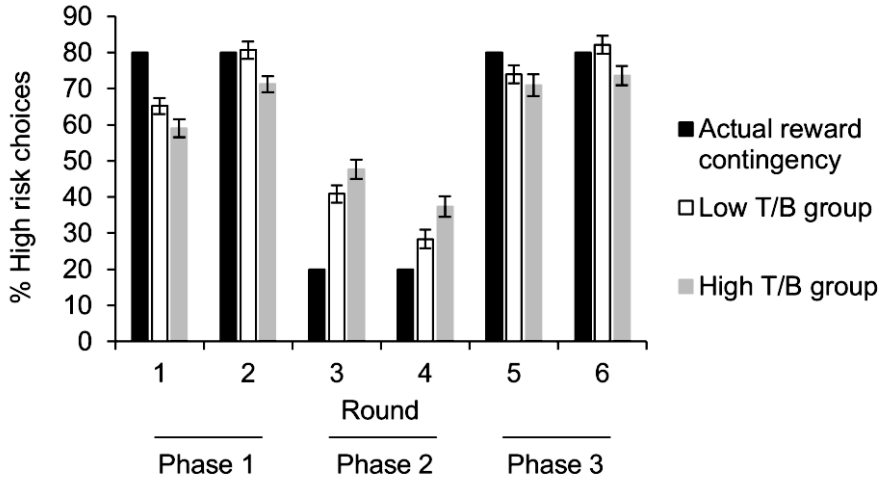


Figure 5. Percentage of high-risk choices during each round for the high and low theta/beta EEG ratio group. Participants were categorized as having either a low (white bars) or a high (grey bars) theta/beta EEG ratio, based on a median-split. The figure represents the percentage high-risk choices during each of the six rounds for both groups separately, relative to the actual reward contingency during each round (black bars). Error bars represent ±1 SE.

Percentage high risk-taking in phase 1 and 3 was furthermore found to correlate negatively with percentage high risk-taking in phase 2, rho = -.199, p=.024; rho = -.189, p =.033, respectively, indicating that individuals exhibiting low risky behavior in phase 1 and 3 are for a large part the same individuals exhibiting high risky behavior in phase 2. These correlations were non-significant (p's > .116) when controlling for theta/beta EEG ratio.

A significant positive correlation between theta and beta power was found, rho = .516, p < .001. However, neither of these measures correlated with the individual quadratic trend scores for the percentage high risk-taking across the rounds, rho = -.056, p = .564 (theta power) and rho = .174, p = .07 (beta power). Results of Steiger tests (1980) indicated that the

strength of the latter correlations, and the correlation between theta/beta EEG ratio and the quadratic trend scores were significantly different; $Z = 4.55$, $p < .01$ (theta/beta EEG ratio versus theta power) and $Z = 3.62$, $p < .01$ (theta/beta EEG ratio versus beta power). These findings show that the theta/beta EEG ratio explains unique variance in reversal learning.

No significant correlations were observed between self-reported reward and punishment sensitivity and reversal learning ratios (p's $> .4$). However, a significant negative correlation between self-reported punishment sensitivity and theta/beta EEG ratio was observed, rho = -.211, $p = .028$. Finally, the correlation between theta/beta EEG ratio and self-reported reward sensitivity was not significant ($p = 883$).

Discussion

The primary aim of the current study was to investigate whether the theta/beta EEG ratio is associated with reversal learning in an environment with changing reward-punishment contingencies. Results showed that participants with a high theta/beta EEG ratio were less able to adapt to the change in reward-punishment contingency.

The negative correlation between theta/beta EEG ratio on the one hand and reversal learning on the other hand is consistent with our predictions. In line with the findings by Schutter and colleagues (2005), we theorized that the theta/beta EEG ratio reflects the inverse of cortical regulating signals relative to subcortical to-be-regulated activity (reward drive). Theta/beta EEG ratio was associated with risky decision-making in our previous studies, and the latter was hypothesized to be negatively related to the ability to adapt to changing choice-reward/ punishment contingencies. Hence, participants with high theta/beta EEG ratio would be less able to flexibly switch between decision-making strategies during the task and less able to adjust behavior. Reversal learning in an environment in which previously rewarded actions suddenly have opposite outcomes requires that individuals switch between exploitation of known responses and exploration of alternatives (Daw et al., 2006). The results on the current task showed that, on average, individuals were able to adapt their behavior both within and between the reward-punishment contingency phases of the

reversal learning task. However, individuals with relatively high theta/beta EEG ratio were less able to do so.

Notably, participants with high theta/beta EEG ratio made in fact less risky decisions when high risk-taking was rewarding. At first sight this finding seems at odds with evidence from prior studies that supports a relationship between increased theta/beta EEG ratio and high risk-taking and approach behavior (Massar et al., 2014; Schutter & van Honk, 2005). Results from these studies show that subjects with a high theta/beta EEG ratio keep choosing from high-risk decks during the IGT, while low theta/beta EEG ratio subjects gradually learn which deck is most beneficial and adapt their choices accordingly.

The present results suggest that theta/beta EEG ratio specifically reflects the ability to adapt choices in response to changing contingencies, and that this may be what drives the association between theta/beta EEG ratio and optimal performance also in the traditional implementation of the IGT. In accordance with our findings, a recent double-blind randomized controlled study applied 5-Hz transcranial alternating current stimulation (tACS) to the frontal cortex, which improved reversal learning in healthy volunteers. Results showed that even though volunteers improved on learning ability, they were less inclined to actually change their risk taking accordingly. Notably, EEG recordings showed a significant lowering of spontaneous theta/beta EEG ratios (Wischnewski, Zerr, & Schutter, 2016).

A large body of empirical work has shown that mid-frontally generated theta oscillations are elicited during signals of punishment and conflict (Cavanagh, Figueroa, Cohen & Frank, 2011; Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015; Cohen, Ridderinkhof, Haupt, Elger, & Fell, 2008). More specifically, mid-frontal theta oscillations are thought to signal reward prediction error signals that originate from the subcortical dopamine system and are elicited when outcomes are worse than expected (Holroyd and Coles, 2002). Cavanagh and colleagues (2011) also showed that theta activity was enhanced when participants were uncertain about their responses during a probabilistic reward task. These theta signals during uncertainty seem particularly pronounced in high trait-anxious individuals (Cavanagh & Shackman, 2015) and are thought to signal the need for cognitive control (Cavanagh & Frank, 2014) or behavioral adaptation (Cavanagh, Frank, Klein & Allen, 2010).

It should, however, be noted that these relations pertain to task-related or induced theta signals. These studies may suggest that task-related theta signals are associated with better reversal learning. The present results, however, demonstrate that high spontaneous theta/beta EEG ratio is associated with less ability for adaption after a reversal. One explanation for this apparent discrepancy is that high spontaneous theta activity is associated with a reduced theta response to task demands or stimuli that prompt adaptation. Even though one study found an inverse relation between pre-stimulus theta power and stimulus induced theta power (Klimesch et al., 2004), to the best of our knowledge there are no studies available that have directly addressed this issue. Note that pre-stimulus theta activity within a task context cannot be equated with resting-state spontaneous theta activity. Another clue was more recently provided by Massar et al. (Massar, Rossi, Schutter, & Kenemans, 2012). These authors reported a negative correlation between resting-state theta activity and the feedback-related negativity (FRN) to negative feedback stimuli, albeit only in individuals with high punishment sensitivity. The FRN is generally considered to be an evoked-theta-dominated response (Cavanagh et al., 2012), and it overlaps with induced theta activity with respect to the association with reversal-learning aspects (Mas-Herrero & Marco-Pallarés, 2014). Yet, a direct comparison between resting-state theta activity and task-related theta responses has not been made. Note that we also found a negative correlation between theta/beta EEG ratio and self-reported punishment sensitivity. Conceptually, this relation is also consistent with the idea of a negative relation between resting-state theta (/beta) activity and feedback-induced theta oscillations (e.g., FRN elicited by punishment).

As already mentioned, participants with a high theta/beta EEG ratio showed a pattern of low risk-taking during phase 1 and 3 of the current task. High risk-taking was rewarded in 80% of the cases during these phases. However, still 20 % of the high-risk choices led to a loss of money. Apparently, in high theta/beta EEG ratio individuals, high risk loss may have promptly resulted in low risk taking, whereas it would have been more adaptive to use reward and punishment feedback information over a larger number of trials. This could indicate that endogenous high theta activity correlates with increased prediction errors after high-risk losses (i.e., under uncertainty). This would in fact predict that endogenous high theta relative

to beta activity is associated with larger negative-feedback-induced theta oscillations, which is contrary to the notion presented in the previous paragraph. This apparent contradiction should be addressed in future research including assessment of feedback-induced theta in reversal-learning gambling contexts.

In addition, endogenous low beta relative to theta activity may be associated with a lack of prefrontal cortical control over decision-making strategies (exploit versus explore) during the task. Decreased cognitive flexibility (i.e., low beta activity) together with uncertainty and large prediction errors after high-risk losses (i.e., high theta activity) may have caused a delay in learning reward-punishment contingencies in participants with high theta/beta EEG ratio.

Furthermore, the correlation between theta/beta EEG ratio and the quadratic reversal learning pattern explained significantly more variance than the correlation for theta and beta power separately. These latter tests revealed no significant results. This finding indicates that theta/beta EEG ratio explains unique variance in reversal learning. However, since the separate tests for theta and beta power were not part of our original hypothesis, these findings should be replicated in an independent sample.

Although it was expected that reversal learning would be predominantly associated with theta/beta EEG ratio recorded at the frontal electrode (Schutter et al., 2006: Massar et al., 2014), the current results indicate that the central and parietal theta/beta EEG ratio also predicts reversal learning. This finding concurs with a prior study in which a relationship was observed between disadvantageous decision-making and theta/beta EEG ratio measured at the mid-frontal (Fz) and parietal sites (Pz) (Schutter & van Honk, 2005).

A second aim of the present study was to explore whether self-report measures of reward and punishment sensitivity predict reversal learning during the RLG task. The expected inverse correlation between reward sensitivity (BAS) and reversal learning, and the positive correlation between punishment sensitivity (BIS) and reversal learning were not found. We did, however, observe a significant negative relationship between self-reported punishment sensitivity and theta/beta EEG ratio.

Our study leaves open a number of issues for further investigation. First, our design cannot differentiate between learning the reward-

punishment contingency and executing the correct strategy. Second, our results raise the issue of alternative reward-punishment contingencies for high risk choices (e.g., R:P 60:40 and 40:60 percent) on reversal learning and its relationship with theta/beta EEG ratio. Third, the RLG task always started with a phase during which 80 percent of the high risk choices were rewarded (R:P 80:20). This leaves open the question of whether starting with the alternative R:P contingency phase (i.e., 80 percent punishment for high-risk choices) would yield comparable results. Context-related differences such as task offset may be important to investigate given that, for example, patients with orbitofrontal cortex damage show reduced learning during the IGT compared to controls, but *only* when the first cards of the high-risk deck consist of wins (Fellows, 2007). Fourth, a relation between risk taking and theta asymmetry rather than overall power has also been reported (e.g., Studer, Pedroni, & Rieskamp, 2013), prompting the question if there are similar relations between risk taking and/ or reversal learning and theta/beta EEG ratio asymmetry.

In conclusion, one's ability to adapt to changing reward-punishment environments by adjusting behavior on the basis of shifts in emotional significance is vital for behavioral flexibility in changing environments (Clark et al., 2004). The present study demonstrates that individuals with an increased ratio between low frequent oscillations in the theta range and high frequent oscillations in the beta range during resting state exhibit lower levels of behavioral flexibility. This was reflected by a reduced ability to respond adaptively and to adjust behavior after a reversal of reward contingencies. Higher levels of theta/beta EEG ratios were associated with poorer reversal learning, which is arguably due to a decreased ability to learn which choice was more likely to yield a reward.

Chapter 5

Disentangling the effects of reward value and probability on anticipatory event-related potentials

Iris Schutte
Ivo Heitland
J. Leon Kenemans

Neuropsychologia (2019)

Author contributions

LK and IS conceptualized and set up the study; IS programmed the task; IS collected the data, IS performed the analyses with input from LK and IH, and wrote the initial draft of the manuscript; all authors contributed to the final manuscript; LK supervised the study.

Abstract

Optimal decision-making requires humans to predict the value and probability of prospective (rewarding) outcomes. The aim of the present study was to evaluate and dissociate the cortical mechanisms activated by information on an upcoming potentially rewarded target stimulus with varying probabilities. Electro-cortical activity was recorded during a cued Go/NoGo experiment, during which cue letters signaled upcoming target letters to which participants had to respond. The probability of target letter appearance after the cue letter and the amount of money that could be won for correct and fast responses were orthogonally manipulated across four task blocks. As expected, reward availability affected a prefrontally distributed reward-related positivity, and a centrally distributed P300-like event-related potential (ERP). Moreover, a late prefrontally distributed ERP was affected by probability information. These results show that information on value and probability, respectively, activates separate mechanisms in the cortex. These results contribute to a further understanding of the neural underpinnings of normal and abnormal reward processing.

Introduction

Optimal decision-making requires humans to predict the value and probability of prospective (rewarding) outcomes (Glimcher & Rustichini, 2004). These predictions have direct implications for subsequent behavior, which is based upon cortical activity. In the present study we evaluate the cortical mechanisms activated in a context of anticipating reward with varying probabilities. We furthermore investigate the extent to which these cortical activations interact and to which they are independent.

Neuroimaging studies have investigated the representation of anticipated subjective value (or reward) in the human brain. Parts of the ventro-medial prefrontal cortex including the medial orbito-frontal cortex (mOFC) and the rostral anterior cingulate cortex (rACC) (Breiter et al., 2001; Howard et al., 2015; Kable & Glimcher, 2007; Padmala & Pessoa, 2011; Smith et al., 2009), as well as more posterior regions of the cingulate cortex (Kable & Glimcher, 2007; Kirsch et al., 2003; Padmala & Pessoa, 2011; Smith et al. 2009) show activity during the anticipation of reward. These cortical regions act in concert with subcortical structures, such as the ventral striatum and midbrain (e.g. Breiter et al., 2001; Knutson et al., 2005; Smith et al., 2009; Yacubian et al., 2006; for reviews see: Haber & Knutson, 2010; O'Doherty, 2004; Rushworth & Behrens, 2008). Some of these studies also investigated the effect of increasing the anticipated probability of a reward and reported affected regions in the posterior cingulate cortex (PCC) (Knutson et al., 2005) and medial prefrontal cortex (Knutson et al., 2005; Yacubian et al., 2006), as well as in the dorsal ACC (dACC) (Smith et al., 2009).

Results of the neuroimaging study by Knutson and colleagues (2005) indicate that the dACC may integrate reward and probability information. However, in that study as well in the other neuroimaging studies discussed above, main effects of reward value and probability manipulations abound as well. This prompts the crucial question of the temporal relations between activity related to either the value or probability manipulation on the one hand, and activity related to their interaction on the other. The event-related potential (ERP) technique has a much higher temporal resolution compared to neuroimaging. It is therefore

much better suited in temporally separating sequential neural activities within the brief time windows that characterize real-life decision-making.

Recently, a number of studies have investigated the effects of reward value (but not so much probability) on ERPs during outcome anticipation. It was found that reward cues (highly) predictive of an upcoming monetary reward (monetary incentive delay task: Doñamayor et al., 2012; Flores et al., 2015; passive gambling task: Holroyd et al., 2011) elicit a larger positivity (or less negative activity, Yu & Zhou, 2006) than cues (highly) predictive of no reward or a loss, with a latency between 200-300 ms after the reward announcing cue and a fronto-central scalp distribution. A reward-sensitive ERP with a similar latency was observed during choice presentation (so before the feedback stage) in a gambling task after subjects learned which of the choice options yielded reward (Krigolson, Hassall, & Handy, 2014). This reward-related positivity has been labelled "reward positivity" (Holroyd et al., 2011). It is observed not only in response to reward-predicting cues but also in response to reward delivery (Holroyd et al., 2011). Especially in the latter context it has also been described as a mirror inverse of the feedback-related negativity (FRN; Krigolson, 2018; Proudfit, 2015). Also in the context of negative feedback and errors, cues predicting errors or non-reward have been observed to elicit a larger FRN/ error-related negativity (ERN) relative to cues predicting correct responses or reward (Baker & Holroyd, 2009; Krigolson & Holroyd, 2007).

With respect to probability, a prior study by our group (Bekker et al., 2004) showed that cues highly predictive of an upcoming target, and therefore containing highly relevant information, elicit a larger posterior P300 ERP than cues less predictive of an upcoming target. This finding is line with a large body of research showing sensitivity of the P300 to internal updating of the subjective probability of relevant events or outcomes (Duncan-Johnson & Donchin, 1982). In another study, probability was manipulated within a reward-anticipation context (Yu et al., 2011). Here, cues signaling a 100% certain future reward elicited a smaller FRN/ larger RRP, relative to cues signaling less than 100% certainty (ranging from 0 to 87.5%).

ERP studies investigating the effect of both reward value *and* probability manipulations during reward anticipation are scarce. Furthermore, it remains to be determined whether ERPs elicited by reward

value and probability manipulations are dissociable. In the current study, reward value and probability were orthogonally manipulated across four task blocks during a cued Go/NoGo experiment (CGN task) (adapted from Bekker et al., 2004) in which cues signaled upcoming targets to which participants had to respond. Unlike paradigms in which performance was based on choices between different reward-probability conditions (e.g. Krigolson et al., 2014; Smith et al., 2009; Yacubian et al., 2006), in the present study cued reward value and probability information was decoupled from task requirements such as choosing a response. This enabled us to isolate specific activations related to reward value and probability from those of task requirements. For example, when a certain response choice is directed at an option of a certain reward being obtained with a certain probability, the probability level directly affects the expected reward value. In our design, probability only concerns whether the action will have to be performed at all. This contributes to strong orthogonality, while at the same time reward obtainment is still dependent on producing the adequate response. This latter feature is important at least with respect to the FRN/ERN (Yeung et al., 2005).

The main aims of the current study were: (1) to gain understanding of the temporal profile of cortical mechanisms activated in a reward anticipation context; (2) to investigate the extent to which activations related to reward value and probability manipulations interact when both are completely orthogonally manipulated. Our main focus was on anticipatory activity within a 180-500 ms post cue window (see Methods section)[2]. Specifically, we tested the hypothesis that reward value affects frontal ERP activity early in time. This ERP activity could be associated with the processing of reward itself, or more with indirect effects of reward on attention networks (Corbetta & Shulman, 2002). Probability manipulations were expected to specifically affect parietal ERP activity later in time (P300) (Bekker et al., 2004; Holroyd et al., 2011; Doñamayor et al., 2012; Flores et al., 2015). We also anticipated the possibility of an interaction between reward value and probability. Reward value could have an additive effect on the probability P300 ERP, given that numerous studies have shown that

[2] As per medical-ethical protocol (<NL39997.041.12>, available from the author on request), this research was directed at the cue-induced cortical activations within this latency window, not at possible later processes such as stimulus-preceding negativity, nor at target-elicited activations

the P300 is also sensitive to reward outcome (e.g. Yeung & Sanfey, 2004) and reward anticipation (Broyd et al., 2012; Flores et al., 2015; Pfabigan et al., 2014). Alternatively, reward value and probability could interact like in the Knutson et al. (2005) study. In this scenario, reward announcing cues were expected to elicit more ERP activity than no reward cues, but only when they also predict high probability of an upcoming target.

It should be noted that while our design ensures orthogonal manipulation at block level of reward value and probability, this does not hold at the level of single trials, as cues indicating high-probability rewarded targets also cue low single-target rewards. An 'adaptive scaling' (see Walsh & Anderson, 2012) perspective predicts a response to low single-target reward cues (cueing 98% probability reward), when the low value is the only available option (in addition to no reward, throughout a block of trials), that is identical to the response to high single-target reward cues (cueing 50% probability reward) when the high value is the only available option. From this perspective, reward values during different probability conditions could be validly compared. To assess the extent to which this perspective is tenable, we performed additional analyses using specific contrasts to isolate 'pure' reward and probability effects (see Methods-statistical analysis).

Methods

Subjects

Forty-nine healthy subjects participated in the experiment. Participants were recruited via advertisement at the campus of Utrecht University. None of the subjects had a history of psychiatric or neurologic disorders and none of the subjects used psycho-active medication. Participants were requested to abstain from consuming caffeine and smoking for at least 12 hours prior to participation and were requested to refrain from drugs for at least 2 weeks prior to participation. All participants declared to have normal or corrected-to-normal vision. The study was approved by the medical ethical committee of the University Medical Centre Utrecht and subjects gave written informed consent prior to participation. Participants received 6 Euros per hour or received study credits instead, and additionally received a monetary bonus with a maximum of 10 euros. The monetary bonus was

dependent on task performance (see Cued Go/NoGo task). ERP data of 1 participant were not stored due to a technical issue. Furthermore, 3 participants were excluded during analysis, because too few segments were left in one or more conditions for multiple neighboring electrodes (see data processing). Therefore the final sample consisted of 45 participants (mean age (SD) = 23.9 (4.2) years, 34 females, 43 right-handed).

Procedure

This experiment was part of a larger study (3 sessions on separate days) on the effect of reward and target probability (anticipation) on various aspects of behavior, and of psycho- and neurophysiology. Participants were informed about the experimental procedure and signed the informed consent form during the first session. Half of the subjects completed the cued Go/NoGo (CGN) task during the second session and the other half during the third session.

The CGN task session started with placement of cap and electrodes. Participants were seated in a chair one meter in front of a computer screen in a dimly lit room adjacent to the control room with the chin placed on a chin-rest. Participants fixated on the center of the screen and the chair was adjusted to a comfortable height accordingly. Task instructions were given and the CGN task started subsequently, which lasted about one hour. EEG was recorded during the task. Five subjects completed a spatial cuing task before (3 subjects) or after (2 subjects) the CGN task. The other participants were not subjected to other tasks during the CGN task-session. At the end of the test session the cap and electrodes were removed and participants were paid and dismissed.

Cued Go/NoGo task

The CGN task was controlled by Presentation® software (version 16.0, www.neurobs.com). During the CGN task (adapted from Bekker et al. 2004), the letters A, C, D, E, F, G, H, J, L, X and Y were presented in the center of a 16 inch Dell CRT screen (resolution: 1280 x 1024) in black on a grey background between two vertical bars (height: 1.03º, width: 0.05º). Letters were presented in Arial font, size 79. The letter stimuli were presented for 150 ms and were interleaved by inter-stimulus intervals with a random duration between 1400 and 1600 ms.

The task is illustrated in Figure 1. Participants were instructed to press the left button with the left index finger when letter X followed letter A and to press the right button with the right index finger when letter Y followed letter A, as fast and accurately as possible. This mapping was reversed for half of the participants. Responses were made on a qwerty keyboard[3] on which all keys were covered by a plastic sheet, except for the "z" key, "/" key and the spacebar. The "z" and "/" key were the left and right target button, respectively, and the symbols on these buttons were covered by a white sticker.

Target probability was either 50% (total of 40 targets, see below) or 98% (78 targets) and either no money or a maximum of 5 Euros in total could be won during each block. Participants were fully informed about the probability and the reward levels for a block before the start of that block. Note that although we manipulated target probability (i.e., probability of X or Y following letter A) and target reward value, we were specifically interested in the electrocortical aspects of reward and probability anticipation as they occur after the cue (i.e., letter A) but before the target (letter X or Y).

The amount of money won in the reward block was calculated by multiplying the percentage of correct and timely responses by 12.5 eurocents (5 Euro divided by 40 trials) and 6.4 eurocents (5 Euro divided by 78 trials), respectively.

The task started with 100 practice trials (letters). The practice block always consisted of the reward-98% target probability condition. The main task consisted of four blocks of 400 trials (letters), comprising the four conditions. Participants were informed about the target probability and the total amount of money that could be won during the block at the beginning of each block. Order of the four blocks was counterbalanced across participants. One-minute rest breaks were provided halfway through each block and participants were reminded of the target probability and reward availability of the current block after the rest breaks. One-minute rest breaks were also provided between blocks.

[3] First a response box was used, but after 9 subjects the response box was replaced by a keyboard due to a technical issue.

The cue (A) appeared 80 times during each block. In the 50% target probability blocks, 20 cues were followed by target X and 20 cues were followed by target Y. Forty cues were not followed by a target and 20 X's and 20 Y's were not preceded by a cue. In the 98% target probability blocks, 39 cues were followed by target X and 39 cues were followed by target Y. Two cues were not followed by a target and 1 X and 1 Y were not preceded by a cue. Each of the other letters appeared 20 times in each block, except for letter H and C, which appeared more often (80 and 40 times, respectively), in order to control for frequency differences between the letter stimuli and cues/targets (Bekker et al., 2004). Letter stimuli were presented in a pseudo-random order within each block, with the following restrictions: (1) stimuli were never directly followed by identical stimuli. (2) cues followed by targets (A-X or A-Y) and cues not followed by targets ("NoGo": A-not X or Y) were always followed by at least one "no cue" (i.e., C, D, E, F, G, H, J, L not preceded by a cue). Only cue- and no cue-related ERP activity was analyzed. Note that nocues were not associated with reward or probability information.

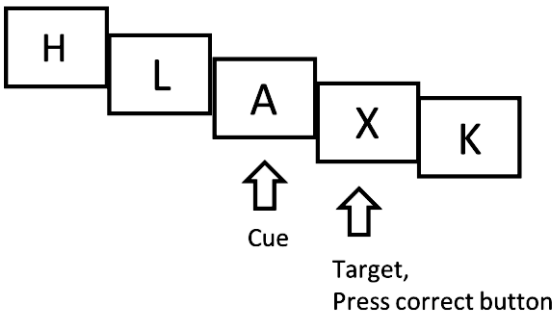


Figure 1. Overview of the cued Go/NoGo task. Letters were presented on the screen and participants were instructed to press a pre-specified button when letter X (target) followed letter A (cue), and when letter Y (target) followed letter A (cue). Four blocks of letter trials were presented, which differed in the amount of money that could be won for correct and fast responses (either 0 or 5 Euros maximally in total) and in the probability of target appearance after the cue (either 50% or 98%).

Nocue related activity was used as a baseline for non-specific effects, as the context of reward/high probability within the current block may sensitize processing. Subtraction of nocue from cue-related activity,

therefore, yields ERP activity specifically related to the temporally specific information on reward and probability (i.e., the cue A signaling that with 50 or 98% probability a reward adding up to 5 Euros could be earned in the reward condition, or of 0 Euro in the no-reward condition).

To briefly recapitulate our hypotheses: We expected cues to elicit a reward-related positivity (RRP) and P300, relative to no-cues (these were either C, D, E, F, G, H, J or L). Furthermore, we expected reward (vs. no reward) to enhance RRP and P300, and high (vs. low) probability to enhance P300. Cues and no-cues were presented pseudo-randomly within a block of trials, and the resulting average no-cue ERP served as to be subtracted baseline for the resulting average cue ERP. This was done to control for non-specific effects of reward and probability that would affect any response to any stimulus in a given block, including irrelevant probes. In total four of these blocks were presented, corresponding to four conditions that resulted from a 2x2 design based on orthogonal manipulation of the probability of target appearance (letter X or Y) after the cue (letter A), and the amount of money that could be won for correct and fast responses.

EEG data acquisition

ERP signals were recorded with the Active-Two system (Biosemi, Amsterdam, The Netherlands) with 64 Ag-AgCl electrodes. Recording electrodes were placed according to the 10/10 system. EOG electrodes were placed above and below the left eye and at the outer canthi of both eyes. EEG signals were online referenced to the Common Mode Sense/Driven Right Leg electrode. EEG data were sampled at 2048 Hz and online low pass filtered at DC to 400 Hz.

Data reduction and analysis

Behavioral data

Mean reaction times (RTs) for valid responses to the target (i.e., single responses within the time window 150-1500 ms after target onset) were calculated for each condition and each subject. Furthermore, the percentage correct responses and percentage omissions were calculated for each condition and each subject. The percentage commission errors to the NoGo stimulus (i.e., a non-target preceded by a cue) was calculated for each

subject only for the 50% target probability blocks, because there were too few NoGo trials during the 98% target probability blocks.

ERP data

ERP data collected during the CGN task were analyzed using Brainvision Analyzer 2.0 (Brain Products GmbH). Data were re-referenced to the average reference, were filtered with a 30 Hz low pass filter (24 dB/oct) and an additional 50 Hz Notch filter, and re-sampled to 256 Hz. Data were segmented into windows from 100 ms before (no)cue onset until 1000 ms after (no)cue onset. Cue-locked segments with pre-mature (< 150 ms) or late responses to the target (> 1500 ms) or with omissions, choice errors, or commission errors were removed from further analyses. Ocular artifacts were corrected by using the Gratton & Coles method (Gratton et al., 1983) and a baseline correction was applied subsequently by using the 100 ms time window before cue onset. Channels were individually inspected for segments with artifacts by using an automatic artifact rejection procedure (maximal allowed absolute difference between two values: 100 μV, lowest allowed activity within a 100 ms interval: 0,5 μV). 1.8 (± 2.1) % of the data segments were lost on average due to the artifact rejection procedure.

Channels with less than half of the segments left within a particular subject/condition were interpolated with a spherical splines method (Brainvision Analyzer 2.0, Brain Products GmbH) using the neighboring electrodes. Data of three subjects were removed from further analyses, because less than half of the cue- (< 40) or nocue-locked (< 100) segments were left in one or more conditions for multiple neighboring electrodes. For 12 subjects data of one or more electrodes within one or more conditions were interpolated (See the table in section 2 of the Supplementary materials). For each subject and condition the average cue-nocue waveforms were computed from -100-700 ms around the cue. The border of the segment was set to 700 ms in order to limit the number of factors obtained with the PCA.

A principal component analysis (PCA) was conducted as this technique allows to separate possibly overlapping ERP components sensitive to reward, probability, or both in terms of spatial distribution and timing. A temporo-spatial PCA was conducted following the guidelines by Dien (2012) and by using the ERP PCA toolkit version 2.63 (Dien, 2010).

Promax rotation with Kaiser loading weighting was used for the initial temporal PCA and nine factors were retained based on a Scree plot (Cattell, 1966). Subsequently, a separate spatial PCA with Infomax rotation was conducted for each of the temporal factors. Five spatial factors were retained for each temporal factor based on the Scree plot averaged over all temporal factors. Both PCA steps were based on the covariance matrix. These steps yielded 45 temporo-spatial factor combinations (TFSF). Based on prior studies (see Introduction) we expected the reward and probability effects to be strongest surrounding the midline of the scalp. The effects were expected to emerge between approximately 180-500 ms post cue onset. Based on recommendations by Dien (2012), PCA factors were selected for statistical analysis in case: (1) they explained more than .5% of the total variance, (2) the temporal loading peaked between 180-500 ms, and (3) the positive voltage was maximal around the midline electrodes. Eight of the 45 TFSF combinations met these criteria. Exploratory analyses were conducted for factors with a temporal loading outside the 180-500 ms post-cue window (i.e., within 0-180 ms or 500-700 ms post-cue), and for factors with a more lateral spatial distribution. This pertained to an additional 10 TFSF factors.

Standard ERP analysis

For comparability with earlier and future studies we supplement the PCA results with results of a standard ERP analysis. Grand-average ERP waveforms for selected midline electrodes are depicted in the results section. The methods and statistical results of the standard analysis are provided in the supplementary materials section.

Statistical analyses

Behavioral data

Repeated-measures ANOVAs (GLM, SPSS version 22) were run for RT, the percentage correct responses, the percentage commission errors to the NoGo stimulus, and the percentage omissions with reward availability (no reward, reward) and target probability (50%, 98%) as within-subject variables. For each contrast (i.e., reward-no reward, high-low probability, and reward effect high-reward effect low probability) deviation from normality was tested using Shapiro-Wilk's tests. Non-parametric Wilcoxon

signed-rank tests were conducted for those contrasts that deviated significantly from normality.

ERP data – PCA

Reward x probability ANOVAs (GLM, SPSS version 22) were run on the TFSF combinations using the average of a 20-ms window around the peak of the factor tested. Alpha was set at .05. For each contrast (i.e., reward-no reward, high-low probability, and reward effect high-reward effect low probability) deviation from normality was tested using Shapiro-Wilk's tests. Non-parametric Wilcoxon signed-rank tests were conducted for those contrasts that deviated significantly from normality. The ten exploratory analyses were corrected for multiple comparisons using a Bonferroni correction. In addition, we analyzed the specific contrasts as mentioned in the introduction (discussion of 'adaptive scaling') in order to identify cortical activations sensitive to gradual increases in single-target reward value. The 'pure reward effect' was estimated from the contrast between 50%-target probability reward versus 50% no-reward conditions. The 'pure probability effect' was estimated from the contrast between no-reward 98% versus no-reward 50%-target probability conditions. Furthermore, in order to identify cortical activations sensitive to gradual increases in single-target reward value (which would NOT be predicted from the adaptive-scaling perspective), rather than just block-level reward versus no reward, we constructed a 3-level factor consisting of 50%-probability-reward (high reward per trial) versus 90%-probability-reward (low reward per trial) versus the average of the two no-reward conditions. This 'gradual-reward effect' was tested with MANOVA (GLM, SPSS version 22).

Results

Behavioral results

None of the contrasts, except the probability contrast for RT, was normally distributed. Reaction times to the targets were significantly shorter during the reward blocks (median (mdn) RT = 470 ms) compared to no reward blocks (mdn = 501 ms), Z = -3.20, p = .001, rank bi-serial r (rrb) = .55. Reaction times were also significantly shorter during the 98% target probability blocks (mdn = 469 ms) compared to the 50% target probability

blocks (mdn = 498 ms), $F(1,44) = 32.41$, $p < .001$, $\eta_p = .42$. Furthermore, participants were more accurate during the reward blocks (mdn = 99.4 %) compared to the no reward blocks (mdn = 98.8%), $Z = -3.05$, $p = .002$, rrb = .62. There was no significant main effect of target probability for the percentage correct responses. The percentage omissions was greater during the no reward blocks (mdn = 1.25%) compared to the reward blocks (mdn = 0.64%), $Z = -2.21$, $p = .027$, rrb = .43. There was no such difference between the high and low probability blocks. Participants rarely made commission errors to the NoGo stimulus. There was no significant difference between the reward and no reward block for the percentage commission errors (no reward block mdn = 0 %; reward block mdn = 0 %), $Z = -1$, $p = .317$, rrb = .5.

ERPs: PCA analysis of the effects of reward and target probability

Figure 2 shows superimposed ERPs from the four reward-probability conditions from selected midline electrode sites. The PCA yielded 8 temporo-spatial factors that met the pre-specified latency and medial-distribution criteria (see Methods paragraph 2.5.2). Table 1 provides an overview of the temporal and spatial distributions of the factors that met the pre-specified criteria and one additional factor that survived the Bonferroni correction. Figure 3 displays the temporal loadings for the components with a significant effect of reward or probability.

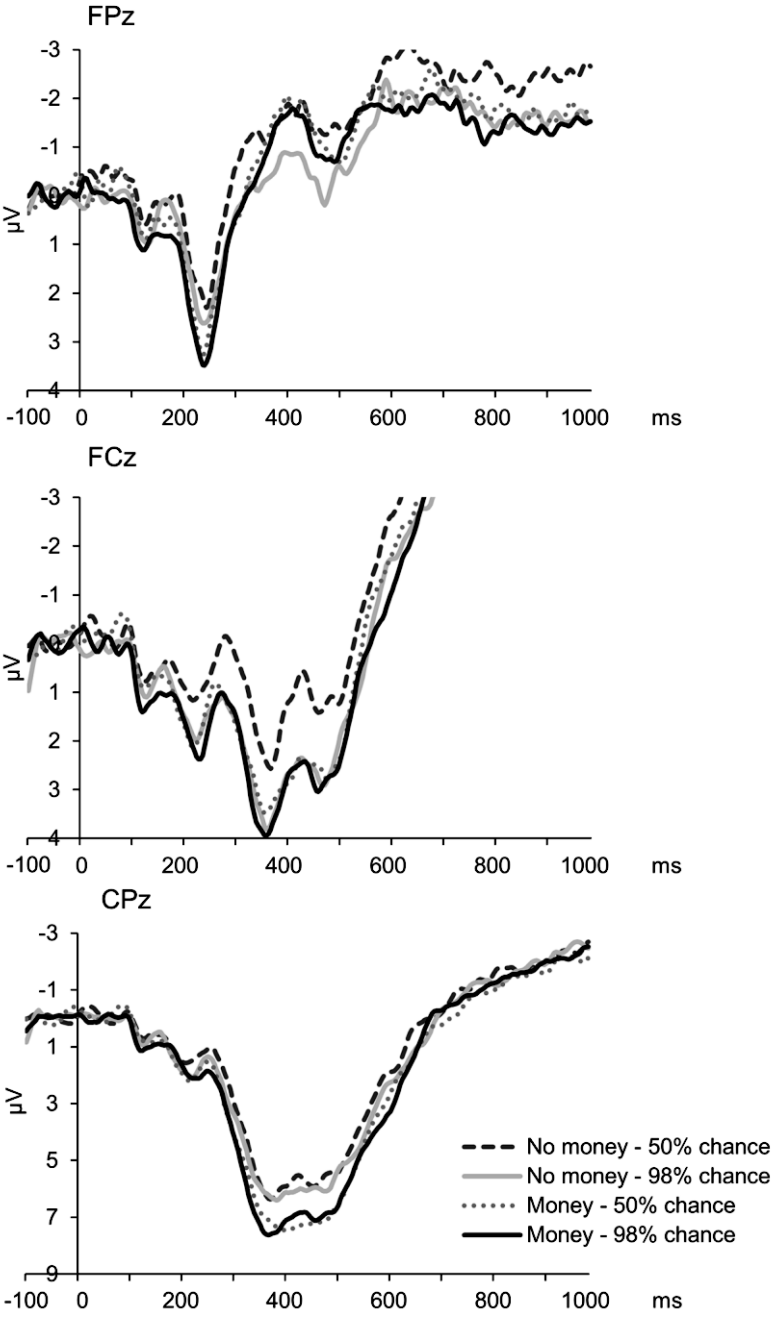


Figure 2. Grand average ERPs. The figure shows grand average ERPs for each of the four conditions, for the electrodes FPz (top), FCz (middle), and CPz (bottom). At timepoint 0 the cue was presented.

Table 1. Temporo-spatial factors

	Latency (ms)	Distribution max. positivity	Distribution max. negativity	Total variance explained (%)	Reward > noreward (p)	High probability > low probability (p)
TF1SF1	439	Pz[1]	Fpz	23.2	n.s.	.048
TF1SF2	439	Cz	P10	10.0	.009	n.s.
TF3SF1	287	POz	Fp2	2.3	n.s.	n.s.
TF3SF2	287	Cz	P9	0.8	n.s.	n.s.
TF4SF2	341	Cz	P9	1.1	n.s.	n.s.
TF5SF1	244	AFz	PO7	1.7	.001	n.s.
TF5SF2	244	CPz	P9	0.7	n.s.	n.s.
TF6SF1	400	Fp1	PO8	1.0	n.s.	n.s.
TF2SF1[2]	681	O1	Fp1	3.9	.001	n.s.

Note

[1] The TF1SF1 component was most positive at electrode Pz. The high > low probability effect, however, was observed at electrode Fpz. At Fpz the amplitude of TF1SF1 was less negative for the high compared to the low probability condition.

[2] TF2SF1 did not meet the pre-specified latency criterion and was therefore tested exploratively. The other factors that did not meet the pre-specified criteria did not survive Bonferroni correction.

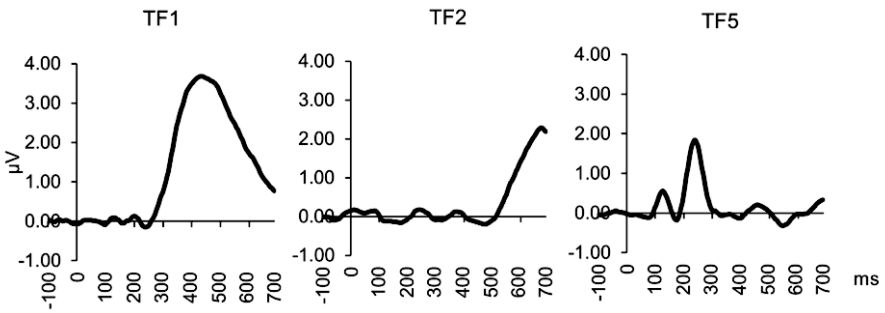


Figure 3. Temporal loadings as obtained before spatial decomposition. The figure displays the temporal loadings associated with the factors with a significant effect of reward or probability. TF1SF1 (probability-related positivity) and TF1SF2 (reward P300) originated from spatial decomposition of temporal factor (TF) 1 (left panel). TF2SF1 (late reward ERP) originated from spatial decomposition of temporal factor TF 2 (middle panel). TF5SF1 (reward-related positivity) originated from spatial decomposition of temporal factor TF 5 (right panel). Temporal loadings are converted to microvolt scaling.

Figure 4 displays the spatial loadings of these components. Figure 5 summarizes the effects of reward and probability on ERP activity. TF1SF1 peaked at 439 ms after the cue and its positive maximum was localized at Pz. This factor was significantly more positive for the low compared to high probability condition. The factor peak was slightly larger at electrode FPz where it was less negative for the high vs low probability condition, $F(1,44) = 4.12$, $p = .048$, $\eta_p = .09$. For the pure probability contrast (no-reward 50 vs 98%) this effect was replicated, $Z = -2.07$, $p = .038$, $rrb = .35$ (the distribution of the pure probability contrast deviated significantly from normality. The outcome of the Wilcoxon signed-rank test was therefore reported instead).

TF1SF2 had the same peak latency (439 ms), but its maximum positivity was localized at Cz. It was significantly more positive for the reward compared to the no reward blocks, $Z = -2.61$, $p = .009$, $rrb = .45$. This effect was replicated for the pure reward contrast (50% reward vs no reward), $t(44) = -2.74$, $p = .009$, $d = .41$. The gradual-reward effect (50%-probability-reward (high reward per trial) versus 90%-probability-reward (low reward per trial) versus the average of the two no-reward conditions) was also significant for this component, $F(2,43) = 3.41$, $p = .042$, $\eta_p = .14$ (note, however, that the 50% reward versus averaged no reward contrast

was not normally distributed. Therefore, the Wilcoxon signed-rank test was used as a follow-up test for this contrast). The gradual-reward effect reflected significant differences between both low and high per-trial reward versus no reward ($t(44) = 2.61$, $p = .012$, $d = .39$ and $Z = -2.29$, $p = .022$, rrb = .39, respectively), in the absence of a low versus high difference ($p = .357$). This is consistent with the adaptive-scaling perspective.

TF5SF1 peaked at 244 ms after the cue and its positivity was localized at AFz. The positivity was larger for the reward compared to the no reward condition, $F(1,44) = 11.81$, $p = .001$, $\eta_p = .21$. This effect was also significant for the pure reward contrast (50% reward vs no reward), $t(44) = -2.97$, $p = .005$, $d = .44$. The gradual-reward effect (50%-probability-reward (high reward per trial) versus 90%-probability-reward (low reward per trial) versus the average of the two no-reward conditions) was also significant for this component, $F(2,43) = 5.77$, $p = .006$, $\eta_p = .21$. This effect reflected significant differences between both low and high per-trial reward versus no reward ($t(44) = 2.90$, $p = .006$, $d = .43$ and $t(44) = 2.77$, $p = .008$, $d = .41$, respectively), in the absence of a low versus high difference ($p = .633$). This is consistent with the adaptive-scaling perspective.

Additional exploratory analyses were conducted for 10 temporo-spatial factors with lateral spatial distributions and/or temporal loadings outside the 180-500 ms interval. Only TF2SF1 survived the correction for multiple comparisons. This factor was significantly less negative for the reward compared to the no reward condition at electrode FP1, $F(1,44) = 12.84$, $p = .001$, $\eta_p = .23$. The special-contrast analysis again revealed a significant pure reward effect, $t(44) = -2.46$, $p = .018$, $d = .37$ and a gradual-reward effect, $F(2,43) = 6.30$, $p = .004$, $\eta_p = .23$, reflecting no difference between low and high ($p = .718$) while both low and high differed from no reward ($p = .012$ and .036, respectively).

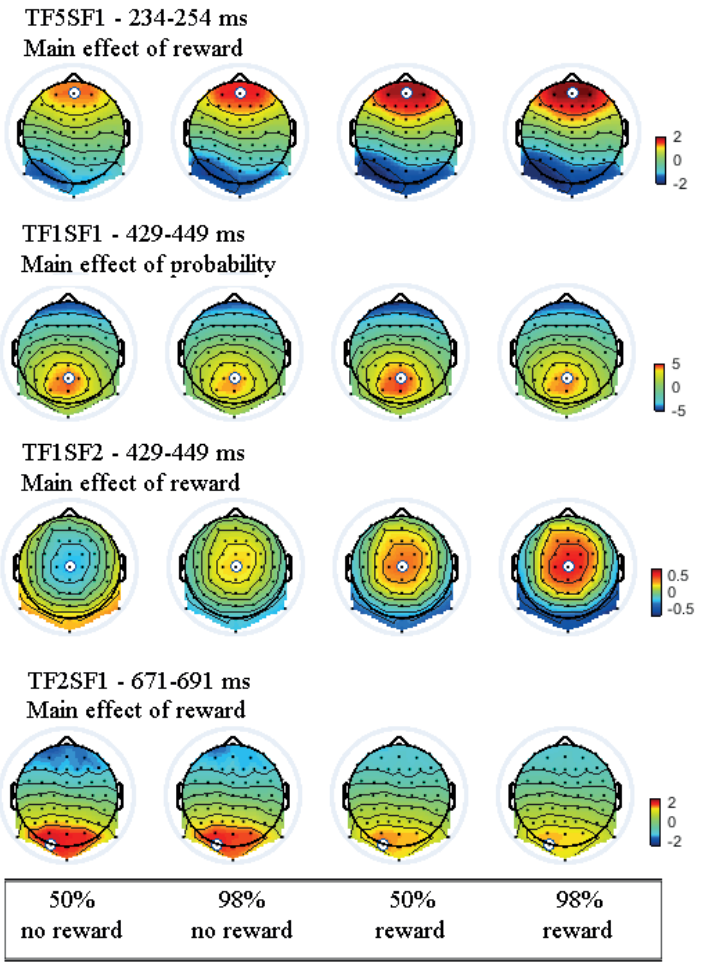


Figure 4. Spatial loadings principal component analysis. The figure displays the spatial loadings of the four temporo-spatial factors with a significant effect of reward (TF5SF1, TF1SF2, TF2SF1) or a significant effect of probability (TF1SF1, second row). Spatial loadings are presented for each condition. The bottom row indicates the conditions. The displayed values have been converted to microVolts.

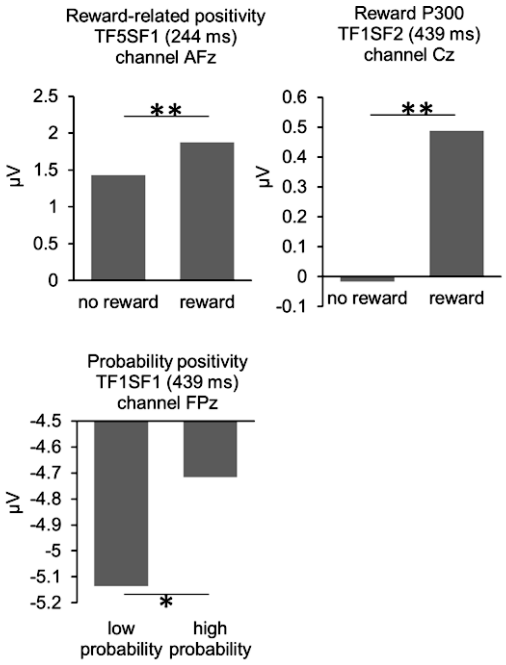


Figure 5. Main effects of reward and probability. The figure displays the factor scores (in microvolt) for the factors of apriori interest. Prospective reward affected a reward-related positivity (top left), and the P300 (top right). Probability affected a frontal ERP component ("probability-related positivity", bottom left). * $p < .05$, ** $p < .01$.

Discussion

The current study aimed to gain understanding in the temporal profile of the cortical processes activated during anticipation of reward with varying probabilities. Another main aim was to investigate whether ERPs elicited by reward and probability manipulations are dissociable. This study provides evidence for separate processing of reward value and probability in the cortex.

Consistent with prior studies (e.g., Bekker et al., 2004; Doñamayor et al. 2012; Flores et al., 2015; Pfabigan et al., 2014), reaction times were significantly shorter during the reward blocks compared to the no-reward blocks and during the 98% compared to 50% probability blocks. Participants were also more accurate during the reward blocks. These performance results show that cue-elicited processing must have been differential depending on reward and probability level. Note that this was the case even while, as in the current paradigm, behavioral choices did not at all concern reward or probability options.

To answer our research question, two main effects of reward were found, as well as one main effect of probability. One reward-related ERP emerged relatively early, and was strongest over the prefrontal electrode locations. A second reward-related ERP emerged later. This P300-like ERP peaked around 400 ms, and was prominent at the central electrode. The probability-related ERP had a similar latency, but the high-low probability ERP was largest over the medial prefrontal cortex. Exploratory analyses revealed an additional reward-related ERP late in the cue-target interval (around 680 ms post-cue). This reward-no reward ERP was largest over the left prefrontal cortex.

As noted, in the present design cues indicating high-probability rewarded targets also cue low single-target rewards. This implies a potential confound between reward and probability effects. Such a confound would not be expected from an 'adaptive scaling' perspective (Walsh & Anderson, 2012). This perspective predicts a response to a relatively low reward value (versus no reward), when the low value is the only available option (in addition to no reward, throughout a block of trials), that is identical to the response to a relatively high reward value (versus no reward) when the high value is the only available option. To evaluate the tenability of the adaptive-scaling perspective, special contrasts between conditions were constructed to assess differences between low and high per-trial reward effects (as in high-probability reward and low-probability reward blocked conditions, respectively). The analyses of the contrasts revealed that reward (versus no reward) effects did not differ at all as a function of per-trial reward magnitude, consistent with the adaptive-scaling perspective. In a similar vein, no indication was found that probability effects depended on the blocked-reward condition.

The PCA revealed an early frontal component that was significantly more positive when reward was at stake compared to when no reward was at stake. It had a latency of 244 ms, which is comparable to reported latencies of reward-related positivities (Doñamayor et al., 2012, Flores et al., 2015, Holroyd et al., 2011, Krigolson et al., 2014). These studies found an increased positivity around 200-250 ms following cues that signal reward compared to cues signaling no reward (Doñamayor et al., 2012; Flores et al., 2015; Holroyd et al., 2011, Krigolson et al., 2014). Similarly, Yu and Zhou (2006) observed less negative ERP activity for cues predicting

that money could be won during an upcoming gamble trial compared to cues predicting that money could be lost, albeit somewhat later in time (around 270 mst post-cue).

The early reward-related positivity observed in the present study and in the studies mentioned above, may be an instance of "the reward positivity". This is an ERP mostly observed *after* positive feedback about performance or a rewarding outcome (Foti et al., 2011; Holroyd et al., 2008; Holroyd et al., 2011; Proudfit, 2015). It usually peaks at mid-frontal electrode sites and is proposed to reflect phasic dopaminergic input from the ventral tegmental area into the dorsal ACC when outcomes are better than expected in order to guide reinforcement learning (Holroyd & Coles, 2002) or to reduce conflict (Holroyd et al., 2008). A similar reward-related positivity has also been observed after reward announcing cues after the reward-predicting value of the cue has been learned (Krigolson et al., 2014). As such, the reward-related positivity associated with cues may reflect an initial estimation of the likelihood of a prospective reward which is adjusted after feedback when necessary (Holroyd et al., 2011).

In the current study the early reward-related positivity (we refer here to any reward manipulation-specific positive deflection peaking before 300 ms) was most prominent at the medial prefrontal electrode site AFz, and resembled the topography of the reward-related positivity as observed by Flores et al. (2015). Nieuwenhuis and colleagues (2005b) observed multiple generators of this component[4] including areas within the rostral ACC. In the study by Doñamayor and colleagues (2012), however, the reward-related positivity peaked somewhat more posteriorly (i.e., at the mid-frontal electrodes), and was source localized to the dorsal posterior cingulate cortex (dPCC).

The early reward-related positivity in the current study may alternatively reflect enhanced attentional capture by the reward cues (Padmala & Pessoa, 2011), or may instead reflect the modulatory effect of reward on sensory processes (Pessoa & Engelmann, 2010). A related phenomenon has been described as the 'frontal selection positivity' (FSP; Kenemans et al., 2002; Bekker et al., 2004). This is a frontally distributed ERP deflection that is stronger for relevant (e.g., cues signaling potential

[4] Note that this reward-related positivity was actually presented as a feedback-related negativity.

reward) relative to less relevant stimuli (e.g., cues signaling no potential reward). Identifying the present reward-related positivity as an FSP would imply a pronounced contribution of posterior-cortex generators (Kenemans et al., 2002), consistent with an interpretation in terms of the modulatory effect of reward on sensory (i.e., visual) processing.

The PCA yielded another component that was significantly more positive for reward compared to no-reward cues. This component has a central distribution and a temporal loading of 439 ms. The latency and central distribution of this ERP may be consistent with an interpretation in terms of P300/P3b. This finding is in line with previous research showing sensitivity of the P300 to the magnitude of wins and losses (Yeung & Sanfey, 2004) as well as the anticipation of reward (Broyd et al., 2012; Flores et al., 2015; Pfabigan et al., 2014). Nieuwenhuis and colleagues (Nieuwenhuis, Aston-Jones & Cohen, 2005a) argued that the P300 reflects the modulatory influence of the locus coeruleus norepinephrine system on information processing in the case of motivationally relevant events. Pfabigan et al. (2014) additionally suggested that the P300 elicited during anticipation of reward may particularly be dependent on dopamine transmission.

The third main effect concerned an effect of target probability. The PCA component sensitive to the probability manipulation had a temporal loading of 439 ms. It consisted of a parietal positivity and prefrontal negativity. The component was significantly more positive (less negative) for highly reliable cues (indicating 98% probability of an upcoming target) compared to unreliable cues (indicating 50% probability of an upcoming target) at the prefrontal electrode site. Based on a prior study of our lab (Bekker et al., 2004) and classical findings on the P300 (Duncan-Johnson & Donchin, 1982) we expected the parietal P300 to be sensitive to target probability. However, the high > low probability ERP effect in the current study was frontally distributed. This frontal distribution is consistent with an fMRI study by Knutson et al. (2005) indicating that probability is also represented in the medial prefrontal cortex (Knutson et al., 2005). It is probable that the precise nature of probability effects in the context of explicit reward is different from that in more implicit-reward contexts (e.g., when subjects just follow the instructions of the experimenter).

It could be argued that any effect of probability on cue-elicited activation reflects enhanced response preparation to highly probable

subsequent targets, compared to low-probability targets. However, this view predicts probability effects on late slow cue-induced potentials such as the contingent negative variation and stimulus-preceding negativity, not on earlier activations such as the currently described probability ERP. In addition, Bekker et al. (2004) reported such early probability effects, in the absence of probability effects on late slow potentials. Furthermore, in the present study the extent of specific-response preparation was probably very limited, as target stimuli embodied a two-choice reaction-time task in which both response alternatives had equal probability.

In conclusion, the current study aimed to dissociate the effect of reward value and target probability manipulations on anticipatory ERPs and provides evidence for separate processing of reward value and probability cues in the cortex. An early reward-related positivity and a late (P300-like) ERP component were specifically affected by reward availability, whereas target probability affected a late frontally distributed ERP. Both reward effects obey the principle of adaptive scaling. The early-reward-related positivity may reflect reward-modulated sensory processing of the reward cues. The probability effect is qualitatively different from analogous effects reported before, perhaps due to the explicit-reward context as maintained in the present paradigm.

Supplementary materials

S1.1 Methods - standard ERP analysis

We limited our standard ERP analysis to the midline sensors, as the PCA analysis yielded reward and probability ERPs for the midline sensors only. ERP data were re-referenced to the averaged mastoids. Individual ERP activity time-locked to nocues was subtracted from cue-locked ERP activity. All other preprocessing steps were as described in paragraph 2.5.2 of the manuscript. Data of six subjects were removed from further analyses, because less than half of the of the cue- (< 40) or nocue-locked (< 100) segments were left in one or more conditions for one or more midline electrodes (FPz, AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz). Grand-averages across subjects were made for each condition. The number of subjects included in the standard analysis amounted to 42.

Cue minus nocue ERPs were divided into four 40-ms time bins around each components peak latency. The ERP peak latency was based on the temporal loadings as obtained with the PCA. Forty-ms time bins were analyzed within the following time windows: 158-321 ms (reward-related positivity), 363-526 ms (reward P300 and probability ERP). Midline scalp distribution plots of the effects of reward and probability were produced in order to visualize where the effects of reward and probability were maximal (Figure S1-S3 below). Reward x probability ANOVAs were run for each of the 40-ms bins for the electrodes with a maximum effect of reward or probability.

S1.2 Results – standard ERP analysis

Figure 2 in the main article displays the grand average ERPs for each condition for the midline sensors with the strongest effects of reward and probability (i.e., FPz and CPz, and FCz, respectively; see Figures S1-S3).

Reward (no reward - reward) x target probability (50% - 98%) repeated-measures ANOVAs were conducted for each of the four 40-ms time bins per ERP component. All contrasts were tested for normality using Shapiro-Wilks tests. In case of significant deviation from normality, the outcome of the non-parametric Wilcoxon signed-rank test was reported. The reward-related positivity was strongest at electrode FPz (Figure S1). ERP activity at FPz was significantly more positive in the reward condition

than in the no reward condition during the interval 158-280 ms, all p's between .003 - .015, $\eta^2 > .14$ (all obtained with F-tests). The reward P300-like potential was strongest at electrode CPz. ERP activity at CPz was significantly more positive in the reward condition than in the no reward condition in the interval 363-526 ms, all p's between .002 - .014, $\eta^2 > .14$ (the distribution of the reward contrast in the 404-444 ms window deviated from normality and was therefore tested with the non-parametric test). The probability ERP was strongest at electrode FCz. ERP activity at FCz was significantly increased for the high compared to the low target probability condition during the time bin 363-403 and 445-485 ms, p's < .049, $\eta^2 > .09$ (F-tests). This effect was marginally significant for the time bin 404-444 ms (p = .065)

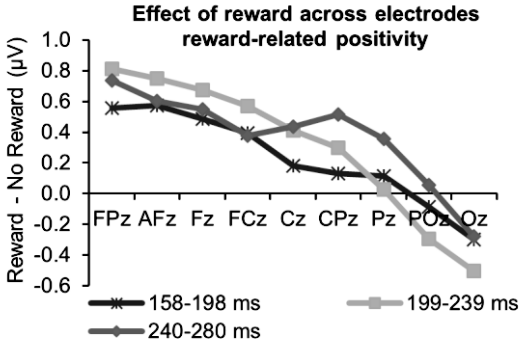


Figure S1. The reward-related positivity across the midline electrodes. The three lines represent the average reward minus no reward ERP activity for each of the midline channels and for the 40-ms time bins with a significant effect.

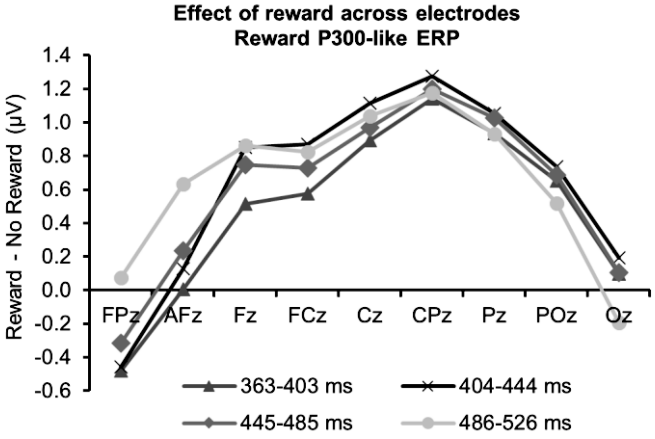


Figure S2. The reward P300-like ERP across the midline electrodes. The lines represent the average reward minus no reward ERP activity for each of the midline channels and for the 40-ms time bins with a significant effect.

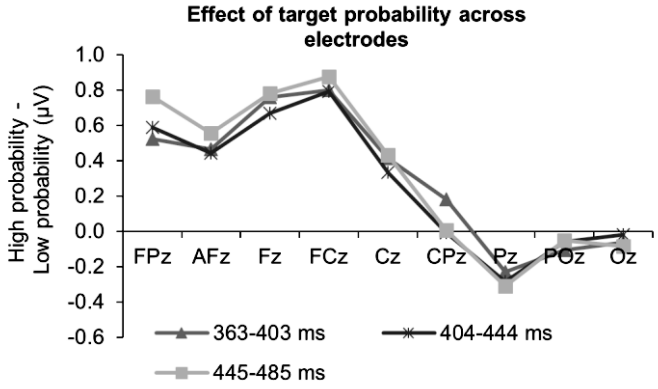


Figure S3. The probability ERP across the midline electrodes. The three lines represent the average high minus low target probability ERP activity for each of the midline channels. ERP activity was significantly increased for the high compared to the low probability condition at FCz for the time bins 363-403 and 445-485 ms, and marginally significant for the time bin 404-444 ms.

S2 Details on the amount of ERP data that were interpolated for the principal component analysis

Table 1
Details on the amount of data that were interpolated. The table contains case numbers for which data were interpolated for specific channels (rows) and task conditions (columns).

	Condition			
	50 % no reward	98 % no reward	50 % reward	98 % reward
Interpolated channel				
Fpz	case 40			
AF7	case 40		case 32	case 32, 49
AF8	case 40, 43		case 8, 43	
F5			case 5	
F7	case 21			case 21
F8			case 37	
FT8			case 9	
P7	case 18			
P10	case 21			
PO3	case 31			
PO7				case 21
O1	case 21	case 16		
O2			case 9	
Iz	case 21		case 21	case 21

Chapter 6

Dopaminergic and noradrenergic manipulation of anticipatory reward and probability event-related potentials

Iris Schutte
Peter K.H. Deschamps
Peter N. van Harten
J. Leon Kenemans

Submitted in a slightly revised form

Author contributions

LK and IS conceptualized and set up the study; IS programmed the task and collected the data with help of a master student; PvH and PD provided (clinical) advise, PD was involved in the clinical supervision; IS and LK contributed to analyzing the data; IS wrote the initial draft of the manuscript; all authors contributed to the final manuscript; LK supervised the study.

Abstract

Predicting what happens in the future in terms of potential reward is essential in daily life. The aim of the current study was to investigate the neurotransmitter systems involved in the anticipation of reward value and probability. We hypothesized that dopaminergic and noradrenergic antagonism would affect anticipation of reward value and probability, respectively. Twenty-three healthy participants were included in a haloperidol (2 mg) x clonidine (0.150 mg) x placebo cross-over design and subjected to a Go/NoGo experimental task during which cues signaled the probability of subsequent target appearance. Reward value (amount of money that could be won for correct and fast responding to the target), as well as probability of target appearance were orthogonally manipulated across four task blocks. Cue-elicited EEG event-related potentials were recorded to assess anticipation of value and probability, respectively. The processing of reward value was affected by dopaminergic antagonism (haloperidol), as evidenced by reduction of the reward-related positivity and P300 to reward cues. The reward P300 was only reduced for subjects with high baseline dopamine levels. In contrast, the processing of reward probability was affected by noradrenergic antagonism (clonidine). In addition, both drugs reduced overall performance (omission rate, response-speed variability). We conclude that at least anticipation of reward value and probability, respectively, is specifically affected by dopaminergic versus noradrenergic antagonism.

Introduction

Predicting future reward value and the likelihood of prospective outcomes enables one to select appropriate actions that maximize gain (Glimcher & Rustichini, 2004). Clarification of the brain mechanisms of decision making as driven by cues signaling reward value and probability is relevant to our understanding of various psychiatric disorders associated with impairments of reward-related decision making (e.g. depression, ADHD, schizophrenia, addiction).

Prior studies have shown that manipulations of anticipated reward value affect a frontal event-related potential (ERP) early in time ("reward-related positivity" (RRP); (Doñamayor, Schoenfeld, & Münte, 2012; Flores, Münte, & Doñamayor, 2015; Holroyd, Krigolson, & Lee, 2011; Krigolson, Hassall, & Handy, 2014; Chapter 5) and a more parietal ERP later in time ("reward P300"; Broyd et al., 2012; Flores et al., 2015; Pfabigan et al., 2014; Chapter 5). Anticipation of the probability of (rewarded) targets elicited a separable frontal ERP component ("probability-related positivity" (PRP); Chapter 5). Here, we aimed to investigate whether these reward and probability ERPs can be dissociated in terms of underlying neurotransmitter systems. Specifically, we tested the hypothesis that, respectively, dopamine (DA) and noradrenaline (NA) are involved in generating these ERPs.

Informative reward cues cause abrupt (phasic) changes in DA level, which have been hypothesized to signal deviations in expected reward value, and are associated with immediate enhancements in motivation and/or task engagement (Collins & Frank, 2016; Hamid et al., 2016). These DA signals are thought to depend on subcortical transient DA increases that have been related to cortical action, i.e., the effect of DA fluctuations on cortical mechanisms underlying subsequent action to an imperative stimulus (Frank & O'reilly, 2006). The present study addresses the sensitivity of these cortical mechanisms, including RRP and P300, to DA and NA manipulations.

Dopaminergic drugs and their effects on (reward) prediction ERPs have so far only been studied in experiments that focused on post-reward mechanisms. These concern the error-related negativity (ERN) and the feedback-related negativity (FRN). Both can be conceived as quick and

transient negative reward-prediction errors, presumably driven by phasic dips in DA release after omissions of (highly) expected rewards. The role of phasic dips in DA release after unexpected reward omission is supported by studies showing that DA agonist treatment leads to increased amplitudes of the ERN (De Bruijn, Hulstijn, Verkes, Ruigt, & Sabbe, 2005; Santesso et al., 2009), whereas DA antagonists had the opposite effect (2.5 and 3 mg haloperidol; De Bruijn, Sabbe, Hulstijn, Ruigt, & Verkes, 2006; Forster et al., 2017; Zirnheld et al., 2004).

Prior studies (Cools et al., 2009; Frank & O'reilly, 2006) have demonstrated that (the effect of DA drugs on) reward-based learning, which is thought to be dependent on these reward prediction ERPs, is modulated by baseline DA functioning. Cools and colleagues (Cools et al., 2009) found that the DA agonist bromocriptine enhanced reward-based reversal learning in subjects with low baseline DA, while it had the opposite effect in subjects with high baseline DA. However, direct evidence for a relationship between baseline DA, and the effect of DA drugs on reward-related ERPs and concomitant reward-based decision making is lacking. The current study aims to shed light on this relationship, which will have important implications for our understanding of the strong individual differences in DA drug efficacy seen in the clinic.

The NA system may have a complementary role in reward processing. NA firing is associated with the amount of effort needed to obtain a reward only *after* the initial choice has been made to work for the reward (the latter was modulated by DA) (Floresco, 2015; Varazzani, San-Galli, Gilardeau, & Bouret, 2015). Therefore, it may be hypothesized that after initial DA-based estimation of future reward value, perhaps aiding the decision to work hard or not, the NA system may be activated in order to mobilize cognitive and physical resources accordingly (Floresco, 2015).

In our prior study (Chapter 5), cues signaling a high-probability subsequent target elicited a frontal-central positivity ERP. It should be noted that this probability-related positivity did not resemble the more posteriorly distributed P300, which has been reported as sensitive to the probability of events and hypothesized to index activity of the NA system (Bekker, Kenemans, & Verbaten, 2004; Duncan-Johnson & Donchin, 1982; Joseph & Sitaram, 1989; Nieuwenhuis, Aston-Jones, & Cohen, 2005). However, there are strong indications that the P300 in the context of reward

processing may be particularly dependent on DA (Pfabigan et al., 2014). In the current study we investigated whether the PRP and reward P300 are dissociable in terms of underlying neurotransmitter systems.

In sum, we aimed to uncover the roles of the DA and NA system in modulating cortical mechanisms activated in the context of reward anticipation. This was investigated by blocking the DA and NA system, respectively, by 2 mg haloperidol and 0.150 mg clonidine. We furthermore investigated whether these modulations were accompanied by a concomitant performance decrease. Another main aim was to investigate whether baseline DA modulates the effect of DA drugs on reward-related ERPs. We expected the reward P300 and RRP to be reduced under haloperidol and the PRP to be reduced under clonidine. A placebo-controlled crossover study was performed using the cued Go/NoGo task from our prior study (Chapter 5) in which cues predicted upcoming targets with varying reward value and probabilities. Spontaneous eye blink rate (EBR) was used as an indirect measure of baseline striatal dopaminergic activity (Jongkees & Colzato, 2016). We expected the attenuating effects of haloperidol on reward-related ERPs to be more pronounced in subjects with high EBR (Cools et al., 2009). Finally, to further account for individual differences in drug responsiveness we compared the effects of clonidine and haloperidol on reward and probability related ERPs between drug responders and non-responders (See Methods).

Methods

Subjects

Twenty-nine healthy males were enrolled in the study (sample size justification can be found in Supplementary materials section 1). Subjects were recruited via advertisement at Utrecht University and via a recruitment website. Exclusion criteria were: (1) a history of relevant medical conditions or mental health issues, (2) current medication use, (3) smoking, (4) a history of cocaine use, (5) daily consumption of > 3 glasses of alcohol (6) more than one occasion of recreational drug use per month, and (7) low blood pressure (< 90 mmHg systolic and/or < 60 mmHg diastolic), low or high heart rate (< 55 or > 100 bpm) during first study visit.

Participants were requested to abstain from consuming xanthines and alcohol for at least 12 hours prior to each session, and to refrain from psychotropic drugs for at least two weeks prior to each session. All participants declared to have normal or corrected-to-normal vision. The study was approved by the medical ethical committee of the University Medical Centre Utrecht and conducted in accordance with the declaration of Helsinki. The study was pre-registered in the Netherlands Trial Registry (NTR), number: NTR5019. Participants received 10 Euros per hour, and could win a maximum of 15 Euros per session during the cued Go/NoGo task.

Two subjects were excluded because of voluntary withdrawal and one because of nausea during the placebo session. Data of three subjects were discarded during the analyses (see data reduction and analysis). The final sample consisted of twenty-three males (mean age (± sd): 22.8 (± 3.7) years).

Pharmacological manipulation

Haloperidol 2 mg is a potent antagonist of the dopamine D2 receptor, and is assumed to attenuate neurotransmission in the mesocortical and mesolimbic DA pathway (Kapur, Zipursky, Jones, Remington, & Houle, 2000). Clonidine 0.150 mg binds to pre-synaptic α2-adrenergic receptors, which has an inhibiting effect on NA release from broadly distributed NA nerve terminals (Svensson, Bunney, & Aghajanian, 1975). All capsules were over-encapsulated by the pharmacy to ensure double-blinding.

Cued Go/NoGo task

The same Cued Go/NoGo task was used as in Chapter 5. The task is illustrated in Figure 1A. Details of the task can be found in the supplementary materials, section 2. Subjects had to press a left or right button when a target letter (letter X or Y) followed a cue letter (always letter A), as fast and accurately as possible.

The amount of money that could be won for correct and fast responses (a total of either 0 Euros or 5 Euros during the block) and the probability of target appearance after the cue (either 50 % or 98 %) were orthogonally manipulated across four task blocks. This reward value and probability information was shown to participants at the beginning of each

block. During reward blocks of the pre-drug task version, which lasted half as long, 2.50 Euros could be won.

Participants were assigned to one of four possible task block orders. For each participant this order was kept the same for the pre- and post-version of the cued Go/NoGo task and for all three sessions. Task block order was counterbalanced across participants and drug order.

EEG – EOG data acquisition

The electro-encephalogram (EEG) and electro-oculogram (EOG) were recorded with the Active-Two system (Biosemi, Amsterdam, The Netherlands) using 64 Ag-AgCl electrodes placed according to the international 10/10 system. The horizontal and vertical EOG were recorded from electrodes placed above and below the left eye and electrodes at the outer canthi of both eyes. Signals were online referenced to the Common Mode Sense/Driven Right Leg electrodes and online filtered with a filter at DC to 400 Hz (default); the sample rate was 2048 Hz.

Cardiovascular measures

Systolic blood pressure change after clonidine administration was used as a proxy for central alpha-2 stimulation (Logemann, Böcker, Deschamps, Kemner, & Kenemans, 2014), under the assumption that the effects of clonidine on cortical systems and the brainstem cardiovascular center are correlated. This is reasonable as the anti-hypertensive properties are probably at least partly mediated by stimulating alpha-2 receptors. Blood pressure and heart rate were assessed by an automatic blood pressure monitor (Microlife BP A6 PC) in a double-blind fashion.

Motoric measures

The potential occurrence of hyperkinesia or akathisia (Kapur et al., 2000) after haloperidol administration was monitored with an accelerometer (Actigraph, GT3X+, Actigraph, LLC, Pensacola, Fl, USA) placed around the right ankle of the participant. Motor activity along three axes (x, y, z) was stored with a sampling rate of 100 Hz. The increase in motor activity following haloperidol treatment was used as a proxy for central responsivity to haloperidol (Logemann et al., 2017).

Subjective measures

The methods and analyses of the subjective effects data can be found in the Supplementary Materials, section 3.

Procedure

During an initial screening session the informed consent form was signed and in/exclusion criteria were checked. Figure 1B presents an overview of the procedure of each of the three main sessions.

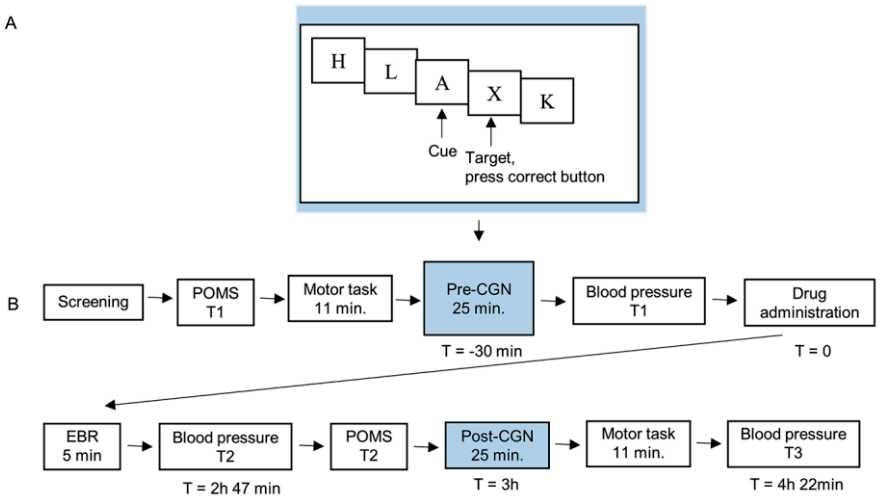


Figure 1. Procedure and Cued Go/NoGo task. A. Participants were subjected to the cued Go/NoGo task twice per session: to a short version before drug administration and to a longer version with EEG recording 3 hours after drug administration. Subjects wore an actigraph around the right ankle during both cued Go/NoGo task versions. Subjects sat in a comfortable dentist chair one meter in front of a computer screen in a separate dimly lit room. The subject's eyes were at the level of the center of the screen. Subjects were instructed to press a pre-specified button when letter X or Y followed the cue (always letter A). The amount of money for correct and fast responding and the probability that a cue would be followed by a target were orthogonally manipulated across four task blocks. The amount of money was either 0 or 5 Euros maximally per reward block (or 0 and 2.5 Euros during the pre-drug task version), and the probability of target appearance was either 50% or 98%. B. Overview of the procedure of one test session of the main study. The main study consisted of three sessions separated by at least one week. During each session subjects received 0.150 mg clonidine, 2 mg haloperidol, or placebo. Session order was balanced across subjects.

A cross-over design was used and during each of the three sessions participants either received 2 mg haloperidol, 0.150 mg clonidine or placebo. Sessions were separated by at least one week. Participants and researchers who carried out the experiments and who performed the analyses were blinded to the drug allocation. Subjects were randomly assigned to a drug order, which was counterbalanced across subjects. Randomization of drug order was performed using Excel by a researcher who was not involved in the study.

Data reduction and analysis

Cued Go/NoGo task – behavioral data

Mean reaction times (RTs) for valid responses to the target (i.e., single responses within the time window 150-1500 ms after target onset), reaction time variability (SDRT), the percentage correct responses, and the percentage omissions were calculated for each condition and each subject. Pre-drug behavioral data of one subject were not stored due to a technical issue. Behavioral data of this subject were therefore not analyzed.

Spontaneous movements

Ten minutes of data were analyzed for the pre- and post-drug condition, starting from ten minutes after Cued Go/NoGo task initiation. For each axis (i.e., the x, y, z direction) data were integrated into 10 s-epochs and subsequently averaged across the three axes.

EBR and ERP data

Eye blink and ERP data were analyzed using Brainvision Analyzer 2.0 (Brain Products GmbH). Data were re-sampled to 256 Hz. Bipolar eye blink signals were obtained by subtracting the lower and upper channels, respectively. EEG data were re-referenced to the averaged mastoids. A 0.5-30 Hz band pass filter (24 dB/oct), and an additional 50 Hz Notch filter were applied. The number of eye blinks were subsequently counted using the Gratton and Coles algorithm (Gratton, Coles, & Donchin, 1983) and transformed to EBR per minute.

ERP data were epoched into windows from -100-1000 ms surrounding cue and nocue onset (nocues were letter stimuli not followed by a target and not preceded by a cue). Epochs with incorrect, pre-mature

(< 150 ms) or late responses to the target (> 1500 ms) or with omissions or commission errors were excluded. For each subject extreme artifacts in the EOG channels and the target midline electrodes[5] were removed automatically without removing normal eye blink activity. This was done in order to improve the subsequent ocular artifact correction step, which used the Gratton & Coles algorithm (Gratton et al., 1983). Two subjects exhibited a very low number of eye blinks, which led to an incorrect estimation of vertical EOG transfer coefficients. Therefore, for these two subjects, the Gratton & Coles method was used to correct only horizontal EOG artifacts, and to remove segments with vertical EOG artifacts. This was done for all task and drug conditions. Furthermore, for six subjects EOG correction was inadequate for the posterior electrodes (Pz, POz, and/or Oz), as evidenced by clearly deviant EOG transfer coefficients for these electrodes. For these subjects we used the EEG signal of the Pz, POz, and Oz leads without EOG correction for all task and drug conditions (the signals of the other electrodes were corrected for ocular artifacts).

Data were baseline corrected using the 100 ms period before cue onset. Segments with activity lower than 0.5 μV over a 100 ms period or with an absolute difference between values exceeding 100 μV were automatically removed. Average ERPs were computed for each condition, and ERP activity time-locked to no cues was subtracted from cue-locked ERP activity for each condition. This was done to isolate blocked-condition-dependent brain activity specifically associated with the cue. Finally, grand averages of the cue-minus-no-cue ERPs were computed for each condition.

Data of two subjects were discarded, because of more than 50% choice errors and/or omissions for one or more conditions. Data of one subject were discarded because of inadequate ocular artifact correction, which was likely due to atypical eye movements. The final number of subjects included in the ERP analyses amounted to 23.

[5] As will be explained in the paragraph "Cued Go/NoGo task - ERP data –selection of time windows and electrodes" we used two methods to select target time windows and electrodes for statistical testing. For the selection method based on the results of our prior study, segments were removed in case of extreme artifacts in only the EOG channels, or the channels of a-priori interest: FPz, FCz, or CPz. For the collapsed localizer approach, segments were removed in case of extreme artifacts in the EOG channels or any of the midline electrodes (from FPz to Oz).

Statistical analyses

Cued Go/NoGo task – behavioral data

Repeated-measures MANOVAs (GLM, SPSS version 22) were run for RT, RT variability, the percentage correct responses, and the percentage omissions, with reward value (no reward, reward), target probability (50%, 98%), drug (clonidine, haloperidol, placebo), and time (pre-, post-drug), as within-subject variables.

Cued Go/NoGo task - ERP data –selection of time windows and electrodes

We originally started by selecting electrodes and time windows for statistical testing that had the strongest effects in our prior study (a-priori selection method; Chapter 5). This yielded: 199-280 ms at FPz (reward-related positivity), 363-526 ms at CPz (reward P300), and 445-485 ms at FCz (probability-related positivity). These intervals are also largely consistent with data of and intervals used by prior studies (Flores et al., 2015; Holroyd et al., 2011). However, based on a recommendation during a previous review process we alternatively implemented the "collapsed localizer" approach (CLA; Luck & Gaspelin, 2017). The latter method involves testing where/when task effects (i.e., the main effect of reward and probability) are strongest across all drug conditions, and using these electrodes/time windows to subsequently test the effects of the drugs. On the one hand, the a-priori selection method may introduce a bias, because the spatial and temporal distributions of the task effects may be slightly different for different experiments. On the other hand, the CLA is very conservative, because task effects are tested across drugs. Task effects may be missed if drugs have opposite effects on the sign of the reward and probability ERP amplitudes. For example, one would not find a main effect of probability across drugs if ERP amplitudes are larger for the high compared to low probability condition under placebo and haloperidol, and if this effect is reversed under clonidine (i.e., low probability > high probability). To foreshadow our results, this is exactly what happened. We could not identify electrode locations and time windows with a main effect of probability using the CLA. Therefore, for the PRP we chose the target electrode and time window with the strongest effect in our prior study (i.e., 445-485 ms at FCz). With respect to the RRP and reward P300, the CLA

yielded: 240-280 ms at FPz for the RRP, and 281-526 ms and the average of the electrodes CPz, Pz, POz, Oz for the reward P300.

Effects of the drugs on the three ERPs were tested using drug x reward value x probability MANOVAs. We noticed, however, that the results obtained with the two methods were slightly different, even for the PRP for which the same time window and electrode was used for testing. The latter was due to a difference in the artifact rejection procedure between both methods leading to a different selection of segments (See also Footnote 5).

There is no specific reason to believe that for this study one of these methods is superior to the other. We, therefore, chose to average the ERPs obtained with the two methods, under the assumption that the ERPs obtained with the a-priori method reflect the same underlying processes as the ERPs obtained with the CLA. Averaging between these two methods suppresses the noise associated with each of these methods, while commonality between the methods is kept.

The ERP results section presents the results of ERP data averaged across selection methods. Effects of the drugs on the three ERPs were tested using drug x reward value x probability MANOVAs. Additional median-split analyses were run to take individual variance in the effect of the administered dose into account. A proxy for individual variance in the effect of haloperidol on the reward P300 and RRP was based on spontaneous movement increase after haloperidol compared to placebo. A proxy for individual variance in the effect of clonidine on the PRP was based on the decrease in systolic blood pressure after clonidine compared to placebo. Another median-split analysis addressed individual variance in baseline DA, for which EBR during the placebo condition was used as a proxy. For all median-split analyses we only compared the drug of interest with placebo.

Alpha was set at .05. The between-subjects factor drug order was initially included in all MANOVAs in order to reduce the variance induced by this factor for the tests of interest. If a given effect of interest did not depend on order, the order factor was removed from the model, so as to increase the dfs for the effect of interest (Kenemans, Wieleman, Zeegers, & Verbaten, 1999).

Drug effects on peripheral measures

Additional analyses were run to test the group-level effects of the drugs on EBR, cardiovascular data and spontaneous movements. These results can be found in section 4 of the supplementary materials.

Results

Cued Go/NoGo task – behavioral results

For none of the four performance measures (RT, RT variability, % correct, % omissions) there were significant interactions among the effects of drug, reward value, and probability respectively (all $p > .079$). With respect to main effects of reward value and probability (details can be found in the supplementary materials, section 5), subjects responded faster and were more accurate in the reward compared to the no reward condition. RTs were also less variable and the percentage omissions was lower for the reward compared to the no reward condition.

As to drug effects (Figure 2), the percentage omissions across all task blocks was significantly increased after clonidine, $F(1,21) = 10.3$, $p = .004$, $\eta_p^2 = .33$, and haloperidol administration, $F(1,21) = 7.1$, $p = .014$, $\eta_p^2 = .25$, compared to the pre-treatment time point (time x drug: $F(2,20) = 5.2$, $p = .016$, $\eta_p^2 = .34$). The increase from pre to post measurement was significantly stronger for haloperidol compared to placebo ($p = .032$), but not significantly stronger for clonidine compared to placebo ($p = .053$). A significant time x drug interaction was found for RT variability, $F(2,20) = 3.9$, $p = .037$, $\eta_p^2 = .28$, indicating that RT variability was increased following clonidine, $F(1,21) = 8.0$, $p = .010$, $\eta_p^2 = .28$ and haloperidol treatment, $F(1,21) = 14.8$, $p = .001$, $\eta_p^2 = .41$. No such increase was observed following placebo ($p = .99$). The increase in RT variability from pre to post treatment measurement was significantly stronger for both the clonidine and haloperidol condition compared to placebo (p values ≤ .042).

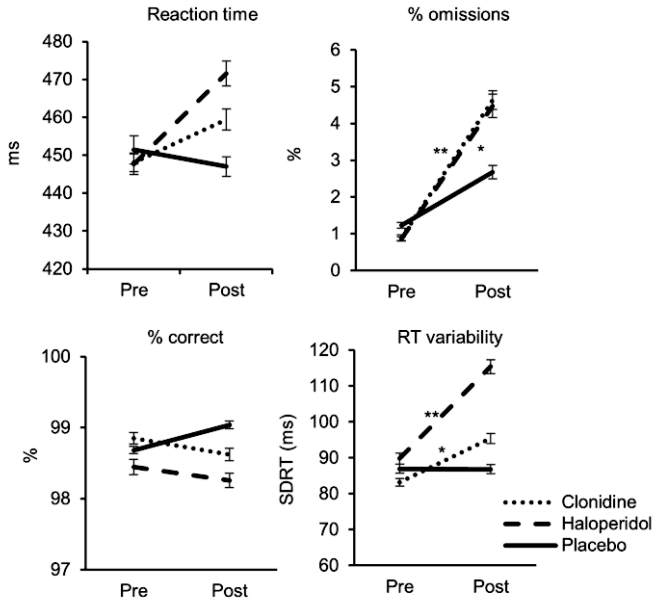


Figure 2. Effects of the medication on behavioral parameters. The figure presents the effects of drug administration on reaction times, the percentage omissions, the percentage correct responses, and the reaction time variability during the cued Go/NoGo task. Pre on the X-axis represents the pre medication measurement; Post represents the post-medication measurement. Error bars represent 1 standard error. Stars represent significant differences between the pre and post-drug measurement; * $p < .05$, ** $p < .01$.

Cued Go/NoGo task – ERPs

Drug effects on the RRP

The RRP was significantly present across drug conditions (main effect of reward), $F(1,17) = 5.6$, $p = .031$, $\eta_p^2 = .25$. There was also a significant drug x reward interaction, $F(2,16) = 6$, $p = .012$, $\eta_p^2 = .43$. Follow-up tests showed that the RRP was significantly and specifically reduced by haloperidol, as evidenced by significant main effects of reward for both clonidine, $F(1,17) = 6.2$, $p = .023$, $\eta_p^2 = .27$ and placebo, $F(1,17) = 6.6$, $p = .020$, $\eta_p^2 = .28$, and absence of the reward effect for haloperidol, $p = .335$ (Figure 3).

Median-split analyses including the factor EBR and spontaneous movement increases did not yield significant interactions between drug, reward, and group, $p = .537$ and $p = .446$, respectively.

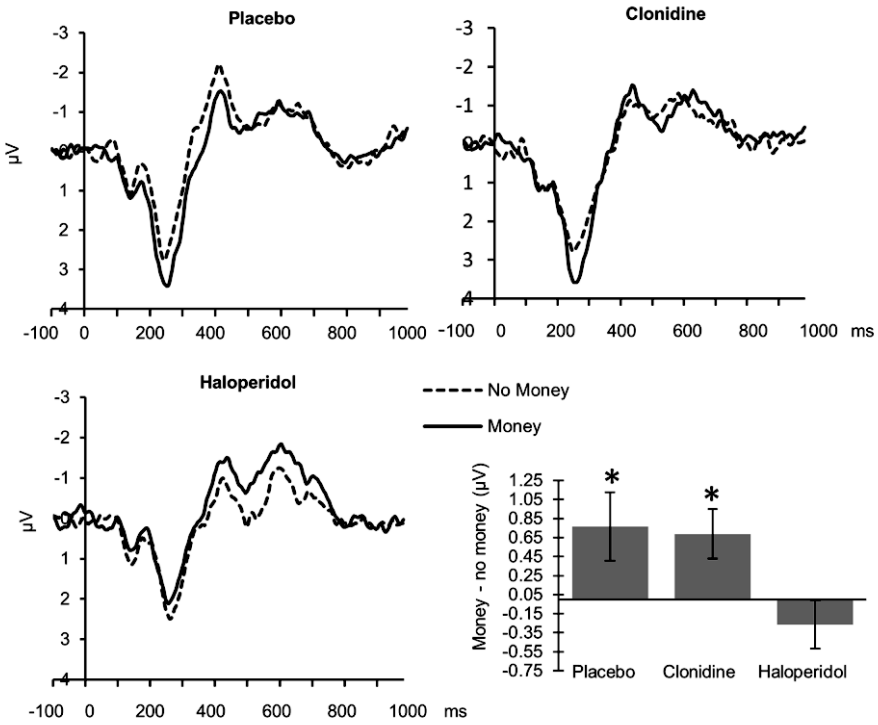


Figure 3. The reward-related positivity displayed for each drug condition. There was a significant drug x reward interaction for the RRP at electrode Fpz (p = .012). The RRP was significantly present for the placebo and clonidine condition (p = .02, and p = .023, respectively), but not for the haloperidol condition (p = .34). The data shown have been averaged across a-priori and collapsed localizer (CLA) selection methods. The bar graph displays the average reward-no reward difference averaged across electrode FPz 199-280 ms (a-priori method) and electrode Fpz 240-280 ms (CLA). Error bars represent ± 1 standard error. * p< .05.

Drug effects on the reward P300

The reward P300 was significantly present across drug conditions when testing on the group level, main effect of reward: $F(1,22) = 6.7$, $p = .017$, $\eta_p^2 = .23$. There was no significant main effect of drug, and also no interaction with drug when testing across all subjects. Figure S6.1 displays the reward P300 for all drug conditions.

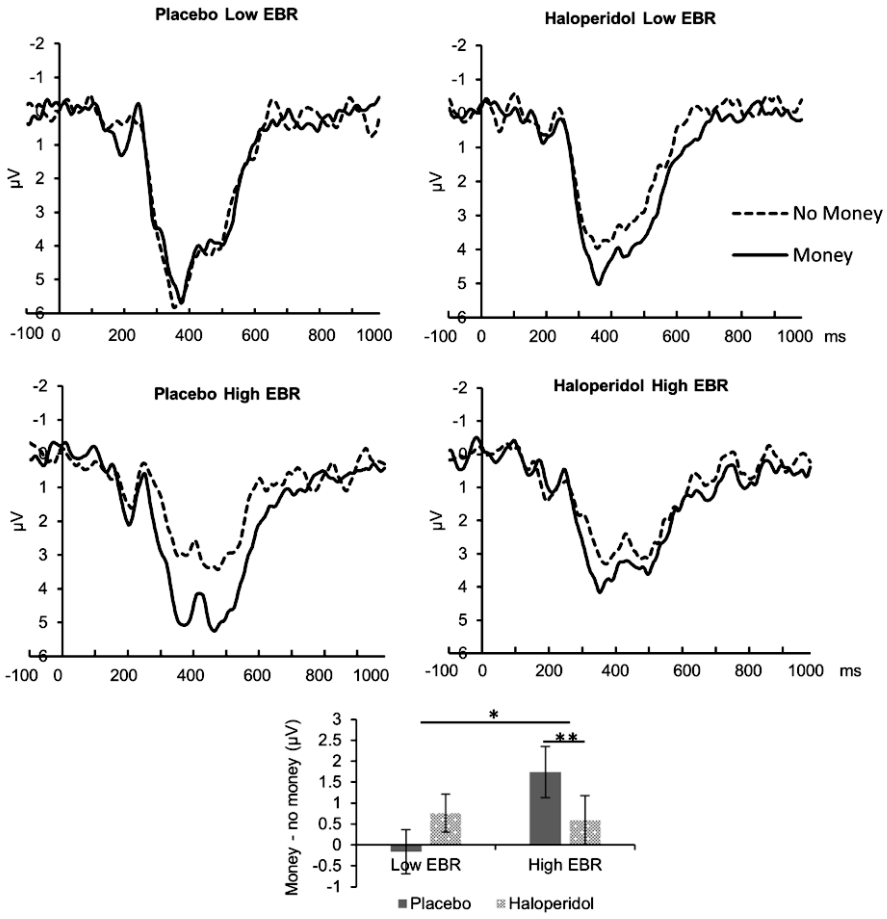


Figure 4. Drug effects on the reward P300 for subjects with high and low baseline dopamine activity. The effect of reward on the P300 was significantly attenuated under haloperidol compared to placebo, but only in subjects with high baseline dopamine activity (bottom row; high EBR). There was no such effect for subjects with low baseline dopamine activity (top row; low EBR). The data shown have been averaged across a-priori and collapsed localizer (CLA) selection methods. The bar graph displays the average reward-no reward difference averaged across selection method 1 (i.e., electrode CPz, 363-526 ms) and method 2 (i.e., averaged signal of electrodes CPz-Pz-POz-Oz, 281-526 ms), for subjects with high and low baseline dopamine separately. Error bars represent ± 1 standard error. * $p < .05$, ** $p < .01$.

However, as shown in Figure 4, the median-split analysis revealed a dissociation between the effect of haloperidol versus placebo on the reward P300 in subjects with high and low baseline dopamine activity, $F(1,11) = 6.7$, $p = .025$, $\eta_p^2 = .38$ (drug x reward x EBR group). Haloperidol significantly attenuated the reward P300 relative to placebo in subjects with high

baseline dopamine activity, $F(1,5) = 18$, $p = .008$, $\eta_p^2 = .78$ (drug x reward), reward > no reward placebo: $p = .011$, $\eta_p^2 = .76$, reward > no reward haloperidol: $p = .394$. There was no significant difference between haloperidol and placebo with respect to the reward P300 amplitude in subjects with low baseline dopamine activity, $F(1,6) = 1.2$, $p = .315$ (drug x reward).

There was no such dissociation between drug responders and non-responders (based on movement increase following haloperidol), $p = .831$.

Drug effects on the probability ERP

There was no main effect of target probability on the cue-elicited ERP in the expected time window and sensor space. There was however a significant interaction between probability and drug (Figure 5), $F(2,21) = 5$, $p=.017$, $\eta_p^2 = .32$. The probability effect was not significant (although in the expected direction, i.e., less positivity with low probability) under placebo ($p = .346$) or haloperidol ($p = .486$). It was significant, but in the opposite direction (less positivity with high probability) under clonidine $p = .024$, $\eta_p^2 = .21$.

There was no significant interaction between drug, probability, and systolic blood pressure increases, $p = .114$.

Discussion

The current study supports the hypothesis that cortical mechanisms associated with the processing of reward value and of probability are directly related to the dopamine and noradrenaline neurotransmitter systems, respectively.

Cortical mechanisms related to the processing of prospective reward value reflected by the reward P300 and RRP were significantly reduced by haloperidol, as expected. The reduction of the RRP as elicited by reward cues is consistent with prior studies showing that DA antagonists reduce the amplitude of the feedback-related negativity (FRN) (De Bruijn et al., 2006; Forster et al., 2017; Zirnheld et al., 2004). The FRN is an ERP component that is more negative following reward omission relative to reward obtainment, and is possibly functionally related to the RRP.

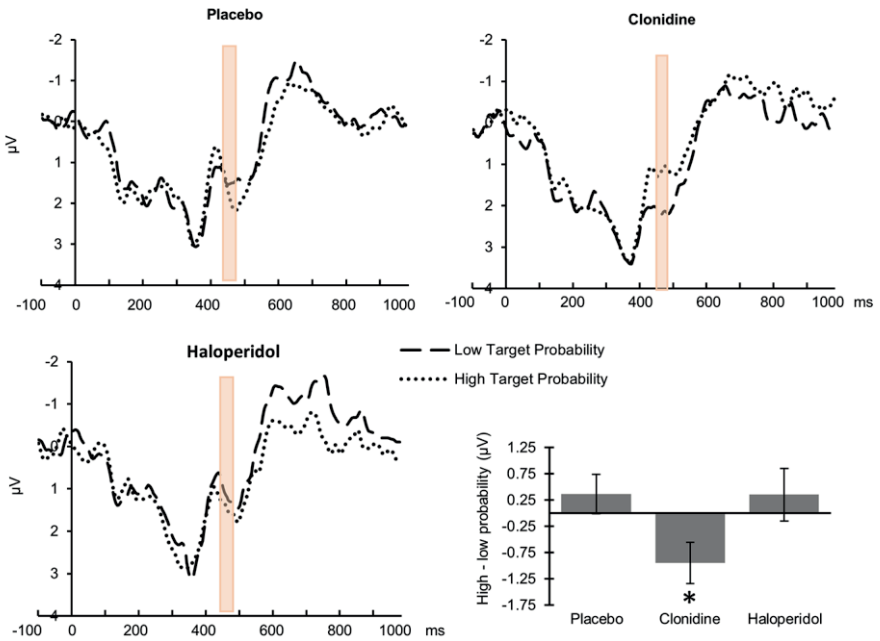


Figure 5. Drug effects on the probability-related positivity. A significant drug x target probability interaction was observed within the target time window of the PRP (displayed in pink) at electrode FCz ($p = 017$). ERP activity was significantly decreased for the high compared to low target probability blocks under clonidine. This pattern was reversed (but not significantly different) for the other two drug conditions. The data shown have been averaged across a-priori and collapsed localizer (CLA) selection methods. The bar graph displays the average high-low target probability difference for electrode FCz between 445-485 ms. Error bars represent ± 1 standard error. * $p < .05$.

The attenuation of the reward P300 by haloperidol was only significant for participants with a high proxy for baseline DA levels (EBR). The use of EBR as an indirect measure of tonic (baseline) striatal DA activity is well-established (Jongkees & Colzato, 2016). The link between EBR and striatal DA activity is supported for example by evidence from genetic studies showing a relationship between EBR and DA gene polymorphisms (Dreisbach et al., 2005), and by effects of DA drugs on EBR (Blin, Masson, Azulay, Fondarai, & Serratrice, 1990; Karson, 1983).

The reduction of the reward P300 specifically in subjects with high EBR is consistent with the notion that the effects of DA medication may differ between individuals depending on baseline DA level (Cools et al.,

2009; Frank & O'reilly, 2006; Martins, Mehta, & Prata, 2017). This finding is particularly interesting in light of the relatively greater efficacy of DA antagonists in patients with high baseline DA levels such as seen in patients with psychotic symptoms (Abi-Dargham et al., 2000). Individual differences in the effect of DA drugs depending on baseline DA may explain the mixed findings in the literature (Frank & O'reilly, 2006) with respect to the direction of the DA effects. It has been suggested that single low doses of haloperidol including 2 mg exclusively stimulate presynaptic D2 receptors and therefore result in an acute increase of DA transmission (Frank & O'reilly, 2006). However, the current data as well as numerous other experimental results (e.g. reduced approach (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), reduced inhibitory control (Logemann et al., 2017) and increased prolactine levels (Frank & O'reilly, 2006) after haloperidol 1 and 2 mg, respectively) indicate that this may not hold so much in general. The pre/post synaptic effect balance may be a matter of differences in baseline DA activity. For example, in high baseline DA subjects there may be relatively more postsynaptic binding because of presynaptic D2 receptor occupancy by endogenous DA. The current results suggest, however, that there was more room for reduction of the P300 amplitude by haloperidol in the high baseline DA group, because of a relatively increased reward P300 under placebo in the high compared to the low baseline DA group. In any case, the observed relationship between DA drug effects and a proxy for endogenous DA demonstrate that it is useful to take endogenous DA into account in future studies.

The ERP component associated with the processing of the probability of a prospective (rewarded) target (probability-related positivity (PRP)) was specifically related to the NA manipulation, as expected. Specifically, the PRP was larger (although not significantly) in the high versus the low probability condition under placebo and haloperidol, and this effect was reversed (i.e., significantly larger in low vs high probability condition) under clonidine. This is consistent with our hypothesis of a PRP being reduced specifically under clonidine. However, this effect pattern also prompts the assumption of an additional influence to account for the absence of a PRP in especially the placebo condition (which was pronounced in our previous study, Chapter 5). One possibility is the difference in gender ratio's between the two studies (all males in the current

study versus a mixed sample in our previous study). Another option refers to context dependency of clonidine effects (Brown, Van der Wee, Van Noorden, Giltay, & Nieuwenhuis, 2015; de Rover et al., 2015). Differences in context between our previous and our current study may concern the current repetition of the various conditions across three sessions (as opposed to just one in the previous study).

With respect to behavioral performance, both drugs reduced detection rates and increased variability in response speed in a non-specific manner, indicative of reduced attention (Logemann et al., 2017). This pattern of behavior may alternatively be explained in terms of increased fatigue, as both drugs induced increased subjective feelings of tiredness compared to placebo. The drugs did, however, not significantly slow reaction times, which speaks against this interpretation in terms of increased fatigue. Furthermore, the attenuation of the reward P300 following haloperidol is not easily explained in terms of decreased attention or increased fatigue as this effect was specific for a group of subjects with high baseline dopamine levels. This suggests that it reflects direct effects of the drug on systems underlying this ERP rather than reflecting a general effect on attention/ alertness during task performance.

A limitation of the current study concerns the repeated sessions, which may in its own right may have altered the pattern of the PRP. Also, and in contrast to our previous study, the exclusive focus on male participants may be seen as a limitation. A future replication should perhaps employ a between-subjects design and included female participants, taking into account natural hormonal fluctuations.

In conclusion, the current study demonstrates that cortical processes related to the anticipation of reward value and probability relate to the dopaminergic and noradrenergic system, respectively. These results may have implications for the pharmacological treatment of patients displaying problems with reward processing and decision-making, such as patients with depression, schizophrenia, and ADHD. Treatments for these disorders often target the NA or DA system (El Mansari et al., 2010). Future studies could investigate whether reward and probability ERPs predict the effectiveness of DA and NA treatment in these patients. For example, patients with an attenuated reward P300 may benefit more from compounds targeting the DA system. Furthermore, the current study

demonstrates that it is useful to take baseline DA functioning into account when investigating the effects of DA drugs.

Supplementary materials

1. Sample size justification

We aimed at a final sample of 28 subjects plus two extra subjects (for drug order balancing purposes). This sample size (n=28) was based on a prior study in which the effects of haloperidol 2 mg were compared to placebo on stop signal reaction times (SSRTs) (Logemann et al., 2017). In this study haloperidol was found to significantly increase SSRTs. Due to time/ end of contract restriction we were forced to at least temporarily suspend inclusion at n=26 and performed interim analyses. This resulted in 3 further exclusions. Given the results with respect to the hypothesized drug x task factor effects we decided that it was no longer necessary to include (and burden) further participants.

2. Cued Go/NoGo task

During this task the letters A, C, D, E, F, G, H, J, L, X and Y (in black, font size 79) were presented in the center of a grey colored screen and between two vertical bars (height: 1.03º, width: 0.05º). Letter stimuli were presented for 150 ms followed by inter-trial intervals with a random duration between 1400 and 1600 ms. Subjects had to press a left or right button when a target letter (letter X or Y) followed a cue letter (always letter A). Each of the target letters (X, Y) was associated with either the left or right button. This mapping was counterbalanced across participants. The response buttons were the "z" and "/" key of a qwerty keyboard. The letters of the keyboard were covered by a sheet and the target buttons were covered by a white sticker. Subjects were instructed to respond (or not) as fast and accurately as possible. The probability of target appearance after the cue (either 50 % or 98 %) and the amount of money that could be won for correct and fast responses (a total of either 0 Euros or 5 (or 2.5) Euros during the block) were orthogonally manipulated across four task blocks. This information was shown to participants at the beginning of each block. Participants could win 2.50 Euros during reward blocks of the pre-drug task version, because this task version was twice as short as the post-drug task version.

During the first session of the study and before the pre-drug cued Go/NoGo task version participants received a practice block consisting of

100 trials (letters). The practice block always consisted of the reward-98% target probability condition. During the second and third session participants received a shorter practice block consisting of 25 trials. The pre-drug cued Go/NoGo task version consisted of four blocks of 200 letter trials and the post-drug cued Go/NoGo task version consisted of four blocks of 400 letter trials. Participants were assigned to one of four possible task block orders. For each participant this order was kept the same for the pre- and post-version of the cued Go/NoGo task and for all three sessions. Task block order was counterbalanced across participants and drug order.

One-minute rest breaks were provided between task blocks and halfway through each task block. Breaks within blocks were followed by a reminder of the target probability and reward availability of the current block.

In both the pre- and post- cued Go/NoGo task cues appeared with a frequency of 20 % during each block. Both the X and Y appeared with a frequency of 10%. In the 50% target probability blocks, half of the cues were followed by a target (50% X, 50% Y). The other half of the cues were not followed by a target (X or Y) and 50% of the X and Y stimuli were not preceded by a cue. Similarly, in the 98% target probability blocks, 98% of the cues were followed by a target (50% X, 50% Y). Two percent of the cues were not followed by a target and 2% of the targets (50% X, 50%Y) were not preceded by a cue. All other letters appeared with a frequency of 5%, except for letter H and C. These letters appeared with a frequency of 20% and 10%, respectively, in order to control for the frequency differences between cues, targets and other letters (Bekker, Kenemans, & Verbaten, 2004).

Letters were presented in a pseudo-random order, with the following restrictions: (1) the same letter was never repeated on the subsequent trial, (2) sequences of cues followed by targets (A-X, A-Y) and of cues not followed by targets were never followed by A, X, or Y.

3. Subjective measures

Methods

Subjective effects of the drugs on mood were assessed by the Dutch short version of the Profile of Mood States (POMS) questionnaire (Wald &

Mellenbergh, 1990). This self-report questionnaire consists of 32 statements describing feelings like “sad” and “tired” and comprises six subscales: depression, fatigue, tension, anger, vigor and total mood disturbance. Each item is rated on a scale ranging from “not” (score 0) to “very much” (score 4). The total mood disturbance score was computed by subtracting the total score from the vigor subscale from the total score of the other subscales.

Data from one subject were removed from the anger subscale, data from two subjects were removed from the fatigue subscale, and data from three subjects were removed from the total mood disturbance scale, because of incomplete responses.

MANOVAs (GLM, SPSS version 22) were run with drug and time (pre-, post-medication) as within-subjects variables.

Results

Figure S3.1 displays the subjective mood scores over time for each drug condition. Only the levels of fatigue, vigor, and total mood disturbance were significantly affected by the drugs. Specifically, participants felt significantly more tired after administration of clonidine and haloperidol, but not after placebo, $F(4,17) = 5.8$, $p = .004$, $\eta_p^2 = .58$ (time x drug), main effect of time haloperidol: $p = .008$, $\eta_p^2 = .40$, main effect of time clonidine: $p < .001$, $\eta_p^2 = .64$. The level of tiredness under clonidine and haloperidol was increased for the post1 compared to pre-drug time points (p’s ≤ .011). The increase was significantly stronger for clonidine compared to placebo and haloperidol (p’s < .001), and did not significantly differ between haloperidol and placebo ($p = .151$).

Mood disturbance scores increased significantly after drug administration for all drug conditions, but the effect was most pronounced for clonidine, which explains the time x drug interaction, $F(4,16) = 5.4$, $p = .006$, $\eta_p^2 = .57$. The increase from the pre to post1 time point was significantly stronger for clonidine compared to placebo and haloperidol (p’s ≤ .026). There was no such difference between haloperidol and placebo ($p = .65$).

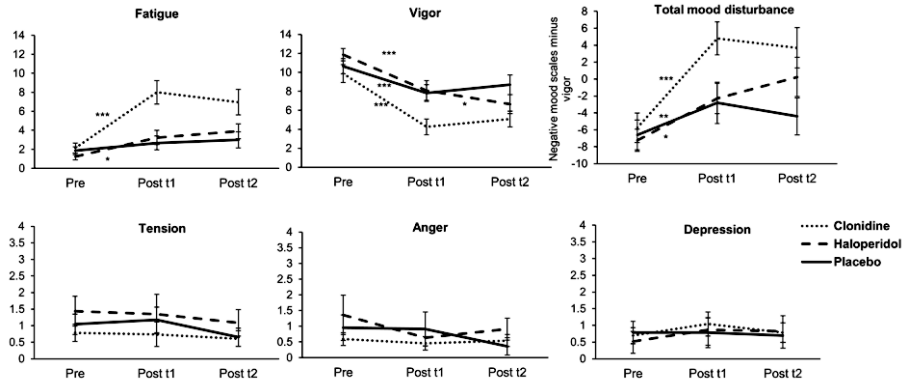


Figure S3.1 Subjective mood scores. Fatigue, vigor and total mood disturbance scores changed over time dependent on the drug condition. Error bars represent 1 standard error. Star symbols mark significant differences between two time points within a drug condition. * $p < .05$, ** $p < .01$, *** $p < .001$.

Subjects felt less vigorous after drug administration in all drug conditions. However, the drop in vigor was most pronounced for the clonidine condition explaining the time x drug interaction, $F(4,19) = 5.1$, $p = .006$, $\eta_p^2 = .52$. A stronger drop in vigor between the pre and post1 time point was observed for clonidine compared to placebo ($p = .001$), but there was no significant difference between haloperidol and, respectively, clonidine and placebo. The level of vigor stabilized from the post1 time point onwards under clonidine and placebo, whereas it dropped further under haloperidol (post1 – post2 time point under haloperidol: $p = .021$).

4. Peripheral control measures

Spontaneous movements and eye blinks

Additional MANOVAs were run to investigate whether the proxy for individual variance in the effect of haloperidol (spontaneous movements) and the proxy for individual variance in endogenous DA (eye blink rate, (EBR)) were affected by the drugs on the group-level. We expected spontaneous movements to increase after haloperidol treatment compared to placebo and clonidine. EBR was expected to decrease specifically after

haloperidol treatment. MANOVAs were run with drug and time (pre-, post-medication) as within-subjects variables.

Spontaneous movements were not significantly increased following haloperidol treatment (time x drug: $F < 1$, $p = .477$). The number of eye blinks during resting state was also not significantly different under haloperidol treatment compared to placebo and clonidine ($F(2,21) = 1.3$, $p = .284$). Figure S4.1 displays EBR and spontaneous motor activity during each drug condition. As outlined above, spontaneous movements and EBR were not significantly affected by haloperidol when testing at the group-level. The effects of haloperidol on spontaneous movements were in the expected direction (post-treatment increase relative to placebo and clonidine). The non-significant effects may be explained by sizeable individual variance in e.g. pharmacokinetics, D2 receptor availability, and endogenous DA, possibly resulting in strong individual variance in the effect of haloperidol on these parameters. At least for EBR strong individual variance in the effects of DA drugs (DA agonist treatment) has been observed before (Cavanagh, Masters, Bath, & Frank, 2014).

Cardiovascular data

Additional MANOVAs with drug and time (pre-, post-t1, post-t2 time point) as within-subjects variables were run for systolic and diastolic blood pressure and heart rate in order to test whether these parameters were affected by clonidine at the group-level. Time point T2 after drug administration was included in the test in order to examine whether the effects of clonidine on systolic blood pressure were evident until after the cued Go/Nogo task. We expected that clonidine would attenuate systolic and diastolic blood pressure. No effect was expected with respect to heart rate.

A significant time x drug interaction was found for systolic and diastolic blood pressure, $F(4,19) = 17.7$, $p < .001$, $\eta_p^2 = .79$; $F(4,19) = 22.1$, $p < .001$, $\eta_p^2 = .82$, respectively. Systolic blood pressure was significantly attenuated specifically following clonidine administration (Figure S4.1). None of the treatments significantly affected heart rate.

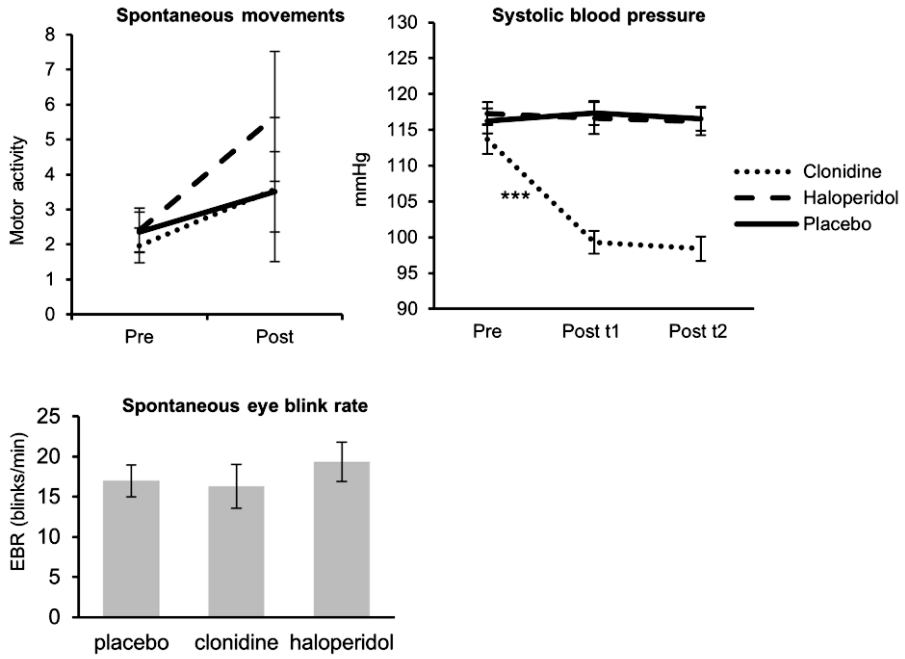


Figure S4.1. Peripheral control measures. Top left: spontaneous motor activity measured during the cued Go/NoGo task before (pre) and 3 hours after (post) drug administration. Data were collapsed across the x, y, and z movement direction. Top right: systolic blood pressure for each drug condition measured at 3 time points: before drug administration (Pre), at t = 2 h 47 min after drug administration (post t1), and at t = 4 h 22 min after drug administration (post t2). Bottom left: eye blink rate in blinks per minute for each drug condition. EBR was measured at t = 2 h 40 min after drug administration. Error bars represent 1 standard error. *** $p < .001$.

The effects of clonidine on cardiovascular variables were as expected. Figure S4.1 (right panel) shows that the attenuation of systolic blood pressure was evident until after the cued Go/NoGo task (time point T2). Systolic blood pressure remained stable between the time point three hours after drug administration (just before the cued Go/Nogo task) and time point T2 (after the task).

5. The effects of reward and probability on behavioral parameters

With respect to main effects of reward value and probability (Figure S5.1), subjects responded faster and were more accurate in the reward compared to the no reward condition; $F(1,21) = 30.7$, $p < .001$, $\eta_p^2 = .59$; $F(1,21) = 9.3$, $p = .006$, $\eta_p^2 = .31$, respectively. RTs were also less variable and the percentage omissions was lower for the reward compared to the no reward condition, especially after treatment (time x reward interaction). Main effect of reward post-treatment: $F(1,21) = 29.0$, $p < .001$, $\eta_p^2 = .58$ (RT variability); $F(1,21) = 7.8$, $p = .011$, $\eta_p^2 = .27$ (omissions). RTs were also shorter and less variable in the high compared to the low probability condition; $F(1,21) = 46.0$, $p < .001$, $\eta_p^2 = .69$; $F(1,21) = 9.2$, $p = .006$, $\eta_p^2 = .30$, respectively.

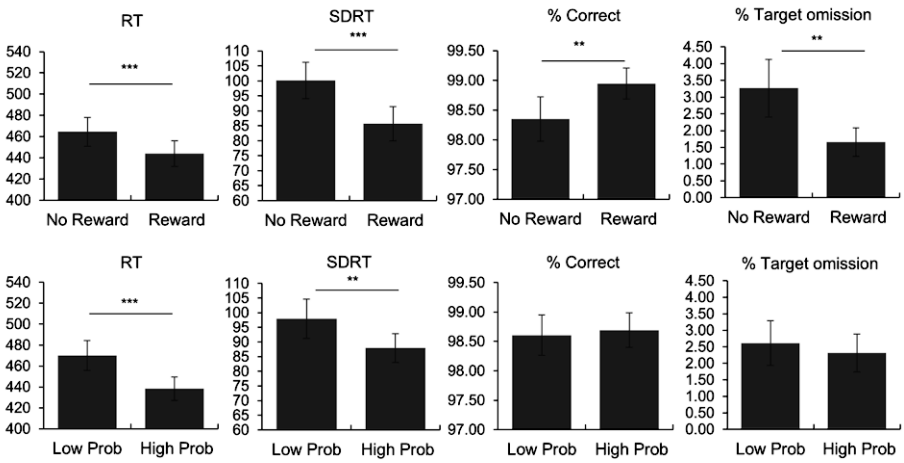


Figure S5.1. Effects of reward and probability on behavioral parameters. The figure presents the main effects of reward (top row) and target probability (bottom row) on reaction times, reaction time variability the percentage correct responses, and the percentage omissions during the cued Go/NoGo task. Data are averaged across all drugs and across the pre and post time points. Error bars represent 1 standard error. * $p < .05$, ** $p < .01$, *** $p < .001$.

6. The reward P300 for each drug condition separately

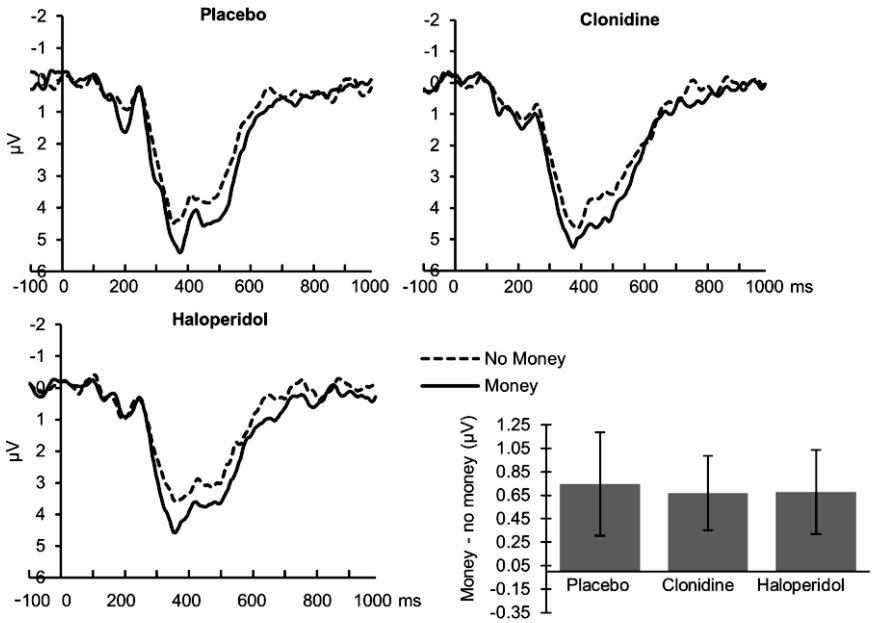


Figure S6.1. Drug effects on the reward P300. P300 activity was in general significantly increased for the money compared to the no money blocks. This reward effect was not affected by drug condition when tested across all subjects. The data shown have been averaged across a-priori and collapsed localizer (CLA) selection methods (see paragraph Cued Go/NoGo task - ERP data –selection of time windows and electrodes of the main paper). The bar graph displays the average reward-no reward difference averaged across selection method 1 (i.e., electrode CPz, 363-526 ms) and method 2 (i.e., averaged signal of electrodes CPz-Pz-POz-Oz, 281-526 ms). Error bars represent ± 1 standard error.

Chapter 7

General Discussion

The aim of this thesis was to elucidate how the human brain processes the value and probability of potential future rewards. In addition, the studies presented in this thesis addressed the topic of individual differences in motivational style and reward-based decision-making, and the assessment of the subjective value of monetary rewards of varying magnitude. More specifically, I addressed the following questions: How can the subjective value of monetary rewards of varying magnitude be assessed? Are there individual differences in reward sensitivity as reflected in learning style? As reflected in spontaneous fluctuations in the human electroencephalogram? What are the cortical substrates of anticipation of reward magnitude and of reward probability? And can these substrates be dissociated neurochemically? This chapter summarizes the extent to which these questions can be answered based on the studies described in the preceding chapters. It also addresses remaining and some new questions and unresolved issues, as well as potential applications and avenues for future research.

7.1 Assessing the subjective value of rewards

The study in Chapter 2 investigated whether increasing reward magnitudes result in an increasing approach tendency. Approach tendency was operationalized as the extent to which the startle reflex was weakened when increasing rewards were anticipated and/ or actually obtained. This methodology is based on the logic of the motivational priming theory (Lang, Bradley, & Cuthbert, 1990). This theory states that a startle reflex is an instantiation of a defensive action tendency, and that this tendency would be reduced when it is incongruent with the motivational state induced by the context in which the startle reflex is initiated (e.g., a context of (prospective) reward obtainment).

In a first experiment I found that, if anything, startle reflexes were enhanced when rewards expected but not obtained (rather than reduced when they were anticipated and/ or obtained). At first sight this pattern appeared to be especially pronounced for the highest reward magnitude. In addition, anticipation of reward (versus no reward) resulted in startle-reflex enhancement non-specifically for all reward magnitudes, suggesting a global increase in arousal rather than a specific increase in approach or

avoidance tendencies. These results were not replicated in a second experiment, in which a design was adopted that, if anything, should result in even stronger startle-reflex modulation with increasing rewards.

Especially in light of the non-replication, it may be concluded that startle-reflex modulation during the anticipation, obtainment, and omission of rewards is not a suitable method to track the subjective value of monetary rewards as used in the studies described in Chapter 2, in spite of some promising results reported previously (Hackley, Muñoz, Hebert, Valle-Inclán, & Vila, 2009; Skolnick & Davidson, 2002). It may be that per-trial rewards of say 20 eurocents (accumulating across all trials) still have higher subjective value on average than e.g. per-trial rewards of 5 eurocents, but that this does not materialize in differential modulation of startle. Startle modulation has been conceived of as an objective measure of approach tendency, but arousal is necessary for stimuli to cause startle modulation (Bradley, Codispoti, Cuthbert, & Lang, 2001). Classical startle-reflex modulation studies (see e.g., Grillon & Baas, 2003) typically find robust startle potentiation in strongly arousing negative contexts such as pictures of mutilated bodies, of threat with electric shock. Startle attenuation is typically found with positive stimuli but such attenuation is less robust than potentiation with negative conditions (Grillon & Baas, 2003). Attenuation is strongest with erotic pictures (Bradley et al., 2001; Gard, Gard, Mehta, Kring, & Patrick, 2007), which may perhaps also relate to the fact that erotic pictures are of a more primary reward nature (Rothschild & Gaidis, 1981). It is possible that monetary rewards in general, or per-trial rewards varying from 0 to 20 cents specifically are too little arousing to have differential impact at all at the various levels of the human CNS, human cognition, or behavior. With respect to the latter, total absence of impact is unlikely, to the extent that in later studies (Chapters 5 and 6) clear behavioral effects were found when manipulating the possibility of reward relative to no possible reward. In fact, this was also found in the first startle-reflex experiment.

7.2 Individual differences in reward sensitivity

It may be hypothesized that there are individual differences with respect to the sensitivity to reward prospect, and that these differences affect the

electrocortical correlates of reward value anticipation. These electrocortical correlates of reward value anticipation were investigated in Chapters 5 and 6. In this thesis two methods were evaluated for their potential to assess individual differences in reward sensitivity. One approach was to determine whether a given individual learns more from rewards (positive feedback) or more from punishment (negative feedback). The other approach was to evaluate the ratio between spontaneous slow-wave theta and fast-wave beta oscillations in the electro-encephalogram. As will be discussed below, the results of Chapter 3 and 4 indicate that neither of these methods could be used as a reliable index of individual differences in reward sensitivity.

The probabilistic selection task (PST; Chapter 3; Frank, Seeberger, & O'reilly, 2004) was used to assess whether an individual is better at learning from reward feedback (reward sensitive) or better at learning from punishment feedback (punishment sensitive). It has been reported that a low relative to a high dopamine status is associated with learning from punishment feedback (negative learning; Frank et al., 2004; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). However, the results of Chapter 3 show that categorization as positive/negative learner (i.e., being reward/punishment sensitive) may not solely reflect motivational style, but also the perceptual discriminability or salience of stimulus characters associated with high and low reward probability in the PST. This indicates that in order to isolate effects of motivational value, discriminability of the different stimuli needs to be carefully equated. Differences in stimulus features across the stimulus set can be balanced out when regarding group-level data, but can markedly skew individual results. In practice, this means that subjects could be incorrectly classified as being reward sensitive if the most rewarded character in their task version is mapped onto a highly salient stimulus character. Future studies should investigate whether this issue can be controlled by matching the discriminability of the task stimuli at the individual level.

Note that the vast majority of studies in this domain using the PST counterbalanced or randomized the stimulus mappings across subjects, and hence previous replicated findings of individual differences can still be considered valid. This may concern genetics (Cockburn, Collins, & Frank, 2014; Doll, Hutchison, & Frank, 2011; Frank, Moustafa, Haughey, Curran, &

Hutchison, 2007; Klein et al., 2007), but also striatal D1 and D2 dopamine receptor binding (Cox et al., 2015), and neural responses to feedback (Cavanagh, Bismark, Frank, & Allen, 2011; Frank, Woroch, & Curran, 2005; Jocham, Klein, & Ullsperger, 2011; Klein et al., 2007). These effects were found despite stimulus counterbalancing, and were largely present without differences in learning curves (in contrast to the effects of discriminability reported in Chapter 3).

Thus, the results of Chapter 3 indicate that the PST is not suitable for assessing individual differences in reward and punishment sensitivity. In Chapter 4 an alternative approach, theta/beta EEG ratio, was evaluated for its potential to index reward sensitivity. Theta/beta EEG ratio has been hypothesized to reflect the ratio between subcortical (reward) drive and cortical signals that inhibit these drives (Schutter & Van Honk, 2005). Prior studies have shown that high theta/beta EEG ratio is associated with risky decision making in the Iowa Gambling task (Massar, Kenemans, & Schutter, 2014; Schutter & Van Honk, 2005) in which subjects pick cards from decks that differ in reward and punishment contingencies. Subjects gradually learn to pick cards from a deck that is associated with moderate immediate rewards, and a small amount of punishment. In these studies it was shown that subjects with high theta/beta EEG ratios keep choosing from a deck that is associated with large immediate rewards, but also large punishment in the long run. The results of the study by Massar and colleagues (2014) suggest that poor decision making as observed in individuals with high theta/beta EEG ratio reflects a strong reward drive. From this perspective, it could be reasoned that high theta/beta EEG ratios are associated with strong representations of reward value. As such, theta/beta EEG ratio could be a good candidate for the classification of subjects as being reward or punishment sensitive. However, the reasoning in Chapter 4 was that the positive relation between theta/beta EEG ratio and poor decision making in the Iowa Gambling task could also be related to a reduced ability to adapt to contingency reversals. This idea was explicitly tested. Indeed, the results of Chapter 4 show that high theta/beta EEG ratio is not related to high reward sensitivity, but rather to poor behavioral adaptation in a context with changing reward-punishment contingencies. Based on prior studies showing a relationship between high/theta beta EEG ratio and risk taking (Massar et al., 2014; Schutter & Van Honk, 2005) it could be expected that

subjects with a high theta/beta ratio would exhibit a pattern of high risk taking throughout the reversal learning task. In contrast, however, subjects with a high theta/beta EEG ratio showed a pattern of low risk taking during phases of the task in which high risk taking was beneficial. This may indicate that subjects with a high theta/beta EEG ratio are less able to learn to make optimal decisions based on reward and punishment feedback.

The results of Chapter 4 are consistent with previous findings that theta transcranial alternating current stimulation (tACS) at the frontal cortex improves the speed of learning after a contingency reversal. Theta tACS is presumed to facilitate the function of frontal theta oscillatory activity, which has been associated with learning from reward and punishment feedback. Notably, this modulation-induced improvement of reversal leaning coincided with lower theta/beta EEG ratios (Wischnewski, Zerr, & Schutter, 2016). The results of Chapter 4 also fit results of recent studies suggesting that theta/beta EEG ratio is related to PFC-mediated executive control over subcortical responses or attentional control (Angelidis, van der Does, Schakel, & Putman, 2016), including the regulation of attention to fearful stimuli (Angelidis, Hagenaars, van Son, van der Does, & Putman, 2018; van Son et al., 2018; van Son, Angelidis, Hagenaars, van der Does, & Putman, 2018). Together, the results of Chapter 4 and these latter studies suggest that theta/beta EEG ratio does not reflect a brain state that promotes reward seeking per se.

We did not succeed in finding a specific, reliable, and objective measure of reward sensitivity. Therefore, it was decided to proceed with the investigation of reward value-related ERPs without classifying subjects as being reward or punishment sensitive. Another alternative approach from the literature that we did not consider is the assessment of reward sensitivity by means of a signal-detection task (Pizzagalli, Jahn, & O'Shea, 2005). In this task subjects decide whether a cartoon face on the screen has a short or long mouth, and subjects receive feedback accordingly. One of these stimulus types is associated with an increased amount of positive feedback. This leads to a response bias towards the more frequently reinforced stimulus, and the extent of this bias is proposed to index reward sensitivity. Importantly, in contrast to the probabilistic selection task used in Chapter 3 discriminability is not an issue in this signal-detection task as

the stimuli are identical, except for the subtle difference in the length of the mouth of the cartoon face.

7.3 Pinpointing anticipated reward value and probability

What are the cortical substrates of anticipation of reward magnitude and of reward probability? In agreement with the hypothesis, the results of Chapter 5 and 6 indicate that the cortical processes related to the anticipation of reward value and probability are dissociable, statistically, in terms of timing, and neurochemically. In Chapter 5 and 6 it was shown that manipulations of reward value affect a frontally distributed positive-going ERP ('reward-related positivity' or RRP), and a more posterior P300-like ERP ('reward P300'). Manipulations of probability specifically affected a separate frontally distributed ERP ('probability related positivity' or PRP).

As expected, it was consistently found that a frontally distributed positive-going ERP was specifically related to the anticipation of reward value (RRP). As discussed in Chapter 5, this ERP may be identified as an instance of the "reward positivity", which has been proposed to reflect brain mechanisms related to the prediction of reward value (Holroyd, Krigolson, & Lee, 2011). In the studies described in Chapter 5 and 6 the RRP was maximal at frontopolar electrode sites, which is consistent with a prior observation. However, in most studies reporting on the reward positivity elicited either by positive feedback or by reward-announcing cues, this ERP was found to peak more posteriorly, although still over frontal-central midline sites (Doñamayor, Schoenfeld, & Münte, 2012; Flores, Münte, & Doñamayor, 2015; Foti, Weinberg, Dien, & Hajcak, 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Holroyd et al., 2011). For a very recent anticipation example see (Wischnewski & Schutter, 2019). This difference between previous and the current results may reflect subtle differences in the equivalent-dipole (ED) orientation of the involved neural generators reflecting differences between studies with respect to the involved neural populations. The latter however residing in the same global brain region.

A more specific aim of this study was to identify neural correlates of the prediction of reward value. However, the possibility that the RRP as observed in Chapter 5 and 6 reflects sequelae of reward prediction with respect to cognitive control, rather than reward prediction itself cannot be

excluded. It is, thus, possible that the observed positivity actually reflects enhanced attentional capture by the reward cues or a modulatory effect on visual processing rather than reward value prediction (Padmala & Pessoa, 2011; Pessoa & Engelmann, 2010). One way to conceive of the respective roles of these systems is that neurons in subcortical and mesocortical regions are instrumental in representing the reward potential of a given context (Kable & Glimcher, 2007), and signal this to more dorsolateral-cortical neurons where attention settings can be adjusted accordingly. Theoretically, ERP source localization could then distinguish between mesocortical sources indicating reward processing as such, and dorsolateral-cortex sources indicating sequelae of reward processing with respect to cognitive control. Initial attempts applying this strategy have indicated either mesocortical (posterior cingulate gyrus; (Doñamayor et al., 2012)) or even subcortical (Foti et al., 2011) generators of RRPs. This may indicate that the RRP as observed in the current study reflects reward processing as such, not so much its sequela in cognitive control.

A further initial hypothesis was that a posteriorly distributed P300-like ERP is sensitive to manipulations of both probability and reward value. In line with this prediction, in Chapter 5 and 6 it was found that the posterior P300 was increased for the reward compared to no reward condition. However, in contrast to what was expected, manipulations of probability did not affect the posterior P300, but instead affected a more frontally distributed ERP (the probability-related positivity (PRP)). As argued in Chapter 5, it is possible that the nature of the effects of cued probability is different in the context of explicit reward compared to contexts without explicit reward. One example is the study by Bekker and colleagues (Bekker, Kenemans, & Verbaten, 2004) where subjects "simply" performed a cognitive task, i.e., a similar cued Go/Nogo task as used in Chapter 5 and 6, but without the reward and probability manipulations. The only reward in their task version was perhaps the prospect of financial or course-credit revenue after the end of the experiment. In that study target probability had a specific and significant effect on the amplitude of the cue-elicited parietal P300. It may be speculated that the anticipation of probability of a target in the context of explicit reward that is or is not available engages prefrontal rather than or in addition to posterior areas such as the PCC. This idea is consistent with the results of a prior study by

Knutson et al. (Knutson, Taylor, Kaufman, Peterson, & Glover, 2005) showing that the anticipation of high compared to low reward probability was associated with increased activations in the medial PFC, but this idea remains to be tested.

The PRP was not significant under placebo in the study described in Chapter 6, which may have been due to differences in the experimental setup between both studies. In the study of Chapter 5 participants were subjected to the task once, while in the study of Chapter 6 they were subjected to the task multiple times across three sessions. The difference in significance of the PRP may have also been due to differences in gender ratios between the studies.

As to behavioral performance, in both Chapter 5 and 6 it was found that reward anticipation and high probability speeded up reactions to subsequent target stimuli in an additive manner. This is consistent with prior reports (Bekker et al., 2004; Doñamayor et al., 2012; Flores et al., 2015; Pfabigan et al., 2014) and confirms the differential processing of cues in the different reward and probability conditions: the preceding cues must have been processed differentially as a function of reward and probability prospects. The additivity of the reward and probability effects further suggests that these reward and probability prospects affect separable cortical processes, which was indeed confirmed by the ERP results as described above.

7.4 A pharmacological dissociation of reward value and probability anticipation

The results of Chapter 5 demonstrate that cortical processes related to anticipation of reward value and probability are dissociable, as they differ in timing and/ or scalp topography they produce. The aim of Chapter 6 was to investigate whether these separate processes can also be dissociated in terms of underlying catecholaminergic neurotransmitters. Consistent with what was hypothesized, the results of Chapter 6 indicate that the anticipation of reward value and probability are associated with, respectively, dopaminergic and noradrenergic transmission. Specifically, it was found that 2 mg haloperidol, aimed at antagonizing dopaminergic (D2) transmission, reduced the RRP and, although only in high-baseline-

dopamine individuals, the reward P300. These effects were relative to placebo and specific to haloperidol, i.e., they did not occur for clonidine. Conversely, 150 μg clonidine, aimed at net antagonizing noradrenergic transmission, reduced the PRP, which was not affected by haloperidol (both relative to placebo). More precisely, the PRP actually reversed polarity under clonidine, while it was not significant under placebo nor haloperidol. We expected that manipulations of reward probability would affect the parietal P300, and that this effect would be sensitive to the clonidine manipulation. However, we observed effects of reward probability on a separate frontally distributed ERP (PRP), which we found to be sensitive to noradrenergic antagonism. As mentioned above, in Chapter 6 the PRP (ERP activity being more positive for the high compared to the low target probability condition) was absent under placebo and haloperidol, whereas it was pronounced in Chapter 5. Furthermore, clonidine reversed the polarity of this ERP. That is, ERP activity became more negative for the high compared to the low probability condition under clonidine. This pattern of results suggests that there was an additional factor in the experiment of Chapter 6 which caused a general reduction of the PRP. One possibility is that the context of repeated sessions in the experiment of Chapter 6 caused annihilation of the PRP and that clonidine pushed the PRP further into a negativity.

As noted in the previous section, in both chapters 5 and 6 reward and high probability induced shorter RTs to subsequent target stimuli. Anticipatory processing of reward value and probability was disrupted by the drugs, but surprisingly neither of the drugs caused specific concomitant performance decreases (i.e., a reduction of the reward-related or probability-related effect on performance). Both haloperidol and clonidine, relative to placebo, enhanced variability of responding (probably due to more occasional very long reaction times) and omission rate. These observations are indicative of an increased number of lapses in attention towards the task as a whole (Logemann et al., 2017). This is at odds with a prior study in which it was found that haloperidol had a specific negative effect on reward-related behavioral performance increases e.g., Pleger et al., 2009. It is unclear why haloperidol and clonidine administration did not result in a reduction of, respectively, reward and probability-related

performance increases in the present study. This issue should be resolved in future studies.

7.5 Clinical implications

Together, the findings of Chaper 5 and 6 reveal a dissociation between anticipatory processing of reward value related to dopaminergic and anticipatory processing of probability related to noradrenergic transmission. These findings may have important scientific and clinical implications. The studies in chapters 5 and 6 demonstrate that the anticipation of reward involves at least three distinct neural processes specifically related to either the coding of value (reward-related positivity (RRP) and reward P300) or the coding of probability (probability-positivity (PRP)), each of which may or may not contribute to the behavioral phenotype of psychiatric patients with motivational deficits. Interesting in this respect is a study demonstrating that negative symptoms in schizophrenia are related to a reduced propensity to exert greater effort for more probable rewards, but not to a reduced propensity to exert greater effort for more valuable rewards (Barch, Treadway, & Schoen, 2014). It may therefore be speculated that the coding of probability, rather than the coding of value of forthcoming reward is affected in schizophrenic patients with negative symptoms. Future studies could investigate whether disturbances in any of these aforementioned processes, as putatively indexed by reductions or increases of the associated ERPs, are related to specific symptoms in psychiatry such as impulsivity or negative symptoms, including anhedonia. For example, it is worth investigating whether a greater severity of negative symptoms in schizophrenia is associated with a relatively stronger reduction of the amplitude of the PRP.

Depression is associated with decreased reward sensitivity and hypersensitivity to punishment. It may be speculated that these traits specifically relate to impairments in the coding of prospective reward value rather than probability in (a subset of) patients with depression (Foti & Hajcak, 2009). As such, the investigation of whether decreased reward and increased punishment sensitivity in patients with depression relates to specific reductions of the amplitude of the cue-induced RRP and reward P300 may be worthwhile. As a first step in this direction the amplitude of

these reward-value related ERPs could be related to individual differences in reward and punishment sensitivity in healthy subjects (see Sections 7.2 and 7.5).

This thesis furthermore demonstrates that the reward value-related and probability-related ERPs respond differentially to different treatments. If motivational deficits in psychiatry are indeed associated with alterations of the amplitude of the RRP, reward P300, and/or PRP this result may have implications for treatment choice. For example, patients with specific deficits in the coding of prospective reward value, putatively indexed by reductions/increases of the RRP and/or reward P300, may benefit most from compounds targeting the dopamine system. Patients with deficits in the coding of reward probability, putatively indexed by reductions of the probability-related positivity, may instead benefit more from compounds affecting noradrenergic signaling. This is an interesting avenue for future research.

7.6 Future directions

The results discussed in this thesis suggest several avenues for future research. First, an open question that I did not succeed in addressing in this thesis is whether the amplitude of the reward-related positivity and reward P300 relates to reward sensitivity. My hypothesis is that subjects with high reward sensitivity have stronger representations of prospective reward value, as indexed by larger reward-related positivities (RRP) and reward P300 ERPs. As noted in paragraph 7.2, the task developed by Pizzagalli et al. (2005) may prove useful for the categorization of subjects as being more or less reward sensitive. Performance in this task has already been related to the magnitude of the reward positivity as elicited by positive feedback. More specifically, high reward sensitivity was found to be associated with larger feedback-related reward positivities (Bress & Hajcak, 2013).

Second, recent studies have shown that the magnitude of the reward positivity as elicited by reward feedback predicts depressive symptoms and risk for depression (Foti & Hajcak, 2009; Foti, Carlson, Sauder, & Proudfit, 2014; Kujawa, Proudfit, & Klein, 2014; Proudfit, 2015). Future studies could investigate whether the reward-related positivity, the reward P300, and

probability ERP as elicited by predictive cues signaling reward value and probability also relate to specific symptoms in psychiatry.

Third, the results of the study described in Chapter 6 demonstrate that the effect of dopamine antagonism on the reward P300 depends on baseline dopamine transmission, which is in line by an earlier report (Cools & D'Esposito, 2011). It is known that phasic noradrenergic responses also relate to baseline noradrenergic firing (Berridge & Waterhouse, 2003). It is therefore possible that the effect of noradrenergic antagonism on the cue-elicited probability ERP is different for subjects with high and low baseline noradrenergic tone. This could be tested in future studies. Trait anxiety has been found to correlate with baseline noradrenergic tone, and hence it may be used as a proxy for baseline noradrenergic signaling in future studies (de Rover et al., 2015).

Finally, I mentioned in the introduction that neurons in frontal and parietal regions in monkeys that are associated with cognitive control respond to information about reward and probability (Platt & Glimcher, 1999). The hypothesis in this thesis was that the activity in these fronto-parietal structures is driven by input from other structures that actually compute prospective reward value and probability. The fronto-parietal network is for example concerned with the direction of attention to locations in space expected to contain relevant information. In future studies it would be worth to investigate whether prospective reward value and probability affect the strength with which this cognitive control function is exerted. This could be tested by evaluating the effect of reward and probability cues on ERP responses associated with the selective direction of attention to a certain location (van der Lubbe, Neggers, Verleger, & Kenemans, 2006).

7.7 Summary/conclusion

The studies presented in this thesis focused on how the human brain anticipates rewards. The aim of Chapter 2 was to investigate how the subjective value of rewards can be assessed. The startle reflex was used as a readout measure of motivational state during the anticipation, obtainment and omission of rewards with varying magnitudes in order to compare the value of these rewards. However, no evidence was found for systematic

increases in startle modulation with increased reward magnitude at stake. It may therefore be concluded that startle-reflex modulation during the anticipation, obtainment, and omission of rewards is not a suitable method to track the subjective value of monetary rewards. The aim of Chapter 3 and 4 was to investigate whether there are individual differences in reward sensitivity as reflected in theta/beta EEG ratio and in the extent to which subjects learn from reward and punishment feedback as assessed by the probabilistic selection task (PST). The results of Chapter 3 and 4 indicate that neither of these methods is suitable for reliably assessing individual differences in reward sensitivity. The results of Chapter 3 indicate that classification as positive or negative learner may be biased by salience or discriminability of the stimuli in the PST. The results of Chapter 4 show that high theta/beta EEG ratio is not associated with reward seeking per se, but rather with poor behavioral adaptation in a context with changing reward-punishment contingencies. Pizzagalli's signal detection task (Pizzagalli et al., 2005) may prove useful for assessing individual differences in reward sensitivity. The aim of Chapter 5 and 6 was to clarify how the human brain processes the value and probability of potential future rewards. The results in these chapters demonstrate that the anticipation of reward involves at least three distinct neural processes specifically related to either the coding of value (reward-related positivity (RRP) and reward P300) or the coding of probability (probability-positivity (PRP)). The value-related processes were found to be under control of dopamine, whereas the probability-related process was found to be under control of noradrenaline. Future studies could investigate whether reductions or increases in the amplitude of these event-related potentials relate to specific symptoms in psychiatry. It could for example be investigated whether depressive symptoms are associated with a reduction of the value-related ERPs. As a first step in this direction, future studies could examine the relationship between the amplitude of the RRP and reward P300 and reward sensitivity in healthy subjects.

Summary in Dutch/ Nederlandse samenvatting

Mensen en de meeste andere organismen hebben een aangeboren neiging om beloning na te streven en straf te vermijden. Een beloning kan worden gedefinieerd als een object dat, of activiteit of situatie die: positieve emoties kan oproepen, ervoor kan zorgen dat gedrag dat leidt tot de beloning wordt bekrachtigd, en tot toenaderingsgedrag kan leiden. Verder worden activiteiten en situaties die straf doen voorkomen ook gezien als beloningen. Het instinct om beloning na te streven is van cruciaal belang, niet alleen voor het psychologisch welbevinden van het individu, maar ook voor zijn/haar overleving en dat van de nakomelingen. Immers, als we bijvoorbeeld niet geneigd zouden zijn om te eten, te drinken en te schuilen voor gevaar zouden we niet overleven.

Processen in het brein die geassocieerd zijn met het verlangen naar een bepaalde beloning sturen het gedrag zodanig dat de kans op het verkrijgen van de beloning wordt vergroot. Om optimale beslissingen te kunnen nemen over wat te doen moet het brein kunnen voorspellen hoe waardevol een bepaalde beloning is en wat de kans is dat deze beloning daadwerkelijk wordt verkregen. Een van de hoofdvragen die in dit proefschrift werd gesteld is: hoe bepaalt het brein wat de waarde en waarschijnlijkheid is van toekomstige beloningen.

Een manier om deze vraag te beantwoorden is om activiteit in de menselijke hersenen bloot te leggen die geassocieerd is met de anticipatie van de waarde en de waarschijnlijkheid van toekomstige beloningen. De basis voor dit onderzoek is het werk van Glimcher (2003). In een studie met apen liet hij zien dat activiteit van neuronen in corticale gebieden die geassocieerd zijn met cognitieve controle (zoals de laterale intra-parietale cortex) werd gemoduleerd door de waarde en waarschijnlijkheid van beloning (voedsel).

Een van de doelen van dit proefschrift was om activiteit in de menselijke cortex (de buitenste schil van de hersenen) gerelateerd aan de anticipatie van respectievelijk waarde en waarschijnlijkheid van een beloning te identificeren, en van elkaar te onderscheiden. Ook is onderzocht of deze respectievelijke processen dissocieerbaar zijn in termen van onderliggende neurotransmitters, de chemische boodschapper-stoffen die noodzakelijk zijn voor hersenactiviteit. Andere onderzoeksvragen die aan bod komen zijn: Hoe kan de subjectieve waarde van geldbeloningen worden bepaald? En zijn er individuele verschillen in de mate van belonings-

gevoeligheid en manifesteert dit zich in verschillen in het leren van positieve en negatieve feedback? En in verschillen in spontane fluctuaties in hersenactiviteit zoals gemeten met behulp van het het electro-encephalogram (EEG)?

De subjectieve waarde van geldbeloningen

Het primaire doel van het onderzoek dat beschreven is in Hoofdstuk 2 van dit proefschrift was om te onderzoeken hoe waardevol variërende geldbedragen zijn voor gezonde proefpersonen. De geldbedragen varieerden tussen 0 en 20 cent voor elke mogelijke enkelvoudige beloning ("trial"), en liepen op tot een totaal tussen 0 en 5.60 euro. Dit heb ik onderzocht door de toenaderingsneiging te meten tijdens een passieve goktaak waarbij verschillende geldbedragen op het spel stonden die vervolgens werden gewonnen of verloren (niet gewonnen). De aanname was dat een sterkere toenaderingsneiging samenhangt met een grotere subjectieve waarde van de beloning. Toenaderingsneiging was geoperationaliseerd in de mate waarin de schrik-reflex werd gedempt tijdens de anticipatie of ontvangst van een geldbeloning. Het is bekent dat een schrik-reflex (opgewekt door een hard geluid) sterker is in een onplezierige context of in de context van gevaar. Het omgekeerde kan het geval zijn in een plezierige context, zoals tijdens de ontvangst van een geldbeloning (Lang, Bradley, & Cuthbert, 1990). De resultaten van Hoofdstuk 2 laten zien dat de schrik-reflex juist vergoot was wanneer er geld op het spel stond, maar niet gewonnen werd (in plaats van gedempt was wanneer het geld wel gewonnen was). Dit patroon leek met name op te treden bij het grootste geldbedrag (20 cent). Tijdens de anticipatie van een mogelijke winst was de schrik-reflex vergroot wanneer er geld op het spel stond in vergelijking met waneer er geen geld op het spel stond. Beide resultaten werden echter niet gerepliceerd in een vervolgonderzoek. Mede door deze non-replicatie is mijn conclusie dat schrik-reflex-modulatie geen geschikte methode is voor het bepalen van de subjectieve waarde van geldbeloningen.

Individuele verschillen in beloningsgevoeligheid

De verwachting was dat er individuele verschillen zijn in de mate van beloningsgevoeligheid en dat deze verschillen samenhangen met de sterkte van de representatie van beloningswaarde in het brein. Meer specifiek, de hypothese was dat corticale potentialen ("*event-related potentials*"; ERPs, zoals gemeten in het eerder genoemde EEG) die geassocieerd zijn met de anticipatie van beloningswaarde groter zouden zijn in individuen met een sterkere beloningsgevoeligheid. In Hoofdstuk 5 en 6 zijn de electro-corticale correlaten van de waarde van een toekomstige beloning onderzocht. In Hoofdstuk 3 en 4 zijn twee potentieel bruikbare methodes voor het meten van individuele verschillen in beloningsgevoeligheid geëvalueerd. Een van deze methodes is het onderzoeken of een individu meer leert van positieve feedback (individu is gevoelig voor beloning) of meer van negatieve feedback (individu is gevoelig voor straf) met behulp van de "*probabilistic selection task*" (PST; Frank et al., 2004). De andere methode is het meten van de verhouding van langzame ten opzichte van snelle fluctuaties in het EEG (theta/beta EEG ratio). De resultaten van Hoofdstuk 3 en 4 laten zien dat aan beide methodes haken en ogen zitten als het gaat om een valide index voor beloningsgevoeligheid.

Tijdens de PST leren proefpersonen stimuli (in dit geval Japanse karakter-tekens) te kiezen die het vaakst leiden tot positieve feedback. Aan de hand van de keuzes in de daaropvolgende testfase kan vervolgens worden afgeleid of een proefpersoon meer heeft geleerd van positieve of negatieve feedback tijdens de leerfase, dus of iemand in hogere mate beloningsgevoelig of strafgevoelig is. De resultaten van Hoofdstuk 3 laten echter zien dat deze uitkomst (of iemand beloningsgevoelig of strafgevoelig is) in sterke mate beinvloed werd door de saillantie of perceptuele discriminabiliteit van stimuli in de PST. Deze resultaten tonen aan dat het voor dergelijke gedragstaken belangrijk is om voor elk individu afzonderlijk de stimuli zodanig af te stemmen dat er geen verschillen zijn in discriminabiliteit en saillantie tussen de stimuli. Uit vervolgonderzoek zal moeten blijken of dit mogelijk is. Het individueel afstemmen van de stimuli is met name van belang wanneer men een uitspraak wil doen over leerprestatie op individueel niveau. In eerdere onderzoeken met de PST zijn leerprestaties van verschillende groepen (bijvoorbeeld patienten versus

controles) met elkaar vergeleken. Hierbij was de koppeling tussen specifieke stimuli en beloning of straf voor elke patient of proefpersoon weer anders. Daardoor kunnen de resultaten met betrekking tot verschillen tussen groepen in beloningsgevoeligheid wèl als valide worden beschouwd.

Uit hoofdstuk 3 blijkt dus dat de PST in zijn huidige vorm niet geschikt is om individuele verschillen in beloningsgevoeligheid vast te stellen. Als alternatief onderzochten we of theta/beta EEG ratio een geschikte maat is voor beloningsgevoeligheid. Theta/beta EEG ratio wordt gezien als de verhouding tussen de zucht naar beloning, gemedieerd door subcorticale (diepe hersen-) gebieden en de remmende invloed hierop vanuit de frontale cortex (Schutter & van Honk, 2005). Uit eerder onderzoek is gebleken dat een hoge theta/beta EEG ratio samenhangt met beloningsgevoeligheid en het nemen van risico's tijdens een gokspel (Schutter & van Honk, 2005; Massar et al., 2014). De resultaten van hoofdstuk 4 laten echter zien dat theta/beta EEG ratio verband houdt met het vermogen om de strategie aan te passen in een gok-context.

In de goktaak van Hoofdstuk 4 kregen proefpersonen steeds twee keuzes voorgelegd (een hoog of laag geldbedrag). In ieder taakblok leidde steeds een keuze vaker tot beloning dan de andere keuze. Tweemaal tijdens de taak moesten proefpersonen hun strategie aanpassen. Op deze momenten moesten zij leren dat de andere keuze vanaf dat moment vaker leidde tot beloning. Deze momenten werden niet aangekondigd. De resultaten van Hoofdstuk 4 laten zien dat proefpersonen met een hoge theta/beta EEG ratio minder goed waren in het aanpassen van hun strategie op deze momenten. Verder vonden we dat individuen met een hoge theta/beta EEG ratio vaker voor het lage bedrag kozen (en dus minder risico namen) in taakblokken waarin een keuze voor het hoge geldbedrag juist vaker beloond werd. Deze resultaten laten zien dat theta/beta EEG ratio niet per se geassocieerd is met het nemen van risico, maar meer met het minder goed leren van positieve en negatieve feedback en het aanpassen van strategie. Theta/beta EEG ratio zoals gemeten in deze taak bleek dus geen geschikte maat voor beloningsgevoeligheid.

De corticale correlaten van de anticipatie van waarde en waarschijnlijkheid van beloning

In Hoofdstuk 5 en 6 is gekeken wat de corticale correlaten zijn van de anticipatie van waarde en waarschijnlijkheid van beloning. De resultaten laten zien dat deze respectievelijke processen dissocieerbaar zijn, zowel in termen van tijd (Hoofdstuk 5) als in termen van betrokken neurotransmitters (Hoofdstuk 6).

In deze hoofdstukken is gebruik gemaakt van een taak waarbij verschillende letters werden gepresenteerd op een computerscherm. De bedoeling was dat proefpersonen op de juiste knop drukten wanneer letter X of Y na de letter A verscheen. Voorafgaand aan ieder taakblok werd aangegeven of een proefpersoon geld kon verdienen voor een correcte knopdruk of niet en wat de kans was dat letter X of Y zou volgen na letter A. De waarde van de beloning (0 Euro versus een totaal van maximaal 5 Euro per taakblok) en de waarschijnlijkheid van letter X/Y (50% versus 98%) werden op een volledige gekruiste manier gemanipuleerd tijdens vier taakblokken. In dit onderzoek werd specifiek gekeken naar de corticale hersenactiviteit die verband houdt met de anticipatie van respectievelijk de waarde van de beloning en waarschijnlijkheid, zoals die optreedt na letter A en voorafgaat aan de (mogelijke) presentatie van letter X/Y.

Deze hoofdstukken laten zien dat de manipulatie van waarde effect had op de amplitude van een frontale positieve potentiaal rond 250 ms na presentatie van letter A (de belonings-gerelateerde positiviteit, oftewel *reward-related positivity*, RRP). De amplitude van deze ERP was groter tijdens de blokken waarbij geld kon worden verdiend in vergelijking met blokken waarbij geen geld kon worden verdiend. Een zelfde effect werd gevonden voor de P300, een parietale positieve potentiaal die na c.a. 400 ms na letter A optreedt. Op basis van eerder onderzoek hadden we verwacht dat de manipulatie van waarschijnlijkheid van het verschijnen van letter X/Y ook effect zou hebben op de P300. Deze manipulatie bleek echter effect te hebben op een andere ERP, die we de waarschijnlijkheids-gerelateerde positiviteit (*probability-related positivity*; PRP) hebben genoemd. Deze PRP had een frontaal (in plaats van parietaal) maximum. De amplitude van de PRP was groter tijdens de blokken met een grote kans op het verschijnen van letter X/Y in vergelijking met blokken met een lage kans hierop.

Neurotransmitters betrokken bij de anticipatie van waarde en waarschijnlijkheid

Zijn de corticale processen gerelateerd aan de anticipatie van waarde en waarschijnlijkheid dissocieerbaar in termen van onderliggende neurotransmitters? In overeenstemming met de verwachting lieten de resultaten van Hoofdstuk 6 zien dat de anticipatie van waarde van een beloning gerelateerd was aan neurotransmissie in het dopamine-systeem en dat de anticipatie van waarschijnlijkheid gerelateerd was aan neurotransmissie binnen het noradrenaline-systeem. Specifiek vonden we dat blokkade van dopaminerge neurotransmissie door middel van de stof haloperidol de amplitude van de belonings-gerelateerde positiviteit significant kleiner maakte. Een zelfde effect werd gevonden voor de amplitude van de P300. Het effect van haloperidol op de P300 was echter alleen significant in een groep proefpersonen die vanzichzelf hoge concentraties van dopamine hebben. In deze groep was er kennelijk meer ruimte voor het dempende effect van haloperidol op de P300.

Blokkade van noradrenerge neurotransmissie door middel van de stof clonidine had een specifiek en significant effect op de waarschijnlijkheids-gerelateerde positiviteit (PRP). In Hoofdstuk 5 was de PRP prominent aanwezig, maar in Hoofdstuk 6 was de PRP afwezig in de placebo en haloperidol conditie. In de clonidine conditie had de PRP een omgekeerde polariteit (de amplitude was groter voor de taakblokken met een lage kans op letter X/Y dan voor de blokken met een hoge kans). Het zou kunnen dat de PRP in het algemeen gedempt was tijdens de studie beschreven in Hoofdstuk 6 als gevolg van de meerdere sessies en dat clonidine de amplitude nog verder naar beneden bracht tot in een negativiteit.

De gedragsresultaten van Hoofdstuk 5 en 6 laten zien dat zowel de anticipatie van een geldbeloning als de anticipatie van een hoge kans op letter X/Y leidden tot snellere reactietijden op een additieve manier. Deze gedragseffecten laten zien dat er wel een verschil moet zijn geweest in de neurale processen die optreden na presentatie van letter A in de verschillende beloning- en kans condities. In strijd met de verwachting vonden we in Hoofdstuk 6 geen effect van haloperidol en clonidine op deze beloning- en kans-gerelateerde gedragseffecten. Wel vonden we algemene

effecten van de medicatie op het aantal omissies en de variabiliteit van reactiesnelheid. Dit houdt waarschijnlijk verband met het effect dat de medicatie heeft op (het wegvallen van) aandacht.

Slotopmerkingen

De studies in deze thesis geven onder andere meer inzicht in hoe het brein op een potentiele beloning anticipeert. De studies laten zien dat meerdere dissocieerbare processen een rol spelen tijdens de anticipatie van een beloning. Dit kan bijdragen aan een beter begrip in de processen die mogelijk verstoord zijn bij psychiatrische patienten met problemen op het gebied van motivatie. Ook kan het interessant zijn om te kijken of patienten met een specifiek probleem met het anticiperen van de waarschijnlijkheid van een beloning (meer) baat hebben bij een medicijn dat gericht is op noradrenerge neurotransmissie. Patienten met problemen met het anticiperen van waarde van een toekomstige beloning zouden daarentegen (meer) baat kunnen hebben bij een medicijn dat gericht is op dopaminerge neurotransmissie. Denk hierbij aan bijvoorbeeld major depressive disorder, waarbij anhedonie (pathologisch gebrek aan genotsbeleving) een belangrijke rol speelt, en waarbij met wisselend succes medicatie wordt toegepast, soms gericht op het noradrenaline systeem, maar soms ook op het dopamine-systeem.

Een ander punt voor mogelijk vervolgonderzoek is wat precies de relatie is tussen corticale hersenactiviteit gerelateerd aan waarde en waarschijnlijkheid, en het op verkrijgen van de beloning gericht gedrag. Zijn de in dit proefschrift beschreven corticale processen afspiegelingen van de representatie of de berekening van waarde en waarschijnlijkheid, of eerder het gevolg daarvan, bijvoorbeeld het richten van aandacht op een mogelijke toekomstige beloning (zie ook de opmerking over cognitieve controle aan het begin van deze samenvatting)?

References

A

Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L. S., . . . Van Heertum, R. L. (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proceedings of the National Academy of Sciences, 97*(14), 8104-8109.

Angelidis, A., Hagenaars, M., van Son, D., van der Does, W., & Putman, P. (2018). Do not look away! spontaneous frontal EEG theta/beta ratio as a marker for cognitive control over attention to mild and high threat. *Biological Psychology, 135*, 8-17.

Angelidis, A., van der Does, W., Schakel, L., & Putman, P. (2016). Frontal EEG theta/beta ratio as an electrophysiological marker for attentional control and its test-retest reliability. *Biological Psychology, 121*, 49-52.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu.Rev.Neurosci., 28*, 403-450.

B

Baas, J., Kenemans, J. L., Böcker, K., & Verbaten, M. N. (2002). Threat-induced cortical processing and startle potentiation. *Neuroreport, 13*(1), 133-137.

Baas, J., Mol, N., Kenemans, J. L., Prinssen, E. P., Niklson, I., Xia-Chen, C., . . . Van Gerven, J. (2009). Validating a human model for anxiety using startle potentiated by cue and context: The effects of alprazolam, pregabalin, and diphenhydramine. *Psychopharmacology, 205*(1), 73-84.

Baker, T. E., & Holroyd, C. B. (2009). Which Way Do I Go? Neural Activation in Response to Feedback and Spatial Processing in a Virtual T-Maze. *Cerebral Cortex, 19*(8), 1708-1722.

Barch, D. M., Treadway, M. T., & Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: Reduced effort allocation predicts amotivation and functional impairment. *Journal of Abnormal Psychology, 123*(2), 387.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7-15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A.R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Sciences*, *9*, 159-162.

Bekker, E. M., Kenemans, J. L., & Verbaten, M. N. (2004). Electrophysiological correlates of attention, inhibition, sensitivity and bias in a continuous performance task. *Clinical Neurophysiology, 115*(9), 2001-2013.

Benning, S. D., Patrick, C. J., & Lang, A. R. (2004). Emotional modulation of the post-auricular reflex. *Psychophysiology*, *41*(3), 426-432.

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus–noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews, 42*(1), 33-84.

Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology, 191*(3), 391-431.

Blin, O., Masson, G., Azulay, J. P., Fondarai, J., & Serratrice, G. (1990). Apomorphine-induced blinking and yawning in healthy volunteers. *British Journal of Clinical Pharmacology, 30*(5), 769-773.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42*(1), 1-15.

Bodi, N., Keri, S., Nagy, H., Moustafa, A., Myers, C. E., Daw, N., . . . Gluck, M. A. (2009). Reward-learning and the novelty-seeking personality: A between- and within-subjects study of the effects of dopamine agonists on young parkinson's patients. *Brain : A Journal of Neurology, 132*(Pt 9), 2385-2395. doi:10.1093/brain/awp094 [doi]

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1999). Affect and the startle reflex. *Startle modification: Implications for neuroscience, cognitive science, and clinical science*, 157-183.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*(2), 619-639.

Bress, J. N., & Hajcak, G. (2013). Self-report and behavioral measures of reward sensitivity predict the feedback negativity. *Psychophysiology, 50*(7), 610-616.

Brown, S. B., Van der Wee, N. J., Van Noorden, M. S., Giltay, E. J., & Nieuwenhuis, S. (2015). Noradrenergic and cholinergic modulation of late ERP responses to deviant stimuli. *Psychophysiology, 52*(12), 1620-1631.

Broyd, S. J., Richards, H. J., Helps, S. K., Chronaki, G., Bamford, S., & Sonuga-Barke, E. J. (2012). An electrophysiological monetary incentive delay (e-MID) task: A way to decompose the different components of neural response to positive and negative monetary reinforcement. *Journal of Neuroscience Methods, 209*(1), 40-49.

C

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology*, *67*(2), 319.

Cavanagh, J. F., Bismark, A., Frank, M. J., & Allen, J. J. (2011). Larger error signals in major depression are associated with better avoidance learning. *Frontiers in Psychology, 2*, 331.

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2011). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, bhr332.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, *18*(8), 414-421.

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

Cavanagh, J. F., Masters, S. E., Bath, K., & Frank, M. J. (2014). Conflict acts as an implicit cost in reinforcement learning. *Nature Communications, 5*, 5394. doi:10.1038/ncomms6394 [doi]

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology-Paris*, *109*(1), 3-15.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology.General, 143*(4), 1476-1488. doi:10.1037/a0035813 [doi]

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220-238.

Chong, T. T., Bonnelle, V., Manohar, S., Veromann, K., Muhammed, K., Tofaris, G. K., . . . Husain, M. (2015). Dopamine enhances willingness to exert effort for reward in parkinson's disease. *Cortex, 69*, 40-46.

Clark L, Cools R, & Robbins, T.W. (2004). The neuropsychology of ventral prefrontal cortex: decision-making and reversal learning. *Brain and Cognition, 55*, 41–53.

Clark, C. R., Geffen, G. M., & Geffen, L. B. (1989). Catecholamines and the covert orientation of attention in humans. *Neuropsychologia, 27*(2), 131-139.

Cockburn, J., Collins, A. G., & Frank, M. J. (2014). A reinforcement learning mechanism responsible for the valuation of free choice. *Neuron, 83*(3), 551-557. doi:10.1016/j.neuron.2014.06.035 [doi]

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in neurosciences*, *37*(9), 480-490.

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E., & Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain research*, *1238*, 127-142.

Collins, A. G., & Frank, M. J. (2016). Surprise! dopamine signals mix action, value and error. *Nature Neuroscience, 19*(1), 3.

Cools, R. (2008). Role of dopamine in the motivational and cognitive control of behavior. *The Neuroscientist, 14*(4), 381-395.

Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in parkinson's disease depends on medication status and outcome valence. *Neuropsychologia, 44*(10), 1663-1673. doi:S0028-3932(06)00107-2 [pii]

Cools, R., & D'Esposito, M. (2011). Inverted-u–shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry, 69*(12), e125.

Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., & D'Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *Journal of Neuroscience, 29*(5), 1538-1543.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201.

Corsi-Cabrera, M., Galindo-Vilchis, L., del-Río-Portilla, Y., Arce, C., & Ramos-Loyo, J. (2007). Within-subject reliability and inter-session stability of EEG power and coherent activity in women evaluated monthly over nine months. *Clinical Neurophysiology, 118*(1), 9-21.

Coull, J. T., Nobre, A. C., & Frith, C. D. (2001). The noradrenergic α2 agonist clonidine modulates behavioural and neuroanatomical correlates of human attentional orienting and alerting. *Cerebral Cortex, 11*(1), 73-84.

Cox, S. M., Frank, M. J., Larcher, K., Fellows, L. K., Clark, C. A., Leyton, M., & Dagher, A. (2015). Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. *NeuroImage, 109*, 95-101.

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating amazon's mechanical turk as a tool for experimental behavioral research. *PloS One, 8*(3), e57410.

Cuthbert, B. N., Bradley, M. M., & Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology, 33*(2), 103-111.

D

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876-879.

De Bruijn, E. R., Hulstijn, W., Verkes, R. J., Ruigt, G. S., & Sabbe, B. G. (2005). Altered response evaluation: Monitoring of late responses after administration of D-amphetamine. *Journal of Psychophysiology, 19*(4), 311-318.

De Bruijn, E. R., Sabbe, B. G., Hulstijn, W., Ruigt, G. S., & Verkes, R. J. (2006). Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers. *Brain Research, 1105*(1), 122-129.

de Rover, M., Brown, S. B., Band, G. P., Giltay, E. J., van Noorden, M. S., van der Wee, Nic JA, & Nieuwenhuis, S. (2015). Beta receptor-mediated modulation of the oddball P3 but not error-related ERP components in humans. *Psychopharmacology, 232*(17), 3161-3172.

Dichter, G. S., Tomarken, A. J., & Baucom, B. R. (2002). Startle modulation before, during and after exposure to emotional stimuli. *International Journal of Psychophysiology, 43*(2), 191-196.

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of neuroscience methods*, *187*(1), 138-145.

Dien, J. (2012). Applying principal components analysis to event-related potentials: a tutorial. *Developmental neuropsychology*, *37*(6), 497-517.

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 31*(16), 6188-6198. doi:10.1523/JNEUROSCI.6486-10.2011 [doi]

Doñamayor, N., Schoenfeld, M. A., & Münte, T. F. (2012). Magneto-and electroencephalographic manifestations of reward anticipation and delivery. *NeuroImage, 62*(1), 17-29.

Donner, T. H., & Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends in Cognitive Sciences, 15*(5), 191-199.

Dreisbach, G., Müller, J., Goschke, T., Strobel, A., Schulze, K., Lesch, K., & Brocke, B. (2005). Dopamine and cognitive control: The influence of spontaneous eyeblink rate and dopamine gene polymorphisms on perseveration and distractibility. *Behavioral Neuroscience, 119*(2), 483.

Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology, 14*(1-2), 1-52.

E

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and emotion*, *30*(2), 111-116.

El Mansari, M., Guiard, B. P., Chernoloz, O., Ghanbari, R., Katz, N., & Blier, P. (2010). Relevance of norepinephrine–dopamine interactions in the treatment of major depressive disorder. *CNS Neuroscience & Therapeutics, 16*(3)

Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology, 20*(2), 156-165.

Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine, 36*(03), 299-312.

F

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*(3), 298-308.

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain*, *126*(8), 1830-1837.

Fellows, L. K. (2007). The role of orbitofrontal cortex in decision making. *Annals of the New York Academy of Sciences*, *1121*(1), 421-430.

Flores, A., Münte, T. F., & Doñamayor, N. (2015). Event-related EEG responses to anticipation and delivery of monetary and social reward. *Biological Psychology, 109*, 10-19.

Floresco, S. B. (2015). Noradrenaline and dopamine: Sharing the workload. *Trends in Neurosciences, 38*(8), 465-467.

Forster, S. E., Zirnheld, P., Shekhar, A., Steinhauer, S. R., O'Donnell, B. F., & Hetrick, W. P. (2017). Event-related potentials reflect impaired temporal interval learning following haloperidol administration. *Psychopharmacology, 234*(17), 2545-2562.

Foti, D., Carlson, J. M., Sauder, C. L., & Proudfit, G. H. (2014). Reward dysfunction in major depression: Multimodal neuroimaging evidence for refining the melancholic phenotype. *NeuroImage, 101*, 50-58.

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

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

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences, 104*(41), 16311-16316.

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science (New York, N.Y.), 318*(5854), 1309-1312. doi:1146157 [pii]

Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.), 306*(5703), 1940-1943. doi:1102941 [pii]

Frank, M. J., & O'reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience, 120*(3), 497.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501.

Franken, I. H., Muris, P., & Rassin, E. (2005). Psychometric properties of the Dutch BIS/BAS scales. *Journal of Psychopathology and Behavioral Assessment*, *27*(1), 25-30.

Friedman, J. I., Carpenter, D., Lu, J., Fan, J., Tang, C. Y., White, L., . . . Flanagan, L. (2008). A pilot study of adjunctive atomoxetine treatment to second-generation antipsychotics for cognitive impairment in schizophrenia. *Journal of Clinical Psychopharmacology, 28*(1), 59-63.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences, 9*(10), 474-480.

Frodl-Bauch, T., Bottlender, R., & Hegerl, U. (1999). Neurochemical substrates and neuroanatomical generators of the event-related P300. *Neuropsychobiology, 40*(2), 86-94.

G

Gable, P. A., & Harmon-Jones, E. (2009). Postauricular reflex responses to pictures varying in valence and arousal. *Psychophysiology*, *46*(3), 487-490.

Gard, D. E., Gard, M. G., Mehta, N., Kring, A. M., & Patrick, C. J. (2007). Impact of motivational salience on affect modulated startle at early and late probe times. *International Journal of Psychophysiology, 66*(3), 266-270.

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

Gillan, C. M., & Daw, N. D. (2016). Taking psychiatry research online. *Neuron, 91*(1), 19-23.

Glimcher, P. W. (2004). *Decisions, uncertainty, and the brain: The science of neuroeconomics* MIT press.

Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science*, *306*(5695), 447-452.

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: A deficit in the representation of value. *Schizophrenia Bulletin, 34*(5), 835-847.

Goldstein, R. Z., Cottone, L. A., Jia, Z., Maloney, T., Volkow, N. D., & Squires, N. K. (2006). The effect of graded monetary reward on cognitive event-related potentials and behavior in young healthy adults. *International Journal of Psychophysiology*, *62*(2), 272-279.

Grace, A. A., Floresco, S. B., Goto, Y., & Lodge, D. J. (2007). Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends in Neurosciences, 30*(5), 220-227.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, *55*(4), 468-484.

Gray, J.A. (1987). The neuropsychology of anxiety: An enquiry into the septo-hippocampal system. *Oxford University Press*, Oxford.

Grillon, C. (2002). Startle reactivity and anxiety disorders: Aversive conditioning, context, and neurobiology. *Biological Psychiatry, 52*(10), 958-975. doi:S0006322302016657 [pii]

Grillon, C., & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology, 114*(9), 1557-1579.

Gründler, T. O., Cavanagh, J. F., Figueroa, C. M., Frank, M. J., & Allen, J. J. (2009). Task-related dissociation in ERN amplitude as a function of obsessive–compulsive symptoms. *Neuropsychologia, 47*(8), 1978-1987.

H

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4.

Hackley, S. A., Muñoz, M. Á, Hebert, K., Valle-Inclán, F., & Vila, J. (2009). Reciprocal modulation of eye-blink and pinna-flexion components of startle during reward anticipation. *Psychophysiology, 46*(6), 1154-1159.

Hahn, T., Dresler, T., Ehlis, A., Plichta, M. M., Heinzel, S., Polak, T., . . . Fallgatter, A. J. (2009). Neural response to reward anticipation is modulated by gray's impulsivity. *NeuroImage, 46*(4), 1148-1153.

Hall, H., Sedvall, G., Magnusson, O., Kopp, J., Halldin, C., & Farde, L. (1994). Distribution of D 1-and D 2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacology, 11*(4), 245.

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., . . . Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nature Neuroscience, 19*(1), 117.

Hawk, L. W., & Kowmas, A. D. (2003). Affective modulation and prepulse inhibition of startle among undergraduates high and low in behavioral inhibition and approach. *Psychophysiology*, *40*(1), 131-138.

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

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport, 22*(5), 249-252.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688-697.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences, 16*(2), 122-128.

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences, 112*(16), 5195-5200.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von restorff really did. *Psychonomic Bulletin & Review, 2*(1), 105-112.

I

Imai, S., & Garner, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology, 69*(6), 596.

J

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

Jongkees, B. J., & Colzato, L. S. (2016). Spontaneous eye blink rate as predictor of dopamine-related cognitive function—A review. *Neuroscience & Biobehavioral Reviews, 71*, 58-82.

Joseph, K. C., & Sitaram, N. (1989). The effect of clonidine on auditory P300. *Psychiatry Research, 28*(3), 255-262.

K

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience, 10*(12), 1625.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263-291.

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127) World Scientific.

Kapur, S., Zipursky, R., Jones, C., Remington, G., & Houle, S. (2000). Relationship between dopamine D2 occupancy, clinical response, and side effects: A double-blind PET study of first-episode schizophrenia. *American Journal of Psychiatry, 157*(4), 514-520.

Karson, C. N. (1983). Spontaneous eye-blink rates and dopaminergic systems. *Brain, 106*(3), 643-653.

Kenemans, J. L. (1991). DAR: data reduction with contrast vectors. *Computers in psychology: applications in education, research, and psychodiagnostics. Amsterdam: Swets & Zeitlinger*, 140-148.

Kenemans, J. L., Lijffijt, M., Camfferman, G., & Verbaten, M. N. (2002). Split-second sequential selective activation in human secondary visual cortex. *Journal of Cognitive Neuroscience*, *14*(1), 48-61.

Kenemans, J. L., Wieleman, J. S., Zeegers, M., & Verbaten, M. N. (1999). Caffeine and stroop interference. *Pharmacology Biochemistry and Behavior, 63*(4), 589-598.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? neural substrates of avoidance learning in the human brain. *PLoS Biology, 4*(8), e233.

Kirsch, P., Schienle, A., Stark, R., Sammer, G., Blecker, C., Walter, B., . . . Vaitl, D. (2003). Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system:: An event-related fMRI study. *NeuroImage, 20*(2), 1086-1095.

Kiss, M., Goolsby, B. A., Raymond, J. E., Shapiro, K. L., Silvert, L., Nobre, A. C., . . . Eimer, M. (2007). Efficient attentional selection predicts distractor devaluation: Event-related potential evidence for a direct link between attention and emotion. *Journal of Cognitive Neuroscience, 19*(8), 1316-1322.

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science (New York, N.Y.), 318*(5856), 1642-1645. doi:318/5856/1642 [pii]

Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., & Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the P1–N1 complex and are related to memory performance. *Cognitive Brain Research*, *19*(3), 302-316.

Klumpers, F., Heitland, I., Oosting, R. S., Kenemans, J. L., & Baas, J. M. (2012). Genetic variation in serotonin transporter function affects human fear expression indexed by fear-potentiated startle. *Biological psychology*, *89*(2), 277-282.

Klumpers, F., van Gerven, J. M., Prinssen, E. P. M., Niklson, I., Roesch, F., Riedel, W. J., ... & Baas, J. M. P. (2009). Method development studies for repeatedly measuring anxiolytic drug effects in healthy humans. *Journal of Psychopharmacology*. *24*(5), 657-666.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci*, *21*(16), RC159.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *Journal of Neuroscience, 25*(19), 4806-4812.

Knyazev, G. G. (2007). Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews, 31*(3), 377-395.

Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being moved valence activates approach-avoidance behavior independently of evaluation and approach-avoidance intentions. *Psychological Science*, *21*(4), 607-613.

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology, 132*, 175-183.

Krigolson, O. E., Hassall, C. D., & Handy, T. C. (2014). How we learn to make decisions: rapid propagation of reinforcement learning prediction errors in humans. *Journal of cognitive neuroscience*, *26*(3), 635-644.

Krigolson, O. E., & Holroyd, C. B. (2007). Predictive information and error processing: The role of medial-frontal cortex during motor control. *Psychophysiology, 44*(4), 586-595.

Kujawa, A., Proudfit, G. H., & Klein, D. N. (2014). Neural reactivity to rewards and losses in offspring of mothers and fathers with histories of depressive and anxiety disorders. *Journal of Abnormal Psychology, 123*(2), 287.

L

Lang, P. J., & Bradley, M. M. (2013). Appetitive and defensive motivation: goal-directed or goal-determined?. *Emotion Review*, *5*(3), 230-234.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377.

Lipp, O. V., Cox, D., & Siddle, D. A. (2001). Blink startle modulation during anticipation of pleasant and unpleasant stimuli. *Journal of Psychophysiology, 15*(3), 155.

Logemann, H. A., Böcker, K. B., Deschamps, P. K., van Harten, P. N., Koning, J., Kemner, C., . . . Kenemans, J. L. (2017). Haloperidol 2 mg impairs inhibition but not visuospatial attention. *Psychopharmacology, 234*(2), 235-244.

Logemann, H. A., Böcker, K. B., Deschamps, P. K., Kemner, C., & Kenemans, J. L. (2013). The effect of noradrenergic attenuation by clonidine on inhibition in the stop signal task. *Pharmacology Biochemistry and Behavior, 110*, 104-111.

Lou, B., Hsu, W. Y., & Sajda, P. (2015). Perceptual salience and reward both influence feedback-related neural activity arising from choice. *Journal of Neuroscience*, *35*(38), 13064-13075.

Low, A., Lang, P. J., Smith, J. C., & Bradley, M. M. (2008). Both predator and prey: Emotional arousal in threat and reward. *Psychological Science, 19*(9), 865-873. doi:10.1111/j.1467-9280.2008.02170.x [doi]

Lubow, R. E. (1973). Latent inhibition. *Psychological bulletin*, *79*(6), 398.

Luck, S. J. (2012). Event-related potentials. *APA Handbook of Research Methods in Psychology, 1*, 523-546.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146-157.

Luft, C. D. B., Takase, E., & Bhattacharya, J. (2014). Processing graded feedback: electrophysiological correlates of learning from small and large errors. *Journal of cognitive neuroscience*, *26*(5), 1180-1193.

M

Maltzman, I. (1979). Orienting reflexes and classical conditioning in humans. *The Orienting Reflex in Humans,* , 323-351.

Marco-Pallarés, J., Münte, T. F., & Rodríguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neuroscience & Biobehavioral Reviews*, *49*, 1-7.

Maril, S., Hassin-Baer, S., Cohen, O. S., & Tomer, R. (2013). Effects of asymmetric dopamine depletion on sensitivity to rewarding and aversive stimuli in parkinson's disease. *Neuropsychologia, 51*(5), 818-824. doi:10.1016/j.neuropsychologia.2013.02.003 [doi]

Martins, D., Mehta, M. A., & Prata, D. (2017). The "highs and lows" of the human brain on dopaminergics: Evidence from neuropharmacology. *Neuroscience & Biobehavioral Reviews, 80*, 351-371.

Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *Journal of cognitive neuroscience*, *26*(3), 447-458.

Maslow, A., & Lewis, K. J. (1987). Maslow's hierarchy of needs. *Salenger Incorporated, 14*, 987.

Massar, S. A., Kenemans, J. L., & Schutter, D. J. (2014). Resting-state EEG theta activity and risk learning: sensitivity to reward or punishment?.*International Journal of Psychophysiology*, *91*(3), 172-177.

Massar, S. A. A., Rossi, V., Schutter, D. J. L. G., & Kenemans, J. L. (2012). Baseline EEG theta/beta ratio and punishment sensitivity as biomarkers for feedback-related negativity (FRN) and risk-taking. *Clinical Neurophysiology*, *123*(10), 1958-1965.

N

Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proceedings of the National Academy of Sciences of the United States of America, 107*(11), 5232-5237. doi:10.1073/pnas.0911972107 [doi]

Neuhaus, A., Bajbouj, M., Kienast, T., Kalus, P., Von Haebler, D., Winterer, G., & Gallinat, J. (2006). Persistent dysfunctional frontal lobe activation in former smokers. *Psychopharmacology, 186*(2), 191-200.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin, 131*(4), 510.

Nieuwenhuis, S., Heslenfeld, D. J., Alting von Geusau, Niels J, Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage, 25*(4), 1302-1309.

Nitschke, J. B., Larson, C. L., Smoller, M. J., Navin, S. D., Pederson, A. J., Ruffalo, D., . . . Davidson, R. J. (2002). Startle potentiation in aversive anticipation: Evidence for state but not trait effects. *Psychophysiology, 39*(2), 254-258.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology, 191*(3), 507-520.

O

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology*, *14*(6), 769-776.

P

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience, 23*(11), 3419-3432.

Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., & Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in parkinson's and tourette's syndromes. *Proceedings of the National Academy of Sciences of the United States of America, 106*(45), 19179-19184. doi:10.1073/pnas.0904035106 [doi]

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol*, *24*(Suppl D), 5-12.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042.

Pessoa, L., & Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Frontiers in neuroscience*, *4*, 17.

Pfabigan, D. M., Seidel, E., Sladky, R., Hahn, A., Paul, K., Grahl, A., . . . Kranz, G. S. (2014). P300 amplitude variation is related to ventral striatum BOLD response during gain and loss anticipation: An EEG and fMRI experiment. *NeuroImage, 96*, 12-21.

Pineda, J. A., Foote, S. L., & Neville, H. J. (1989). Effects of locus coeruleus lesions on auditory, long-latency, event-related potentials in monkey. *Journal of Neuroscience, 9*(1), 81-93.

Piray, P., Zeighami, Y., Bahrami, F., Eissa, A. M., Hewedi, D. H., & Moustafa, A. A. (2014). Impulse control disorders in parkinson's disease are associated with dysfunction in stimulus valuation but not action valuation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 34*(23), 7814-7824. doi:10.1523/JNEUROSCI.4063-13.2014 [doi]

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biological Psychiatry, 57*(4), 319-327.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*(6741), 233.

Pleger, B., Ruff, C. C., Blankenburg, F., Klöppel, S., Driver, J., & Dolan, R. J. (2009). Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biology, 7*(7), e1000164.

Polanía, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron, 82*(3), 709-720.

Price, T. F., Dieckman, L. W., & Harmon-Jones, E. (2012). Embodying approach motivation: Body posture influences startle eyeblink and event-related potential responses to appetitive stimuli. *Biological psychology*, *90*(3), 211-217.

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449-459.

R

Rothschild, M. L., & Gaidis, W. C. (1981). Behavioral learning theory: Its relevance to marketing and promotions. *Journal of Marketing, 45*(2), 70-78.

Rushworth, M. F., & Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature neuroscience*, *11*(4), 389-397.

S

Sabatinelli, D., Bradley, M. M., & Lang, P. J. (2001). Affective startle modulation in anticipation and perception. *Psychophysiology, 38*(04), 719-722.

Salamone, J. D., & Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. *Neuron, 76*(3), 470-485.

Sandt, A. R., Sloan, D. M., & Johnson, K. J. (2009). Measuring appetitive responding with the postauricular reflex. *Psychophysiology*, *46*(3), 491-497.

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping, 30*(7), 1963-1976.

Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., & Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International Journal of Psychophysiology, 67*(3), 242-251.

Schmajuk, N. A., & DiCarlo, J. J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review, 99*(2), 268.

Schonberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 27*(47), 12860-12867. doi:27/47/12860 [pii]

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95*(3), 853-951.

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

Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion, 18*(5), 593-611.

Schutter, D. J., de Weijer, A. D., Meuwese, J. D., Morgan, B., & van Honk, J. (2008). Interrelations between motivational stance, cortical excitability, and the frontal electroencephalogram asymmetry of emotion: A transcranial magnetic stimulation study. *Human Brain Mapping, 29*(5), 574-580.

Schutter, D. J., Leitner, C., Kenemans, J. L., & van Honk, J. (2006). Electrophysiological correlates of cortico-subcortical interaction: A cross-frequency spectral EEG analysis. *Clinical Neurophysiology*, *117*(2), 381-387.

Schutter, D. J., & Van Honk, J. (2005). Electrophysiological ratio markers for the balance between reward and punishment. *Cognitive Brain Research, 24*(3), 685-690.

Sescousse, G., Caldú, X., Segura, B., & Dreher, J. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 37*(4), 681-696.

Sege, C. T., Bradley, M. M., & Lang, P. J. (2014). Startle modulation during emotional anticipation and perception. *Psychophysiology*, *51*(10), 977-981.

Singh, S., Lewis, R. L., Barto, A. G., & Sorg, J. (2010). Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Development, 2*(2), 70-82.

Skolnick, A. J., & Davidson, R. J. (2002). Affective modulation of eyeblink startle with reward and threat. *Psychophysiology, 39*(6), 835-850.

Smith, B. W., Mitchell, D. G., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J. R., & Ernst, M. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *NeuroImage, 44*(2), 600-609.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*(3), 161-168.

Sonuga-Barke, E. J. (2005). Causal models of attention-deficit/hyperactivity disorder: From common simple deficits to multiple developmental pathways. *Biological Psychiatry, 57*(11), 1231-1238.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin*, *87*(2), 245-251.

Stevens, F. L., Hurley, R. A., & Taber, K. H. (2011). Anterior cingulate cortex: Unique role in cognition and emotion. *The Journal of Neuropsychiatry and Clinical Neurosciences, 23*(2), 121-125.

Studer, B., Pedroni, A., & Rieskamp, J. (2013). Predicting risk-taking behavior from prefrontal resting-state activity and personality. *PLoS one*, *8*(10), e76861.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Stai. *Manual for the State-Trait Anxiety Inventory (Self Evaluation Questionnaire).Palo Alto California: Consulting Psychologist, 22*, 1-24.

Svensson, T. H., Bunney, B. S., & Aghajanian, G. K. (1975). Inhibition of both noradrenergic and serotonergic neurons in brain by the α-adrenergic agonist clonidine. *Brain Research, 92*(2), 291-306.

T

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology,* , 212-222.

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology, 121*(3), 553.

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: Lessons from translational neuroscience. *Neuroscience & Biobehavioral Reviews, 35*(3), 537-555.

V

van der Lubbe, Rob HJ, Neggers, S. F., Verleger, R., & Kenemans, J. L. (2006). Spatiotemporal overlap between brain activation related to saccade preparation and attentional orienting. *Brain Research, 1072*(1), 133-152.

Van de Vijver, I., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of cognitive neuroscience*, *23*(12), 4106-4121.

van Driel, J., Sligte, I. G., Linders, J., Elport, D., & Cohen, M. X. (2015). Frequency band-specific electrical brain stimulation modulates cognitive control processes. *PloS one*, *10*(9), e0138984.

van Driel, J., Swart, J. C., Egner, T., Ridderinkhof, K. R., & Cohen, M. X. (2015). (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 787-807.

van Honk, J., Hermans, E. J., Putman, P., Montagne, B., & Schutter, D. J. (2002). Defective somatic markers in sub-clinical psychopathy. *Neuroreport,13*(8), 1025-1027.

van Son, D., Angelidis, A., Hagenaars, M. A., van der Does, W., & Putman, P. (2018). Early and late dot-probe attentional bias to mild and high threat pictures: Relations with EEG theta/beta ratio, self-reported trait attentional control, and trait anxiety. *Psychophysiology, 55*(12), e13274.

van Son, D., Schalbroeck, R., Angelidis, A., van der Wee, Nic JA, van der Does, W., & Putman, P. (2018). Acute effects of caffeine on threat-selective attention: Moderation by anxiety and EEG theta/beta ratio. *Biological Psychology, 136*, 100-110.

Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys. *Journal of Neuroscience, 35*(20), 7866-7877.

Vogt, B. A. (2016). Midcingulate cortex: Structure, connections, homologies, functions and diseases. *Journal of Chemical Neuroanatomy, 74*, 28-46.

Von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychologische Forschung, 18*(1), 299-342.

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The startle probe response: A new measure of emotion? *Journal of Abnormal Psychology, 97*(4), 487.

W

Wald, F. D., & Mellenbergh, G. J. (1990). De verkorte versie van de nederlandse vertaling van de profile of mood states (POMS). *Nederlands Tijdschrift Voor De Psychologie En Haar Grensgebieden.*

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, *36*(8), 1870-1884.

Waltz, J. A., Frank, M. J., Robinson, B. M., & Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biological Psychiatry, 62*(7), 756-764.

Wischnewski, M., & Schutter, D. J. (2019). Electrophysiological correlates of prediction formation in anticipation of reward-and punishment-related feedback signals. *Psychophysiology*, e13379

Wischnewski, M., Zerr, P., & Schutter, D. J. (2016). Effects of Theta Transcranial Alternating Current Stimulation Over the Frontal Cortex on Reversal Learning. *Brain stimulation*, *9*(5), 705-711.

Wright, J. S., & Panksepp, J. (2012). An evolutionary framework to understand foraging, wanting, and desire: The neuropsychology of the SEEKING system. *Neuropsychoanalysis, 14*(1), 5-39.

Y

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience, 26*(37), 9530-9537.

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*(5), 535-544.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience, 24*(28), 6258-6264.

Yu, R., & Zhou, X. (2006). Brain potentials associated with outcome expectation and outcome evaluation. *Neuroreport, 17*(15), 1649-1653.

Yu, R., Zhou, W., & Zhou, X. (2011). Rapid Processing of Both Reward Probability and Reward Uncertainty in the Human Anterior Cingulate Cortex. *PLoS ONE, 6*(12), e29633. doi: 10.1371/journal.pone .0029633

Z

Zirnheld, P. J., Carroll, C. A., Kieffaber, P. D., O'donnell, B. F., Shekhar, A., & Hetrick, W. P. (2004). Haloperidol impairs learning and error-related negativity in humans. *Journal of Cognitive Neuroscience, 16*(6), 1098-1112.

Acknowledgements/ Dankwoord

En dan is nu de tijd aangebroken dat ik gewoon het dankwoord van mijn proefschrift mag gaan schrijven! Het was een traject waarin ik heel veel heb geleerd en ik kijk er met een goed gevoel op terug. Hier wil ik graag alle mensen bedanken die (in welke vorm dan ook) een bijdrage hebben geleverd aan de totstandkoming van dit proefschrift.

Leon, bedankt voor je vertrouwen in mij als AIO voor het project en de vrijheid en prettige begeleiding die ik daarbij heb gekregen! Ik heb zoveel van jou geleerd: van schrijven tot statistische toetsen waar ik nog nooit van had gehoord (nog regelmatig ~~heb ik nachtmerries~~ denk ik met liefde terug aan DAR 😉). Jouw creatieve blik en gedrevenheid om alles tot in de puntjes te doorgronden zijn echt heel waardevol geweest en hebben de verschillende hoofdstukken naar een hoger niveau gebracht.

Joke en Dennis, veel dank voor jullie ondersteuning en kritische blik op mijn schrijfsels. Dennis, fijn dat je ondanks je verhuizing naar het Donders toch nog tijd hebt gevonden om betrokken te blijven bij mijn project. En heel leuk dat je terug komt op het oude nest. Joke, ook bedankt voor het meedenken over mijn vervolgcarrière. Ik ben heel blij met de nieuwe samenwerking/ aanstelling.

Ivo, kamergenoot van het eerste uur, altijd zo behulpzaam. Bedankt voor de gezelligheid, mega choco-konijnen en alles wat je me geleerd hebt op het gebied van startle en EEG analyses.

Peter Deschamps, ooit was jij mijn begeleider toen ik stage liep bij de afdeling kinder- en jeugdpsychiatrie in het UMC. Door die stage is mijn "onderzoeks-vlammetje" aangewakkerd en ben ik uiteindelijk via via AIO geworden. Jouw betrokkenheid als studiearts en co-auteur mbt de medicatiestudie is van groot belang geweest. Veel dank!

Peter van Harten wil ik hartelijk bedanken voor de goede samenwerking met betrekking tot de medicatiestudie.

Heleen Slagter, Michael Frank and Anne Collins, thank you for the great collaboration and for sharing your expertise.

Inge Volman, tijdens mijn onderzoeksstage bij jou (laaaang geleden) heb ik zo veel geleerd over statistiek en onderzoek doen. Dat heeft een goede basis gelegd voor mijn AIO-schap. Dankjewel daarvoor.

Alle collega's en voormalig collega's van Psychologische Functieleer wil ik bedanken voor de prettige sfeer op de afdeling en de koetjes en kalfjes bij de koffieautomaat. Een aantal wil ik hieronder nog speciaal noemen.

Alexander, je hebt me goed op weg geholpen met de METC en EEG zaken. METC aanvragen hadden zonder jouw hulp ongetwijfeld een stuk langer geduurd.

Chris J, fijn dat je altijd geinteresseerd bent in hoe het gaat. Ik ben blij met de nieuwe aanstelling en ik kijk er naar uit om nog een tijdje samen te werken aan de EEG/rijsimulator experimenten.

Peter B, dank dat je monitor wilde zijn voor de medicatiestudie.
Carlijn, bij jou kon ik aankloppen met EEG vragen. Dankjewel daarvoor.

In de loop van de jaren heb ik veel verschillende kamergenoten gehad. Bedankt voor de gezellige tijden allemaal! Jullie hebben mijn AIO-bestaan leuker gemaakt: Ivo, Richard, Jurre, Michiel (wat een tijd geleden alweer, benieuwd hoe het met jullie gaat), Estrella, Rinske en Febe (dank voor alle leuke klets- en klaag-momenten), Remo, Isabell, Jim, Joris, Jacob. En last but not least Nico, mijn huidige kamergenoot. Ik verbaas me altijd hoe goed jij met studenten bent. Altijd interessante gesprekken met jou over werk en niet-werk en over onze plant Bob.

Grote dank gaat ook uit naar alle collega's en voormalig collega's van het ondersteuningsteam van functieleer, in het bijzonder: Eveline, Ria, Renée, Ellen, Janet, en de mannen van de technische ondersteuning, in het bijzonder: Dennis H (voor de EEG vragen), Son en Jim (voor alle keren dat de BVA software weer eens niet meewerkte), Siarhei en Pieter (voor de kalibraties), en Roy (voor het maken van een script).

De collega's van de bio-emo-social groep wil ik bedanken voor hun goede feedback en ideeën tijdens de labmeetings. Daar heb ik regelmatig van kunnen profiteren.

All students that helped me with piloting, screening subjects, data collection, etcetera. Thanks for all your hard work and effort!: Alexandra Marques, Albena Vassileva, Inge de Kruijf and Carlijn Hintzen. Tijdens de afronding van dit proefschrift heb ik ook nog een aantal studenten mogen begeleiden tijdens hun stage of creative challenge. Bedankt voor jullie begrip dat ik soms wat minder beschikbaar was: Nina, Ilva, Valerie, Thomas, Esmee, Dennis Trommelen, Willem, Ignas, Nathaniël, de externe studenten, de studenten waarvan ik co-begeleider was en iedereen die ik nu vergeet in dit rijtje te zetten. I wish you all the best for your future career!

En dan nu de mensen uit mijn privé-leven:

De dames van de watervrienden/damesuitje en de biomedische meisjes (Annemieke, Michelle en Jantine). Leuk dat we nog altijd contact met elkaar houden! We moeten snel maar weer eens een uitje plannen.

Doortje, af en toe uit t oog, maar niet uit t hart.
Lieve Marleen, zijn we gewoon al bijna 30 jaar Bè en Boe. Wat hebben we veel meegemaakt. Op naar de volgende 30 jaar! Super leuk dat jij mijn paranimf bent.

Ooms, tantes, neven en nichten, leuk dat jullie mijn AIO-avonturen altijd met belangstelling hebben gevolgd.

Randy, mijn allerliefste hamsterfanboy. Dank je voor alle mooie jaren samen. Ik hoop dat jij mij nog heeeel lang plaagt.
Pa, ma, Linda, Erik, Denise en Bjorn, jullie zijn gewoon de allerbeste familie die ik kan wensen, punt. Zonder jullie was ik vast gillend gek geworden tijdens het schrijven en was dit alles niet gelukt. Linda, ik ben blij dat jij paranimf wilt zijn voor je kleine zusje.

Curriculum vitae

Iris Schutte was born on the 27th of June 1985 in Enschede, The Netherlands. In 2003 she graduated from Kottenpark College in Enschede (pre-university education; VWO). She decided to study nursing at Saxion University of Applied Sciences in Enschede, but switched to Biomedical Sciences at Utrecht University, because of her interest in research and science. In 2011 she graduated from the master program Neuroscience & Cognition. After graduation she started as a PhD candidate under supervision of Prof. dr. Leon Kenemans, Prof. dr. Johanna Baas, and Dr. Dennis Schutter at the department of Experimental Psychology at Utrecht University. Iris is currently employed as postdoctoral researcher by Dr. Chris Janssen, Prof. Johanna Baas, and Prof. Leon Kenemans at Utrecht University.

List of publications

Part of this thesis

Published/in press

Schutte, I., Slagter, H. A., Collins, A. G., Frank, M. J., & Kenemans, J. L. (2017). Stimulus discriminability may bias value-based probabilistic learning. *PloS one*, *12*(5), e0176205. https://doi.org/10.1371/journal.pone.0176205

Schutte, I., Kenemans, J. L., & Schutter, D. J. L. G. (2017). Resting-state theta/beta EEG ratio is associated with reward-and punishment-related reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 754-763. https://doi.org/10.3758/s13415-017-0510-3

Schutte, I., Heitland, I., & Kenemans, J. L. (2019). Disentangling the effects of reward value and probability on anticipatory event-related potentials. *Neuropsychologia*, 107138. https://doi.org/10.1016/j.neuropsychologia.2019.107138

Submitted for publication

Schutte, I., Baas, J.M.P., Heitland, I., Kenemans, J.L. (submitted). Startle-reflex modulation due to anticipation, withholding and obtainment of reward – a non-replication.

Schutte, I., Deschamps, P.K.H., van Harten, P.N., Kenemans, J.L. (submitted). Dopaminergic and noradrenergic manipulation of anticipatory reward and probability event-related potentials.

Not part of this thesis

Deschamps, P. K. H., **Schutte, I.**, Kenemans, J. L., Matthys, W. C. H. J., & Schutter, D. J. L. G. (2012). Electromyographic responses to emotional facial expressions in 6–7 year olds: A feasibility study. *International Journal of Psychophysiology*, *85*(2), 195-199. https://doi.org/10.1016/j.ijpsycho.2012.05.004

van de Groep, I. H., de Haas, L. M., **Schutte, I.**, & Bijleveld, E. (2017). Spontaneous eye blink rate (EBR) predicts poor performance in high-stakes situations. *International Journal of Psychophysiology*, *119*, 50-57. https://doi.org/10.1016/j.ijpsycho.2017.01.009


Helmholtz Instituut


Utrecht University