Influence of positive emotion on probabilistic learning: An experimental approach

Jasmina Bakić

Promotor: Prof. dr. Gilles Pourtois

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologie

2015

Influence of positive emotion on probabilistic learning: An experimental approach

Jasmina Bakić

Promotor: Prof. dr. Gilles Pourtois

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologie

2015

Guidance Committee:

Prof. dr. Gilles Pourtois (promotor)

Department of Experimental-Clinical and Health Psychology, Ghent University

Prof. dr. Rudi De Raedt

Department of Experimental-Clinical and Health Psychology, Ghent University

Prof. Dr. Nico Böhler

Department of Experimental Psychology, Ghent University

Prof. Dr. Sven Müller

Department of Experimental-Clinical and Health Psychology, Ghent University

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Gratefulness is one of the best feelings us humans can harvest. And today, I have so much to be grateful for.

First and foremost, to the force of life that goes by the name of Gilles. For the diligence, relentless enthusiasm, patience, impeccable work ethics, intellectual honesty, courage, where I lacked it. The trouble is, the inner voice in my head assigned to my cognitive control function will sound itself off with a very distinct wallonian accent for quite a while, I am afraid. Thank you for all the generous lessons of the past four years. If I managed to return at least a ten percent of what you have given me, then I leave this period happy.

Chris, it was truly a pleasure to work by your side. Thank you for providing me the opportunity to do so, and showing me along the way how to masterly organize a large and complex study in a sound way. Prof. Dr. Tom Loeys, thank you for all the time invested, for precious statistical advice and ideas that motivated a lot of my early steps as a researcher. I would also like to thank Marieke Jepma, our collaborator on many of our projects, for all the help related to data modeling. Prof. Dr. Rudi De Raedt, thank you for making the GOA possible, and for being a part of my guidance committee, along with Prof. Dr. Nico Böhler, and Prof. Dr. Sven Müller, who always provided insightful critique, help and advice over the course of my doctorate. Marie-Anne, thank you for the opportunity to work on your project. It was truly a breath of fresh air to work with you.

I am particularly grateful that during this period I made friendships to last for life. Naomi, what can I say? It just would not be it without you. Pain and Panic, reporting for duty! Vale,

thank you for being the way you are, a picture of a true scientist, independent, smart, and amazingly influential. I have learned so much from you girls, and it changed forever the way I think about research (and life). One of the biggest gifts of this period was that I got to know Margo, my now dear friend, a Greek math-genius. I cannot wait to see you in Toulouse! And Ro! Thank you for all the organization of the UZ testing. We made it through all those exasperating late night and weekend scanning sessions! Would have probably been impossible, and way less fun if you were not there. I am also grateful for a myriad of moments that made my stay in this lovely city so memorable: giggly, cheerful lunches with Elke, Lien's hilarious voice imitations, Ine's contagious, irresistible laughter, game nights at Sofie's, Igor's immensely charming sarcastic comments, Maud's bright spirit, Jonathan and Selene's cat stories, Steph's genuine Aussie candidness.

I am also grateful for the entirety of Panlab. I hope this little miracle will continue to exist in the precious form I encountered it in. Ghent University was smart enough to get under its wing a group of amazingly hard working and brilliant young researchers, and it was nothing but my honor to be a part of this team. I am only hoping that it was not just a group of ex-colleagues, but also future collaborators.

Special thanks for Prof. Dr. Nancy Nicolson for being a role model and for all the support in my early steps.

Posebno hvala mojim dragim roditeljima, koji su na nebrojeno mnogo načina zaslužni za ovaj doktorat. Hvala vam za sve čemu ste me naučili, za svu nesebičnost, požrtvovanje i podršku. Hvala Neveni, koja evo već godinama maestralno igra ulogu mlađe sestre, i time me beskrajno uveseljava. Mojoj dragoj porodici, baki Ljubi i dedi Iliji, baki Evi i dedi Jovi, koji na žalost nije više sa nama. Ovo je naš zajednicki uspeh, i da nije bilo vas, takvi kakvi ste, ni ja danas ne bih bila ovde. Hvala mojoj dragoj Vukici, koja je lako postala deo moje porodice.

10 Acknowledgements

Hvala Gojku, na sapu(a)tništvu. Mojim dragim prijateljima, koji su to ostali uprkos daljini i manjku vremena. Na kraju, posebno hvala Mr. Spec. Psych. Merici Milavić-Vujković, na svim profesionalnim i životnim lekcijama.

To all of you who said: “Go on, get back to writing; you need to finish that thing, we’ll talk later”.

Thank you all!

CHAPTER 1

GENERAL INTRODUCTION

1. Does tonic affect modulate reinforcement learning?

"The anterior midcingulate cortex (aMCC) did not evolve to optimize performance on laboratory measures of 'cold' cognition..." (Shackman et al., 2011, pp 9). Even the plausibility of separation between "cognitive" and "affective" brain areas and functions has been brought to question (Etkin, Egner, & Kalisch, 2011; Pessoa, 2008), especially with consideration of the integrative hubs such as the midcingulate cortex (MCC), characterized by high connectivity with other brain areas, and involvement in processes of the highest complexity, such as executive functions. Notable examples of "state dependent" changes in affectivity that lead to disruptions in executive functions are mood disorders, like unipolar major depression. The need for research of modulatory power of affect is very often recognized (Beeler, 2012; Botvinick & Braver, 2015; Pessoa, 2008; Rangel, Camerer, & Montague, 2008; Ullsperger, Danielmeier, & Jocham, 2014), mostly as extension of motivational phenomena. One of the executive functions that deserves particular attention when it comes to affect modulation is reinforcement learning (RL). It is in essence a highly flexible, complex, ongoing process of creating associations based on value. Value in itself is information about the ever changing environment, it is updated and estimated against a history of previous information (Quilodran, Rothe, & Procyk, 2008). And in humans, value is, at least partially, an affective category.

A nascent tendency in affective neuroscience is to use the RL models to gain insight into the effects of affective components on cognition. In this chapter, I would like to introduce

these first ideas of affect-cognition interaction and set the background against which our own experimental findings can be compared. In the first part, I would like to present the RL theory in more detail, show the specific predictions arising from its postulates, and place it in the context of competing theories. But in order to do that, we need to work our way bottom to the top, and first introduce some basic concepts and classical electrophysiological measures used to form predictions in different theories. To this aim, we first review neurophysiological mechanisms of error monitoring and their link to mechanisms of RL. Then, we examine what (positive) mood actually is, how it is usually defined in the literature, before we bring these two separate areas together in order to formulate predictions regarding the nature and extent of modulatory effects of (positive) mood on RL.

1.1 Error monitoring components

Electroencephalogram (EEG) is a non-invasive, high temporal resolution technique for recording electrical brain activity. Usually (but not exclusively), EEG signal is used to extract event related potentials (ERPs) related to certain internal (e.g. response preparation) or external (e.g. visual feedback on the computer screen) events. ERPs provide a pretty precise and high temporal resolution in a millisecond scale of the recordings made on a scalp level, whereas the actual sources of activity (i.e., intracranial generators) are localized deeper in the brain. ERPs are embedded in the raw electrical EEG signal, which is some 10 to 50 times larger in amplitude than the ERP of interest. This is the reason why ERPs are obtained through averaging across many trials, to extract the signal from the noise (Rugg & Coles, 1995). Importantly, this means that the ERP component of interest is in reality a compound of many trials that all have the same structure. This technique is based on the idea that the neuronal activity that generates ERPs is stable and reproducible, and the averaging procedure removes the noise (which is by definition random) from the signal (which is

constant/systematic across trials) and reliably identifies the systematic variation related to individual events or presentations. This concerted neuronal activity is a response to an event, and it corresponds to the synchronous activation of the post-synaptic dendritic potentials of a large population of neurons. These (pyramidal) neurons are arranged in a geometrical configuration such as to yield a dipolar field. ERPs provide a direct estimate of these dipolar fields. Despite limited spatial resolution of the signal obtained, the ERP technique provides valuable insight into the temporal dynamic of brain processes related to the events of interest, giving precise information about certain neurophysiological properties, through their latency (following the onset of the event), amplitude, polarity, topography (i.e. the distribution of the electric field over the scalp surface), and underlying brain generators (Luck, 2005). In the studies presented here, I will use and focus on two ERP components that have been well documented in the past in the context of RL and performance monitoring, germane with regard to their characteristics, source, and putative meaning. These two ERP components are known as the "error-related negativity" (ERN) and the "feedback-related negativity" (FRN), and their specific electrophysiological properties and putative psychological functions are discussed hereafter.

1.1.1. The error related negativity (ERN/Ne)

The ERN/Ne (Figure 1) is a negative deflection in the ERP related to error commission in reaction time (RT) tasks (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, & Coles, 1993). The ERN/Ne is usually corresponding to a difference wave that is computed by subtracting the amplitude of the response for the correct response, from the incorrect (erroneous) one. Its onset is coinciding with the moment of the response initiation as measured by the electromyogram (EMG) of the finger/hand (effector), and reaching the peak

activation some 50~80ms after motor execution. It has a fronto-central scalp distribution (FCz or Fz channels), consistent with the assumption that the dorsal anterior cingulate cortex (ACC) (or alternatively, the supplementary motor area; see Bonini et al., 2014) is the main generator of this early response-locked negative deflection (Dehaene, Posner, & Tucker, 1994; Gehring, Himle, & Nisenson, 2000; Scherg & Picton, 1991; Scherg, 1990; Vocat, Pourtois, & Vuilleumier, 2008). Consistent with this assumption, lesions to the ACC have resulted in reduction in ERN/Ne amplitudes, as well as increased difficulties for these brain-damaged patients to show normal trial-by trial adjustments during conflict detection or monitoring (Ullsperger, Danielmeier, & Jocham, 2014). The ERN is thought to reflect the neural activity of a large dopaminergic-dependent network, connecting midbrain structures and the basal ganglia to the ACC via specific or dedicated fronto-striatal loops (Falkenstein et al., 1991; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Frank, Woroch, & Curran, 2005; Gehring et al., 1993; Gehring & Willoughby, 2002; Holroyd & Coles, 2002;de Wit et al., 2012). Given its ultrafast time-course following error (or conflict) detection, the ERN is thought to reflect a highly generic error-processing system, that does not depend on the type of sensory information presented, nor the modality in which the response error is committed, and that is probably a subcomponent of a larger executive control system enabling goal-directed behavior and behavioral adaptation in case worse than expected events are timely detected (Holroyd & Coles, 2002). Accordingly, the ERN has sometimes been associated with behavioral adjustments following error (or conflict) detection, such as the post-error slowing effect (Rabbit, 1981). It is considered that the heritability of the ERN component is around 45 to 60% (Anokhin, Golosheykin, & Heath, 2008). Its amplitude is associated with subjective judgments of response accuracy (Scheffers & Coles, 2000), is increased when response accuracy is emphasized over speed (Falkenstein et al., 1991; Gehring, Goss, Coles, et al., 1993; van Driel, Ridderinkhof, & Cohen, 2012) and is reduced following incorrect responses

to stimuli that are presented relatively infrequently, conditions in which errors are particularly likely (Holroyd & Coles, 2002). Thus, all these studies converge and indicate that the ERN is a generic and reliable electrophysiological marker of (internal) error monitoring or detection, enabling the rapid and efficient processing of an important and motivationally salient event, namely (unwanted) response error. In line with this idea, some ERP studies have shown that error awareness is actually not a necessary condition for the generation of the ERN component, which can thus appear in specific experimental conditions or contexts where participants are not aware of committing response errors (Endrass, Reuter, & Kathmann, 2007; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001).

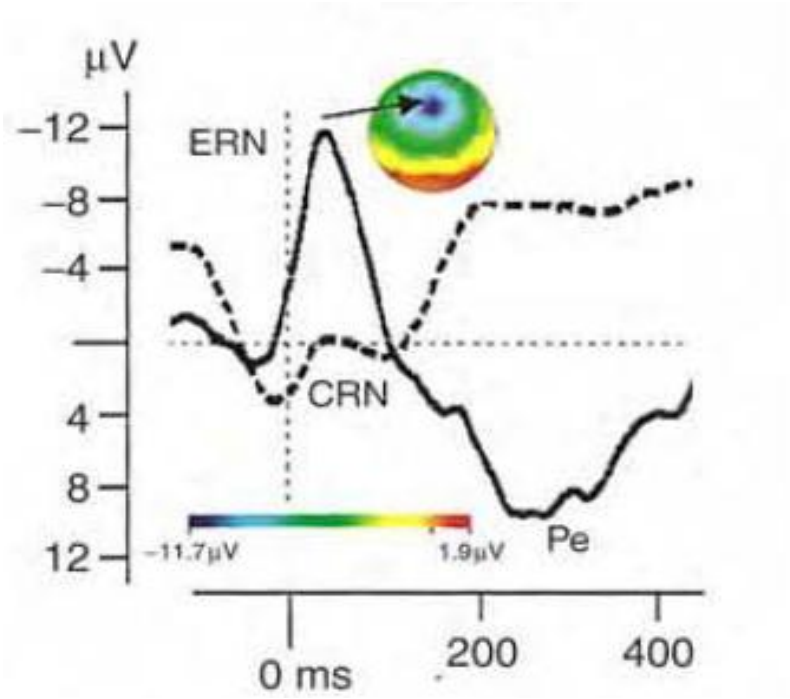

Figure 1. Illustration of the ERN/Ne and CRN component in healthy controls (from Hogan, Vargha-Khadem, Saunders, Kirkham, & Baldeweg, 2006; Figure 3)

1.1.2.The correct response negativity (CRN)

The CRN component (Figure 1) shares many electrophysiological similarities with the ERN/Ne, but is usually smaller in amplitude than this latter deflection. As its name implies, it is generated following correct responses. Because of these similarities, it has been suggested that these two ERP components actually represent the same early cognitive control process with common neural generators localized in the ACC in these two cases (Falkenstein et al., 2000; Roger, Bénar, Vidal, Hasbroucq, & Burle, 2010; Vidal, Burle, Bonnet, Grapperon, & Hasbroucq, 2003; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). The amplitude of the CRN varies depending on changes in the speed-accuracy trade-off (Gehring, Goss, & Coles, 1993), as well as on stimulus-response (S-R) congruency effects (Bartholow et al., 2005). Its amplitude goes down on (correct) trials preceding errors, which has been interpreted as a result of keeping up resources allocated to the task at hand (Allain, Carbonnell, Falkenstein, Burle, & Vidal, 2004; Hajcak, Nieuwenhuis, Ridderinkhof, & Simons, 2005; Ridderinkhof, Nieuwenhuis, & Bashore, 2003). It is considered that this component is also probably present (together with the ERN) in partial errors, corresponding to situations (as can be seen or inferred from the corresponding hand-related EMG activity) where the incorrect response/action is activated, but timely inhibited by the subjects in order to lead to the execution of the correct and competing (motor) response (Burle, Roger, Allain, Vidal, & Hasbroucq, 2008; Carbonnell & Falkenstein, 2006; Gehring, Goss, Coles, et al., 1993; Marten K. Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996). However, the CRN can be recorded in the absence of partial errors, suggesting that both the ERN and CRN components actually reflect accumulation of evidence (in terms or motor or internal representation) in favor (ERN) or against (CRN) the (automatic) detection of error commission (Ullsperger et al., 2014).

1.1.3. The feedback related negativity (FRN)

While the ERN and CRN are generated by an internal (error) monitoring system, sometimes even just before the action/response takes place, the FRN (Figure 2) provides the counterpart of these response-locked ERPs at the feedback level. It appears 250 to 300 ms after the onset of an evaluative visual feedback (following the response), which informs the participant about the occurrence of an unexpected response error or (monetary) loss. Hence, the FRN reflects the processing of an external visual stimulus that timely informs about the outcome or value of a specific event (usually an action preceding the feedback) which in turn disambiguates it or its perceived meaning. The FRN is thought to share the same neural generators as the ERN in the dorsal ACC (Gehring & Willoughby, 2002; Miltner, Braun, & Coles, 1997). The similarity of the FRN with the ERN stems from its clear association to negative compared to positive information/feedback, for which it is usually increased in amplitude. Unexpectedness or unpredictability of the feedback (regardless of its actual valence) also influences the amplitude of the FRN (Ferdinand, Mecklinger, Kray, & Gehring, 2012). Interestingly, an FRN component has also been recorded in situations where the negative feedback was associated to external source of error but not to the actual performance of the subject (Gentsch, Ullsperger, & Ullsperger, 2009). Importantly, the ERN and FRN components are strongly inter-related during the trial sequence and they react in opposite directions (Bediou, Koban, Rosset, Pourtois, & Sander, 2012; Van de Vijver, Cohen, & Ridderinkhof, 2014). If the accumulation of evidence in favor of error commission can operate based on an internal or motor representation, then the ERN, but not the FRN, shows systematic amplitude variations depending on accuracy. By contrast, if feedback is needed to decipher accuracy (because motor or internal evidence is lacking or has not been accumulated yet), then the FRN, but not the ERN, shows systematic amplitude variations with the perceived accuracy. Translated to

learning situations or contexts, which are central to the present dissertation, it is therefore usually found that these two ERP components evolve or behave in opposite ways: while the FRN prevails at the beginning of learning (exploration phase) with a progressive increase of the ERN (when task rules are learned and exploited), the ERN does so towards the end of learning with a rapid decay of the FRN (once the task rules have been internalized) (Chase et al., 2010; Eppinger, Kray, Mock, & Mecklinger, 2008). Accordingly, "a first indicator hypothesis" can be formulated (Stahl, 2010) that states that error or action monitoring is a thrifty process that timely uses the information or value available at a given moment in time (either response/motor or feedback-based in the present case), thereby avoiding the need to perform multiple checks or monitoring. Just like the ERN, the FRN is often computed as a difference wave between positive/correct and negative/incorrect feedback.

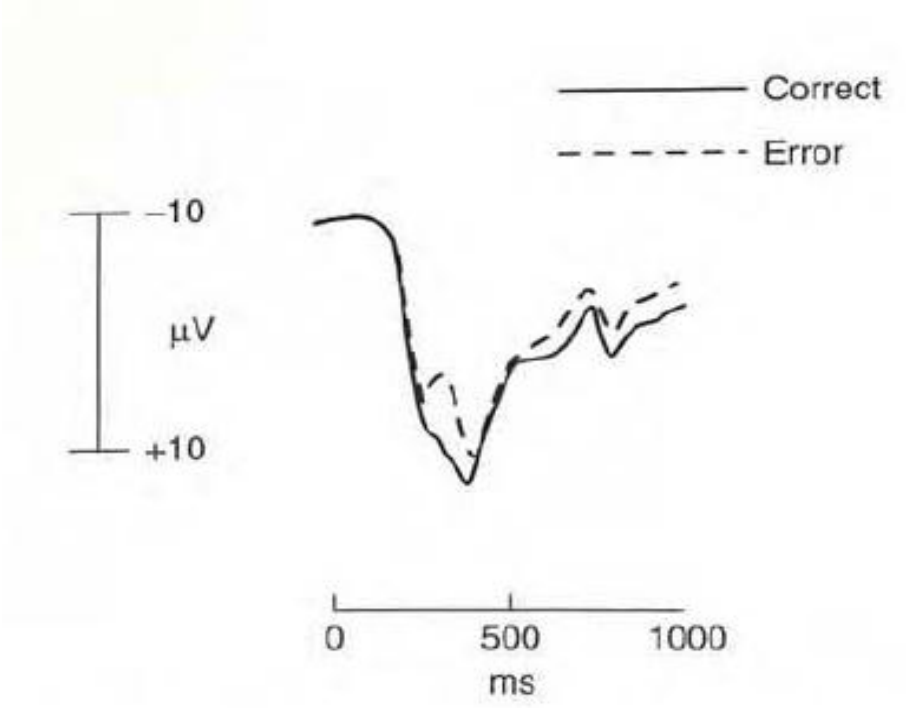

Figure 2. FRN component elicited following visually presented negative feedback during a standard behavioral task (from Miltner, Braun, & Coles, 1997; Figure 3)

These three ERP components (ERN, CRN and FRN) are characterized by a selective power increase in the theta-band frequencies, presumably produced by the pyramidal cells in

the ACC sulcus (Cohen, Wilmes, & van de Vijver, 2011; Yeung, Botvinick, & Cohen, 2004). Some novel integrative work combining evidence from different imaging and anatomical studies (Cavanagh & Shackman, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012) suggests that these, but also some other ERP components that are beyond the focus of this thesis (such as the N2, Pe and P300), are actually reflecting electrical activity/depolarization arising from the MCC. In this model, the MCC is therefore conceived as an important nexus where cognitive and affective aspects related to conflict, error, punishment, or even pain are timely integrated and in turn broadcasted to other interconnected brain areas (such as the interior insula) with the aim to foster goal adaptive behavior by warning the organism about the occurrence of a worse than expected event as well as the need to deal efficiently with it, and eventually by exerting (enhanced) cognitive control (Koban & Pourtois, 2014; Shackman et al., 2011; van Driel et al., 2012).

1.2. The mesencephalic dopamine system and the functional meaning of the ERN, CRN, and the FRN

The mesencephalic dopamine system (see Figure 3 for a schematic overview of the role in performance monitoring) projects to the basal ganglia and the (frontal) cortex (including the ACC), where this neurotransmission serves different purposes or functions. In the basal ganglia, the incoming dopamine suppression or expression is used to estimate events as better or worse than expected (and is thus given the role of "actor" in RL theory). In the ACC, this dopamine signal is used to choose the motor controller (or "critic") which should be activated in order to improve the behavioral outcome (i.e., to correct or rectify the error) (Holroyd & Coles, 2002; Ullsperger et al., 2014).

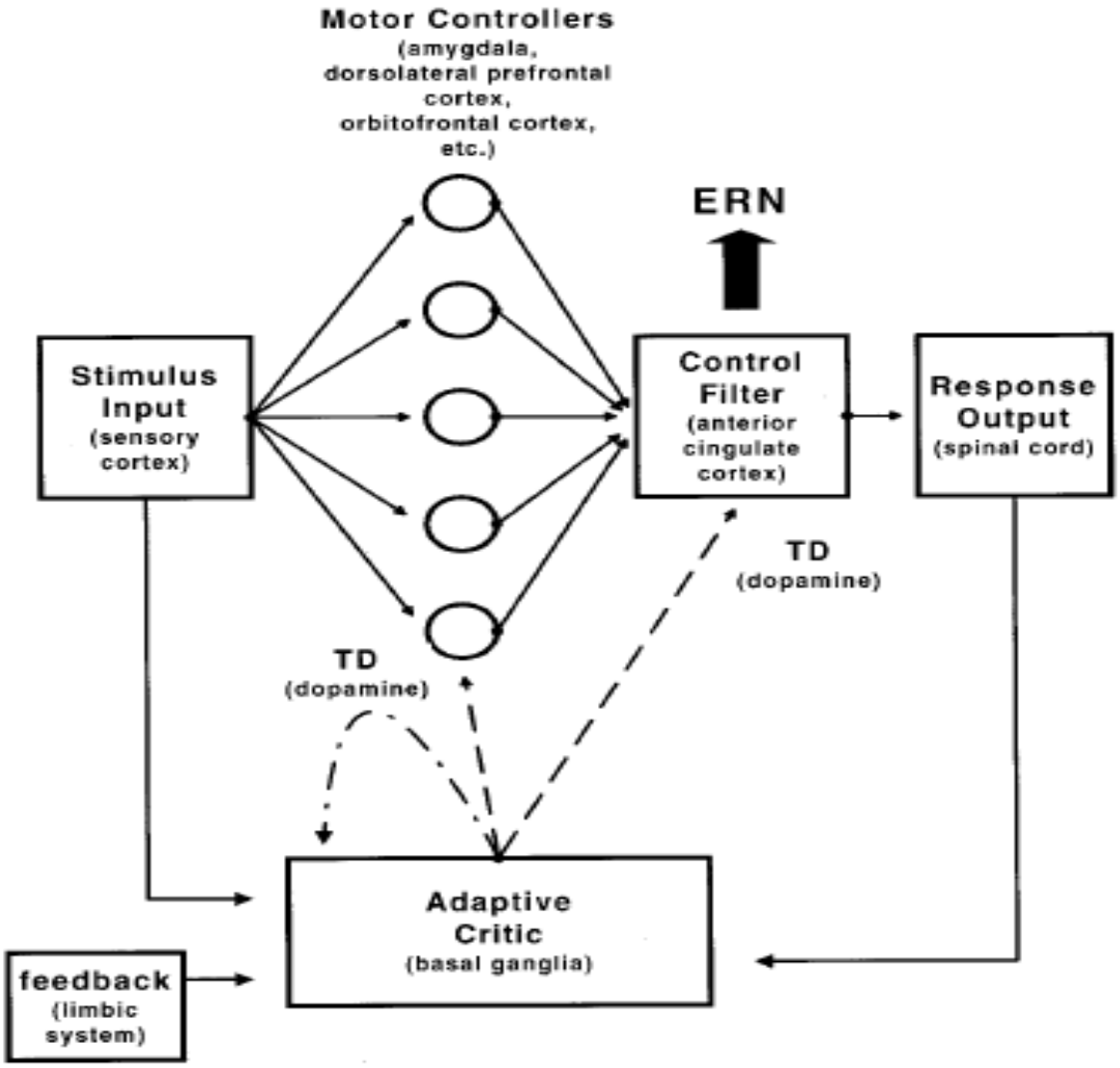

Figure 3. Schematic view of the main constituents and the main pathways belonging to the dopaminergic mesencephalic system, as integrated in a dominant error detection reinforcement learning model available in the literature (Holroyd and Coles, 2002).

Interestingly, the post-error slowing was found to vary depending on the ERN, a measure of control exerted after error commission in order to improve performance (Rabbitt, 1966). In the dominant model of Holroyd and Coles (2002) (which is disputed or challenged by some other competing theories, as we will see later in this chapter), the ERN, generated in the ACC, represents the error prediction signal per se. Initial and groundbreaking experiments performed in animals showed that dopaminergic neurons in the midbrain were reward sensitive (Fiorillo, Tobler, & Schultz, 2003; Hollerman & Schultz, 1998; Schultz, Dayan, & Montague, 1997; Schultz, 1998), presumably coding (as reflected by systematic changes in their firing rates) the pleasure elicited by the reward, with a positive dopamine signal being elicited when the event was better than expected, and a negative dopamine signal (or dip) when the outcome was worse than expected. Importantly, these neurophysiological effects were compatible with earlier computational models, which linked learning to the processing of unpredicted events (Rescorla & Wagner, 1972; Sutton & Barto, 1990).

2. Reinforcement learning

2.1. RL from classic animal conditioning studies

About a century ago, Thorndike performed a series of elegant learning experiments, with cats as his main subjects. Typically, he would put a hungry (i.e. motivated) cat in a cage, in front of which he placed food. He also provided an exit method (a lever to be pulled by the animal), which was not readily obvious (or immediately perceived as an easy solution) to cats. Initially, without any sensible idea about what might remove the obstacle between them and food, cats would try out anything and everything from their usual behavioral repertoire, or habits, to open the cage, exploring all the possible options available to them. Usually, after a while, all the frustrated tossing, scratching and hitting would lead to a positive outcome (i.e.,

the inadvertent release of the lever, the opening of the cage, and reward consumption). Interestingly enough, Thorndike noticed that upon each following return to the cage, cats' behavior was less and less hectic. Randomness was gradually reduced and the time before escaping the cage was found to become shorter and shorter. Finally, the animal reached a stadium where, immediately after being put in the cage, it went directly to the lever and pulled it, exhibiting directedness and knowledge regarding the link between this precise action and the desired outcome. In 1911, Thorndike formulated the law of effect based on these and other observations: "Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond." (Thorndike, 1911, pp 244).

Thorndike's experiments in operant conditioning paved the way for the development of reinforcement learning theory, putting forward the idea that rewards (and punishments) reinforce certain actions, profoundly shaping in turn learning and behavior. Unlike what is usually at stake in classical conditioning, where learning is primarily achieved simply because of time and space association between the two events, here it is the active striving, the examination of the consequences of actions that increase the probability of association. As much as we would like to think that we, as human beings, are more sensible and insightful in our daily goal pursuit compared to Thorndike's cats, much of what we do can nicely be subsumed in learning by trial and error, as the straight application of the law of effect. As hectic as it might appear from the example with the cat in the cage outlined here above, it

should never be underestimated how effective, powerful and ubiquitous this type of learning actually is.

2.2. RL in artificial intelligence

RL theories and their insights into a multitude of (cognitive) processes and behaviors, as found today in the field of learning psychology and neuroscience, did not only stem from research on operant conditioning during the last century, but also from groundbreaking work carried out in the area of Artificial Intelligence. As it stands nowadays, RL classically refers to a computational model and accompanying theory, originally devised by Sutton and Barto (1998), which later led to a rapid and extensive development in the field, including the validation and publication of many other and different models, all able to account for specific facets of RL. These two authors initially developed a body of mathematical algorithms that describe and predict operating in complex and uncertain environments. In the authors' words on laying the foundations of the RL framework: "…we came to realize that perhaps the simplest of the ideas, which had long been taken for granted, had received surprisingly little attention from a computational perspective. This was simply the idea of a learning system that wants something, that adapts its behavior in order to maximize a special signal from its environment. This was the idea of a "hedonistic" learning system, or, as we would say now, the idea of reinforcement learning." (Sutton & Barto, 1998, para 1).

Apparently, those that work with machines can see the clearest what is the specific, unique force behind human learning. The above quote shows in a cogent and simple way what strong forces pleasure and displeasure are in shaping learning. As could be seen from the behavior of Thorndike's cats, there are several specificities to this type of learning: first, it is motivated by *reinforcers* (rewards, but also punishments). Second, unlike what happened in supervised learning, for example, the solution that needs to be applied is not given in advance,

but the agent (i.e. human, animal or machine) has to unveil it by *trial and error*. This means that there is a constant, lively *interaction with the environment*. In the example used here above, the environment is deemed static (the lever and the cage are always the same), but oftentimes this is not the case: the actor needs to constantly examine or monitor it for potential changes that may have important consequences, effects or even values. This creates the need for adjustable systems able to constantly monitor the environment, to *explore* potential novelties, while at the same time flexibly keeping the goal in mind and performing (*exploiting*) the most useful action known thus far. The balance between these two complementary processes or forces is intrinsically a part of RL, as we will review later/below in this chapter. Lastly, a defining characteristic of RL is using previous experience, learning from the *history of previous errors*, real-time tracking of previous interactions with the environment that leads to behavioral optimization and to exploitation of the best possible option.

To provide a better definition, delineation and differentiation of the process under scrutiny here, it is perhaps valuable to say what RL is not. For example, Rangel, Camerer and Monague (2008) made a distinction between Pavlovian, habitual, and decision-based learning. Based on this insightful division, RL was equated with habitual learning. It is characterized by assigning value to a large number of actions, unlike what happens with more limited, Pavlovian-based learning systems. But also, unlike what can be evidenced with goal directed systems, action value is not updated in accordance with the current state of the system, but, instead, in accordance with the state of the system at the moment of learning. This poses certain limitations towards the predictions that can be made using the RL system. However, as we will see from further considerations of Sutton and Barto's model, today instrumental learning goes well beyond what Thorndike's cats were able to do, and does not refer to just this habitual system, but also to goal directed actions that possess much more plasticity, and

also fall under the umbrella of learning from the outcomes of one's actions (Beeler, 2012). On the other hand, Botvinick and Braver (2015) recently emphasized the importance of differentiation between reward-based systems, such as RL, and force-field, or resource models. This recent dichotomy is of importance for our own framework, as it specifies the motivational factors that underlie actions. Since RL is driven by the motivational nature of reward and punishment, these authors propose that each reward and punishment will be weighed against the history of previous reinforcements to determine whether the action is worth the effort or not. Unlike what is proposed in the force-field model (the more important the goal, the bigger the motivation) or in the resource models (the more resources available, the bigger the motivation), RL theory advocates motivation change in a history dependent way, which necessarily leads to poignantly different predictions at a practical level.

As it stands today, the RL theory has spread out and is recognized as useful in many fields, from cognitive psychology, control theory, Artificial Intelligence, neuroscience, and even neuroeconomy. Supplemented by findings from anatomy and function of certain neural systems or structures in the brain, it grew to be a useful model of human learning, capable of producing exact and testable predictions.

2.3. Human electrophysiology of RL

In 2002, Holroyd and Coles (Holroyd & Coles, 2002) published a very influential paper in which they expanded on the early computational work of Sutton and Barto, as well as on the Thorndike's law of effect, and proposed their own RL theory of error processing, which they supported with computational modeling and experimental evidence from ERP research. In their view, the function of the mesencephalic dopamine system is to signal the error. In the ACC, this error is expressed as a dopamine dip which is marked by the ERN (hence the ERN is perceived as a rather late error signaling effect, or it corresponds to the backdoor of a

deeper mesencephalic dopamine system where the primary error signal is computed and extracted), and the ACC uses it then to correct by choosing the appropriate motor action. Interestingly, the ventral ACC is believed to be activated by the evaluative processing, or the affective considerations of the importance of error. Notably, this model brings together findings related to the phasic activity observed (in animal studies) in the mesencephalic dopamine system on the one hand (Hollerman & Schultz, 1998; Schultz et al., 1997) and findings pertaining to the temporal difference error, a reinforcement learning algorithm, on the other hand. To corroborate these predictions, they used a probabilistic learning task to show systematic amplitude variations of the ERN (and/or FRN) depending on the strength of the association between the stimulus and the response. They found that the FRN amplitude was the largest if in the probabilistic - alternating setting ("random condition") there was a disconfirmation of the assumption/prediction created by the feedback in the previous trial. The authors interpreted this neurophysiological effect as lending support to the idea that the ERN or FRN precisely code prediction error.

However, it should be noted that despite its high scientific appeal and validation across many (ERP) studies, alternative views on the actual meaning of the workings of the mesencephalic dopamine system, and by extension the ERN (or FRN) component, have been proposed in the literature. Historically, some of the first interpretations were based on the rather simplistic idea that dopamine burst equals pleasure, as stipulated by the hedonic hypothesis (Wise, Spindler, deWit, & Gerberg, 1978). However, because of groundbreaking neurophysiological experiments, such as the one performed by Shultz and colleagues (1997), which showed that dopamine signal can propagate back in time and mark not the actual consummation, but the prospect of future reward (i.e. the temporal difference), this hypothesis was soon abandoned.

Alternatively, the mismatch theory proposed that the ERN is merely a representation of the mismatch between the intended and the actual response. This theory was supported by findings like the one that shows that the ERN amplitude grows when the difference between the intended and the actual action increases (Bernstein, Scheffers, & Coles, 1995; Falkenstein et al., 1991), as well as when there is a high certainty about unwanted error commission (Scheffers & Coles, 2000). In this view, response errors were often considered to be the product of an impulsive response mode. By comparison, in the RL theory, errors are not simply a consequence of (too much) impulsivity or a transient breakdown in impulse control, even though it should be stated that in many respects, these two theories accord with one another. Interestingly, the mismatch theory was able to accommodate older ERP findings showing that the ERN amplitude could also capture changes in the motivational state of the participant, besides accuracy only or per se. If altered ERN components are observed in specific emotional disorders or conditions, according to this theory, this could be due to the impaired motivational value associated to response errors in these individuals or patients, resulting in a larger or lower mismatch (depending on the specific emotional disorder identified), and hence leading to a larger or lower ERN component relative to control subjects (Luu, Collins, & Tucker, 2000).

Recently, these ideas have been extended in the motivational significance theory of the ERN (Hajcak & Foti, 2008; Hajcak, McDonald, & Simons, 2003, 2004). Central to this model is the assumption that response errors are (perceived as) aversive events, and as such, they provide strong motivational signals. Because response errors may have a different emotional/motivational significance depending on specific trait characteristics or situation factors, the amplitude of the ERN could increase or decrease depending on them. For example, if subjects' cognitive style is characterized by worry, harm avoidance or neuroticism (i.e., internalizing disorders), then the ERN to response errors is overactive, presumably

because these events become highly relevant for them. A similar motivation-related account suggests that dopamine's effects on learning could be viewed as invigorating, such that the boost of dopamine in the ACC increases the salience of the incentives or events (Berridge & Kringelbach, 2013; Berridge, 2004, 2007). Salamone et al. (2007) considered a very interesting alternative by which dopamine does not modulate the salience of reward as such, but has instead a more generic, overall energizing effect on behavior. In this view, the effect of dopamine is meant to reduce the response cost, giving energy to deal with obstacles. If this is the case, and under the assumption that the induction of positive mood leads to a boost or tonic increase of dopamine in the ACC (see Ashby et al., 1999), then we could very well hypothesize that positive mood could augment motivation and in turn increase the size of the ERN component, as this response-locked ERP component is related to the processing of the motivational salience or significance of these events. This specific prediction was directly tested in this thesis, as explained here below in this chapter.

Besides the RL theory, changes in the amplitude of the ERN have often been interpreted against the conflict monitoring theory deriving from the cognitive control literature (Yeung et al., 2004). In this framework, the ERN is not conceived as the product of an error predicting system, that is, a monitoring system that timely compares the expected and the actual (motor) outcome, but rather as a conflict detection or monitoring system. After all, response errors could be viewed as specific examples or instances of conflict (or competition between opposite tendencies or actions). These authors asked themselves a simple question: how is it possible that a putative monitoring system already "knows" somehow what is the correct response, and if it does, why would it not offer it in the first place? The main idea is that during error making or commission, two alternatives, the correct and the erroneous one, have been activated concurrently; a situation which in turn creates conflict. In this context, the ERN is therefore seen as the consequence of the continued processing of the stimulus, even

after the response has been initiated. Some initial support for this model came from somewhat counter-intuitive observations in behavioral tasks, showing that conscious error detection was in fact a slower process than the corrective action (Rabbitt, 1966), suggesting that the correct response was activated all along. The existence of the CRN component additionally lends support to this theory: why would a system that is highly sensitive to error commission produce a signal bound to a correct response? The RL theory was able to explain this apparent discrepancy by stating that the CRN is in fact an instance where both correct and incorrect responses are concurrently activated (without the need to necessarily postulate the presence of a conflict between them). Presently, it is often acknowledged in the psychophysiology literature that these three theories ("pure" RL, motivational significance or conflict monitoring) are each able to account for some of the ERP findings, but not all, with a large number of unsolved issues regarding the exact functional meaning of the ERN (or FRN component) that still need to be addressed in future studies. Recently, new solutions have been proposed, attempting to put together the RL and the conflict monitoring account. The action outcome prediction model is a generalization of the temporal difference error algorithm, so that it can predict multiple action-outcome relationships at the same time (Ullsperger et al., 2014). Furthermore, the action updating account posits that RL theory should be extended with a claim that not only error, but any outcome, regardless of valence, will evoke aMCC activation as long as it is relevant for guiding future decisions. Importantly, these new trends also incorporate considerations of "reward environment" or the state of the system at the moment of the event, which might therefore open the door for a better consideration and understanding of modulatory effects related to the current affective state of the individual (such as his/her current mood state) on (reinforcement) learning (Jocham, Neumann, Klein, Danielmeier, & Ullsperger, 2009).

2.4. Exploration-exploitation trade-off

In the RL theory, it is hypothesized that learning occurs by means of two complementary phases or forces, exploration and exploitation, and that the ACC/MCC plays a critical role in the trade-off between them (Amiez et al., 2013; Quilodran et al., 2008). If the think back to Thorndike's cats, their initial behavior was very exploratory: it consisted of blindly trying out a lot of different alternatives to solve the problem posed, "secretly" hoping that one of them would work out (i.e., lead to reward consumption). Mid-way through the experiment, the exploration was not as blind, they had presumably some ideas regarding which actions might have a higher probability of being eventually rewarded. Lastly, once they made the S-R association, and could determine with absolute certainty that the lever actually needed to be pulled to open the cage and reach/obtain the (food) reward, they did nothing else on subsequent encounters. Since the environment was static and did not pose new challenges for each new exposure, they stopped exploring completely and instead started exploiting the only solution they had learned by trial and error, with the aim to maximize reward or decrease the burden to get it. However, in real life, static environment is more the exception than the rule (Behrens, Woolrich, Walton, & Rushworth, 2007), and therefore, since the pioneer work and observations of Thorndike, clever experimental paradigms have been developed and validated to incorporate this important aspect. These two competing processes, exploration and exploitation, represent complementary tendencies in human behavior that in turn determine the learning success. Properly balancing them appears essential to foster learning, maximize reward or opportunities and avoid losses or threats. Therefore, more recent upgrades of the RL theory have dealt precisely with the modeling of these two components in order to provide a more parsimonious and realistic framework eventually able to explain or model a wide range of (learning) behavior.

According to such recent theoretical developments in the field of RL it is assumed that the basal ganglia are responsible for the exploitation of the known rewards (Frank, Doll, Oas-Terpstra, & Moreno, 2009), which may be why it inhibits dopamine firing once the reward is not there anymore, for whatever reason. The prefrontal cortex is thought to guide the exploratory behavior, as top-down control is necessary to override the learned exploitative tendency (Duncan & Owen, 2000). Importantly, it is considered that the change in the exploration-exploitation trade-off is driven by the information about the stability of the environment (Behrens et al., 2007; Browning, Behrens, Jocham, Reilly, & Bishop, 2015; Humphries, Khamassi, & Gurney, 2012), which opens the possibility that the subject's internal, affective environment, or the tonic mood level/state at a given moment in time, could very well influence this trade-off during RL (possibly, in a valence specific way). This prediction too lies at the heart of the present dissertation, as outlined hereafter in this chapter.

In the RL framework, exploration and exploitation are usually modeled formally through two parameters that stem from the temporal difference rule and track how (learning) behavior is shaped by the outcomes (and values) of actions. One of them is the learning rate, or α, which denotes how much value the new information has in changing the previously established values (Woergoetter & Porr, 2008). α represents the levels of uncertainty in the estimate of the action's value. In volatile environments, where a lot of exploration is needed, α will be high, as the recent/proximal experience is more predictive of the future than the distant experience. On the contrary, when the environment is more or less stable, α should be low, as new actions do not contribute much to what is already known (Beeler, 2012). Secondly, the exploration parameter, or β, tracks how much probing participants did for each potential response option before they stabilized their response on one of them. There are different models that try to approximate the most optimal behavioral strategy by applying different modulations of this parameter. For example, greedy model supposes that the

modeled behavior will always choose the most lucrative option known thus far. Obviously, this completely excludes or undermines the exploration and does not always present a viable model of behavior in changing environments. An alternative, so called "softmax" model, is made to choose the most rewarding option most of the time, while still offering sometimes the possibility to explore other ones (Jepma & Nieuwenhuis, 2011). Extracting α and β parameters from this model can provide us with a lot of information regarding the actual strategy (and not only the amount or rate of learning) used by the subjects to foster accuracy in a specific environment. Using this sort of modeling and gaining insight into the exploration-exploitation trade-off during RL, an elegant study performed with macaques previously showed that (toxic) lesions to the ACC sulcus led to the inability to use the (full or detailed) history of previous outcomes to guide (future) learning, eventually creating a deleterious condition whereby the animal could only plan the next action based on the very last outcome (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006). Additionally, Behrens and colleagues (2007) were able to show that ACC activity was augmented with increasing volatility, and that it also correlated neatly with higher learning rates, as revealed by the corresponding behavioral data.

2.5. Probabilistic learning task

All the research conducted and presented in this dissertation has been based on the use of a previously validated probabilistic learning task, inspired directly by the work of Holroyd & Coles (2002). It is common practice in RL experiments to use such probabilistic (learning) tasks because these paradigms offer a valid way to imitate (or provide a proxy of) volatility encountered in everyday life situations during decision making and behavior optimization (for

examples see Bhanji, Beer, & Bunge, 2010; Frank et al., 2005; Holroyd & Coles, 2002; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2009; Unger, Kray, & Mecklinger, 2012). As has been pointed out by Beeler (2012), Torndike's law of effect is actually a law of probability – it marks that an association becomes more or less likely depending on the positive or negative reinforcer encountered therein. Probabilistic (learning) tasks, which capitalize directly on the use and exploitation of incentive cues (or feedback) to guide learning provide thereby suited experimental paradigms to examine closely this phenomenon (RL). The specific task settings used throughout the empirical chapters gathered in this thesis have been previously devised by Eppinger et al. (2008) to study age related changes in RL at the level of the ERN and FRN components (with focus on their amplitude variations as a function of learning and reward probability; see also Frank et al., 2005 for a similar procedure and Van de Vijver et al., 2014 for the use of theta-band oscillations using a very similar task/procedure). More recently, this exact same paradigm was also used to study effects of behavioral inhibition (defined as a trait disposition; see Unger, Heintz, & Kray, 2012) on RL. An asset of this paradigm is that, besides standard indices of learning at the behavioral level (RT and/or accuracy), learning can also be titrated in terms of changes in the learning rate and the exploration (even though we have to be aware of the fact that because a two-alternatives forced choice task is used in this experimental paradigm, this latter parameter reflects more randomness of decisions than exploration per se, unlike what has been done, for example, in n-arm bandit tasks where multiple choices are feasible for each and every trial; see Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006).

Typically, there is a limited and fixed number of visual stimuli used in one block and shown many times in random order, amounting to 6 (each repeated 40 times) in each of our different experiments (the number of blocks varies from one study to another and is meant to

increase the signal and power for ERP data analysis). These 6 stimuli are assigned to different reward probability conditions, unknown to the subjects (Figure 4a).

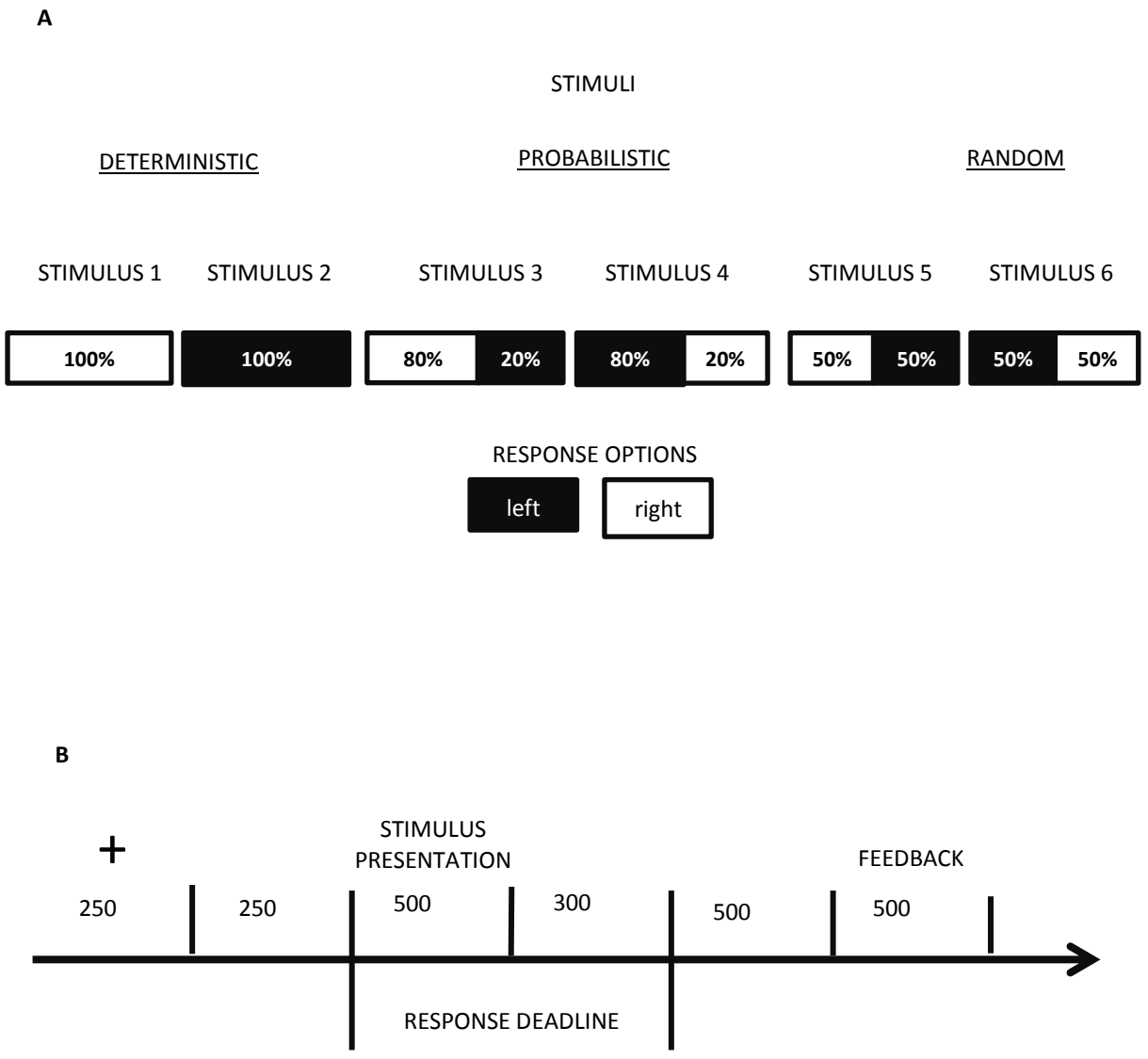

Figure 4. The task (a) and the trial structure (b) of the probabilistic learning task previously devised by Eppinger et al. (2008) and used throughout this thesis.

The subjects are required to perform a 2-alternatives forced choice task with time limit (to avoid the use of uncontrolled strategies): for each and every stimulus presentation, they

have to decipher to which response (either 1/left or 2/right) this stimulus is arbitrarily associated to. The stimuli are always presented one by one, and in a random order. Two of the stimuli from the set belong to a so-called deterministic condition, which means that one of them is always associated with one response option (say, response 1), while the second one always with the alternative/other one (say, response 2). In the so-called probabilistic condition, one stimulus is associated with one response option 80% of the time (i.e., 32 trials out of 40 possible), and 20% with the alternative one (i.e., 8 trials), while the other stimulus has a symmetric probabilistic S-R association. Finally, in the random condition, both stimuli are associated equally often (50%) to each response option. This condition is used as a control condition, because by definition no learning is possible here. Usually, learning is the largest for the deterministic condition, intermediate for the probabilistic and merely absent for the random one. Importantly, in the standard version of this paradigm, each time the participant gives a response, an evaluative (visual) feedback follows after a short interval (usually set to 1000 or 1500 ms) and it informs him/her about the accuracy of his/her last decision. Because the three conditions (differing regarding their actual reward probability) are always shown in random order, some uncertainty (or even volatility when considering the probabilistic and random conditions only) necessarily arises at the level of decision making on each and every trial (with a decreasing gradient for this uncertainty depending on trial number encountered), which is a pre-requisite to foster exploration of the feedback and a potential shift to exploitation after some time or trials, hence yielding optimal conditions to study (the dynamic of) RL.

Remarkably, using this paradigm, the actual amount and efficiency of learning can be titrated for each subject separately, not only at the behavioral level (including computational modeling methods to extract the learning rate and exploration), but also at the electrophysiological level concurrently, because reward probability and RL strongly influence

the amplitude of the ERN and FRN components in this task (see Eppinger et al., 2008; Frank et al., 2005). In accordance with the tenets of the RL theory, it is expected that in the deterministic condition, the ERN (response-locked component) will be the largest, while the FRN (feedback-locked) will be the lowest. By comparison, in the random condition, the pattern of results should be mirror-symmetric (small ERN but large FRN), while in the probabilistic condition both ERN and FRN should have intermediate and comparable sizes. Indeed, in this latter condition, externally provided feedback on task performance (and thus the FRN) continues to be informative throughout the task, while some internal knowledge (and thus the ERN) may help to guide decision making in this condition (see also Van de Vijver et al., 2014 for convergent evidence using theta-band oscillations, rather than ERP amplitudes).

Strikingly, to the best of our knowledge, all the empirical research carried out so far using this kind of experimental paradigm (probabilistic learning) assumed implicitly that these learning effects somehow unfolded in a vacuum. In other words, the specific affective state of the participant was not considered to play any significant role in the dynamic changes in the exploration or exploitation (or both) of incentive cues, and it was therefore deemed neutral, or sometimes even uninteresting. Even though it appears logical to draw such a strong conclusion when exploring RL performed by algorithms or machines, when turning to animals, and more specifically humans, the reality or learning situation/context may actually be much more complex. Human subjects are very rarely experiencing a genuine neutral mood or affective state for a long duration. Emotions and moods dynamically shape a wide range of cognitive and emotional processes, ranging from low level visual perception to higher-level reasoning or decision making, as evidenced in the healthy brain, but also in some pathological conditions characterized by mood disorders (Barrett, 1998; Damasio, 1989; Pessoa, 2008). In this context, RL is also probably subject to modulations depending on specific affective state

or mood factors. Accordingly, one of the main goals of this dissertation was to fill this gap and explore possible changes brought about by mood during RL (when this process was characterized at multiple levels concurrently), instead of assuming it merely did not play any role, and was at best a nuisance. As a matter of fact, pleasure seeking and harm avoidance have become important evolutionary mechanisms, powerful motivational drives, which shape our behaviors in a way which allows maximizing survival and/or reproduction (Darwin & Ekman, 1872). Presumably, and speculatively at this stage, RL does not simply operate based on learned values regarding the reward probability of specific stimuli or external events/incentives available in the environment. Instead, this process probably depends also (at least in part) on the specific dispositions and affective states characterizing the learning agent at a given moment in time. Hence, RL efficiency is probably depending on a complex interaction effect between learned values (stimulus-related effects) and affective states (subject-dependent effects). Before I turn to the specific predictions that could be formulated regarding possible modulatory effects of mood on mechanisms of RL (including in pathological conditions where negative affect prevails, such as unipolar major depression), I will first briefly review in the next section some theoretical reasons and the experimental evidence available, supporting the idea that RL is not devoid of state-dependent (mood) influences.

3. Modulatory effects of mood on RL: support from behavioral, electrophysiological, and neuroanatomical findings

In the previous sections, we reviewed evidence (and theoretical models) showing the prominent role of the mesencephalic dopamine system in RL. The concept of reward (and punishment) has been used extensively to conceptualize the interplay of motivation with

learning. In this framework, (phasic) dopamine firing (or the lack thereof) in mesencephalon (and ACC) shapes learning, because it reflects or encodes the expectation and salience of pleasure. But besides this critical role in mechanisms of RL, these same dopaminergic pathways are also involved in the experience and maintenance of specific affective states. A paradigmatic finding is that blocking these dopaminergic-dependent pathways actually leads to the loss of rewarding/hedonic properties/experiences usually associated with specific reinforcers, such as food or water (Lyness, Friedle, & Moore, 1979), suggesting that dopamine mediates positive affect and influences cognition (Ashby, Isen, & Turken, 1999; Goschke & Bolte, 2013; Ridderinkhof et al., 2012). This idea, however, has received little empirical support, probably because it has been considered as an oversimplification, as affect generation and experience depends on a large scale brain network and the coordination of many areas (both cortical and subcortical), where dopamine (and the mesencephalic pathways underpinning this neurotransmitter) is an important player, but not the only one obviously (several hormonal systems, glucorticoids, or opioids likely contribute to affect generation too; see Yue, Vessel, & Biederman, 2007). But before going into further considerations regarding the potential effects of positive (or negative) affect on RL, it appears important to define what affect (or mood) exactly is, to delineate it clearly from other concepts that are sometimes used interchangeably, such as motivation, reward or arousal, and finally to clarify the commonalties and differences between the two affective categories most frequently used in the affective science literature: mood and emotion.

3.1. Different types of affective experiences

Affect has been defined as a hedonic quality of pleasure or displeasure (Berridge & Kringelbach, 2013). Here it is used as an umbrella term to cover different concepts that include this hedonic quality, but may differ in other aspects (like in Ashby et al., 1999; Berridge & Kringelbach, 2013; Goschke & Bolte, 2014); sometimes it is also used simply as a synonym for emotions (as in Dreisbach, 2006; Lang & Bradley, 2010; Schwager & Rothermund, 2013).

Affect and motivation. Motivational systems are concerned with maximizing pleasure and minimizing pain (Madan, 2013). They are often considered to be evolutionary pristine and basic, based mainly on a "fight or flight", or "approach or avoidance" simple dichotomy. As such, they have a strong motor component of physical approach or avoidance directed towards external stimuli (Lang & Bradley, 2010; Mobbs et al., 2009). Berridge and Kringelbach (2013) consider affect-related phenomena to be founded on brain's motivational circuits and thus they are an extension of motivation. In this view, affect is seen as a disposition to action, based on motivation towards or away from something. Every affective experience can be qualified on the intensity of motivational mobilization. Motivation has also been described as the invigorating impact of prospective reward on cognition and behavior (Botvinick & Braver, 2015). These authors propose that motivation can have short term effects on cognitive control, as in post-error slowing, or increase in accuracy, and long term consequences as well, as in ego depletion phenomena for example (but see Inzlicht, Schmeichel, & Macrae, 2014). Importantly, there is now evidence accumulating that the motivational effect of reward can be seen as increased activity in the (dorsal) ACC, together with increase in behavior towards a more relevant target dimension (Vassena et al., 2014).

Affect and reward. In operant conditioning, reward is defined as an appetitive outcome of an action, opposite to punishment, which is an aversive outcome. It is a reinforcer, in the sense that it strengthens certain actions while eliminating/pruning other ones. Reward is considered to have three core aspects: the hedonic component, sometimes referred to as the consummatory pleasure (or "liking"); the motivation component (or "wanting"); and finally reinforcement learning, that follows as an outcome of these two components (Berridge & Kringelbach, 2013). Affect shares the first component with reward, and the second to a certain extent, while the third aspect is the research topic under scrutiny in this dissertation. In an attempt to separate the effects of positive mood and reward, some recent studies compared the effects of task-contingent reward, task non-contingent reward (thus, a rewarding outcome that would come unrelated to whether the action performed was correct or not), and task-non-contingent positive mood (induced by showing positively-valenced images) on the proactive and reactive components of cognitive control (Braem et al., 2013; Fröber & Dreisbach, 2014). The results showed that only task-contingent reward was increasing proactive control, while task-non-contingent positive mood actually reduced it. In a similar design, with a different mood induction, Chiew and Braver (2014) have found the opposite - increased proactive control related to positive mood. These findings are of importance here because the concept of cognitive control and flexibility shares similarities with the exploration-exploitation dilemma in RL. In a way, we can conclude that reward is a special case of positive affect that represents information about the environment. It is bound to an action, because it is a result of it, and it is likely in itself to produce the same action again. Importantly, a very recent review (Botvinick & Braver, 2015) suggests that reward modulates cognitive control via preparatory, proactive processes, where midbrain dopamine is assumed to play a crucial role.

Affect and arousal. All affective phenomena have a certain inherent arousal or (physiological or body-related) activation level according to which they can be differentiated. In fact, arousal

and valence have been identified as the two most important defining dimensions of affective experiences, often conceptualized as orthogonal to each other, and used for their classification along a Circumplex model (Barrett, 1998; Knutson, Katovich, & Suri, 2014; Russell, Weiss, & Mendelsohn, 1989). Sometimes, it was considered that the higher the arousal, the higher the disposition to an action or motivation change/mobilization related to a certain affect (Harmon-Jones, Gable, & Peterson, 2010). Affect is related mostly to dopamine and serotonin, while arousal is dependent more on norepinephrine and acetylcholine (Aston-Jones & Cohen, 2005; Berridge, 2007). In the past, there have been research attempts undertaken to separate these two concepts by inducing affect-free arousal (as in physical exercise, for example) that showed that arousal by itself does not have the same effect as induced affect (Rowe, Hirsh, & Anderson, 2007). Further, Van Steenbergen, Band and Hommel (2010) also showed that arousal did not influence conflict adaptation in a flaker task, whereas valence of affective experience did so.

As a side note, or a short disclaimer, even though we believe it is important to assess the similarities and differences between mood, arousal and emotion (from a conceptual point of view), it is not our intention in this work to differentiate them systematically and eventually explore their respective (and possibly different) contributions to changes in mechanisms of RL. As will be reviewed later in this chapter, we used specific mood inductions in our experiments meant to induce either positive or negative mood in healthy adult participants (or alternatively, to study pathological cases where positive mood is strongly reduced, such as in major depressive disorder, MDD), but these procedures also changed inevitably "uncontrolled" factors, such as arousal and/or motivation. In our own research, we therefore created different mood groups or conditions based on the content of the affect currently experienced and maintained, but this manipulation never entailed that no concurrent changes in arousal or motivation between these conditions took place. In fact, the ecological validity

of our mood inductions would probably have suffered substantially if we had tried to induce either positive or negative mood without influencing motivation or arousal concurrently. Presumably mood or affect is always made up of a special blend of valence, arousal and specific motivational ingredients. The above mentioned studies of Rowe et al. (2007) and van Steenbergen et al. (2010) already convincingly showed that certain modulatory effects of cognition (or attention) traditionally imputed to affect can best be explained by a change in both valence and arousal, as opposed to valence alone for example. But clearly, this valuable research line focused on the differentiation between different affective or emotional components was not the one we elected or favored in our own research in this thesis.

Different subtypes of affective experiences. Following Bolte and Goschke (2010), emotions are defined here as complex psycho-physiological response patterns, that include evaluations of events with considerations of the organism's current needs, motives and goals. They include changes in the peripheral nervous system (sweating, increased heart rate, pupil dilation) that are often used in experiments as measures of emotion-related autonomic activity. They include activation of dopaminergic and serotonergic systems and many brain areas (at the cortical level, including the orbitofrontal cortex and ACC; at the subcortical level, encompassing the insula, nucleus accumbens (NAc), ventral pallidum, amygdala, and mesolimbic tegmentum). They are usually followed by a certain distinct facial expressions and sometimes also posture, gait or movement. Even though there are some emotions with low action tendencies, for example, amusement or contentment, traditionally it is considered that intrinsic to emotions is the disposition to act (Frijda, Kuipers, & ter Schure, 1989), expressed as emotional engagement of sensory systems, increased attention, facilitated perceptual processing, initiated reflex responses, and motor actions, and finally, mobilization of the whole organism (Lang & Bradley, 2010).

By contrast, moods are classically a type of affective experiences marked by lower intensity and longer duration than emotions. Initial emotional reaction is quick and brief (short lived, and usually bound to an object, event or stimulus, either externally presented or internally available), reflecting a rapid burst of hormones, (which can be somewhat prolonged by additional firing from other indirect and direct projections, modulating the initial signal), while mood (that can be devoid of an object or medium) is considered to be related to more tonic, longer and milder shifts (Goschke & Bolte, 2014). It can be described as a "non-focal, colorizing" experiential quality (Goschke & Bolte, 2014), a diffuse affective state, not directly elicited by a certain specific event or object, and is usually considered to be of lower intensity than emotion (Beedie, Terry, & Lane, 2005; Gross, 1998; Larsen, 2000). Additionally, it is also expected that emotions are always conscious and always have attention focused on an emotion-eliciting event, whereas moods are diffuse and can develop without awareness (Larsen, 2000). Importantly, what differentiates both moods and emotions as affective experiences from similar concepts is the subjective experience, the conscious, appraised hedonic quality.

Emotions have been more prevalent (and in turn much better studied) than moods in the affect-cognition literature (Bolte & Goschke, 2010). Even though mood and emotion research lines share similar goals, the models, predictions and induction techniques differ. It has never been tested directly, but because of the differences in time course, strength, and object-relatedness, moods and emotions probably do not lead to the exact same effects in terms of the way they bias the neural system and homeostasis. For example, there is a lot of research using emotionally valenced stimuli (mostly visual), presented for a few hundred milliseconds, integrated in a task procedure, and repeated across several trials (for example, see Fröber & Dreisbach, 2014; Pizzagalli et al., 2009). These brief, phasic presentations of affect-salient stimuli are typically considered to elicit fast, automatic, strong reactions, that are

then tied to a certain component of the task the subject is performing (e.g., it can facilitate or interfere with the task at hand depending on the actual embedding of this emotional event in the experimental design). Sometimes in such studies, there is no clear or overall change in the reported emotion levels between the emotion/experimental and the neutral group. Using this procedure, a certain number of studies (Braem et al., 2013; Fröber & Dreisbach, 2014) was guided by the hypothesis that only task-contingent emotion could influence learning. Or, for example, it is expected that because emotions include active maintenance of emotion-related information in working memory, then, they could lead to congruency effects, while that is not (or less clearly) the case with moods (Schwager & Rothermund, 2013). On the other hand, mood induction studies usually use techniques that "set the tone" before the beginning of the task (or a part of it) and that are not contingent on the task or task performance. Even though there is no task-relatedness, it is considered that the shift towards more positively or negatively valenced mood can nevertheless create differences in the way the task is eventually performed. For example, in their dominant theory informing about modulatory effects of positive affect on cognition (with a focus on creativity, or reasoning), Ashby et al. (1999) propose that "mild positive affect, of the sort that most people can experience every day" can lead to increase in cognitive flexibility if the situation is at least minimally engaging or involving, emphasizing how subtle or labile influences of mood on cognition may be.

Thus, focusing on moods or emotions leads to somewhat different expectations about the outcomes of their influence. Both are undoubtedly important for understanding the modulatory capacity of affect onto cognition broadly defined, and more specifically RL in the present work. Studying mood has the appeal of showing that affect does not have just momentary influence, but can be accumulated through stretches of time and exert its effect in a long run, for better or worse. After all, the Major Depressive Disorder (MDD), for example, is still classically defined primarily as a mood disorder by the Diagnostic Statistic Manual

(DSM V, American Psychiatric Association, 2013), further emphasizing that mood (alterations) may be a central component of major changes or disturbances in cognition and affect. Nowadays, a steady and growing interest for the effects of positive mood on building psychological resources in many diverse fields of human activity ("a happy worker is a good worker", to name one) is observed in the psychology literature. As we will see in the following part, there is now a substantial amount of research focused on mood-related modulations on cognition, which is particularly concerned with valence-related effects, and more specifically, with the putative beneficial or protective effects created by positive emotions (see Ridderinkhof et al., 2012 for a recent application in the context of RL and Parkinson disease).

3.2. The (valence-related) effects of mood on cognition

Much of the research focused on the effects of mood on cognition relies on a valence-based dichotomy. One of the first elaborated formulations of these ideas came from the "broaden and build" theory of Fredrickson (2004), which brought to light two important themes. First, positive and negative moods have different evolutionary purposes: following the idea that affect is associated with specific action tendencies (Frijda et al., 1989; Kemper & Lazarus, 1992; Tooby & Cosmides, 1990), it was proposed that positive affect instigates engagement, activity, and facilitation of approach behavior. On the other hand, negative affect does the opposite, it focuses, narrows the attention and commits the capacities fully to an event of importance for survival. This association of positive affect and broad and flexible cognitive styles will live to have its different emanations and developments. For example, Fiedler et al. (2001) proposed that positive mood supports active generation of new inferences and

engagement in creative thoughts, while negative mood supports the conservation of already known material, focusing on details and factual information. Relatedly, Bless and colleagues (1996) proposed that in benign situations subjects rely more on their general knowledge structures, and in problematic situations attention is focused more on specific details. Because problematic situations are a deviation from routine, it is a bad idea to rely on the knowledge that is usually applied and therefore the focus should probably be shifted to the specifics of the current situation in order to cope with them efficiently.

Secondly, even though both positive and negative affect is evolutionary important on a short scale, when considered in the long run, they lead to poignantly different outcomes, with direct (and opposite) impact on health and wellbeing. Fredrickson's theory contrasted positive and negative affect related to their long-lasting health effects. According to an "undo" hypothesis, positive affect was thought to accrue resources that help fight off detrimental effects of future stress, with support from experimental studies that showed that positive compared to neutral mood induction leads to faster cardiovascular recovery (Fredrickson, 2004). In such framework, it is never imagined that there is such a thing as too much positive mood or that it can be in any way deleterious on the long-run. Negative affect, on the other hand, if it maintains the prolonged focus on threat, can lead to hopelessness, helplessness, withdrawal, passivity, impaired ability to generate defensive actions, self-denigration and pessimism, and finally, by "circling the drain" in a downward spiral, a fall into a disorder such as depression or anxiety, characterized by an excessive dread of negative feedback, rumination and severe emotion regulation problems (Disner, Beevers, Haigh, & Beck, 2011).

The largest empirical support for this theory came from creativity research, where the beneficial effects of positive mood on flexible performance were found in word association tests (Isen, Johnson, Mertz, & Robinson, 1985), categorization and classification tasks (Isen, Daubman, & Nowicki, 1987; Isen & Means, 1983; Kahn & Isen, 1993), creativity and

innovative problem solving (Isen et al, 1987) to name only a few. Further experimentation showed that by inducing mild positive mood one could influence cognitive fluency (Phillips, Bull, Adams, & Fraser, 2002), functional fixedness during problem solving (Gasper & Clore, 2002), create a broader attention scope (Vanlessen, Rossi, De Raedt, & Pourtois, 2013, 2014), improve ability to overcome dominant responses, and shift from local to global perceptual features in a flexible way (Huntsinger, 2012).

But if positive mood indeed signals safety and encourages exploration as well as a flexible search of the environment for potential opportunities or rewards, then it should influence processes such as goal maintenance, flexible set shifting, and working memory updating as a part of cognitive control and executive function (Duncan & Owen, 2000). For example, can positive mood contribute to increased flexibility in overriding behavioral automatisms? In experiments that tap into these processes it was brought to question for the first time that positive mood is indiscriminately beneficial for performance. Goschke and Bolte (2014) propose a control dilemma hypothesis, where they show that the positive mood resulting in increased flexibility, if not contingent to the task at hand, might lead to less engagement with the current task and more distractibility, something that in real life can have both positive and negative consequences. So these authors propose that task-relatedness is an important concept to use when rethinking the role of positive mood in cognitive control. The positive mood is signaling safe environment in which there is no need to be frugal or to count each reward (Harmon-Jones, Gable, & Price, 2013). In such context, subjects will allow themselves more exploration than in a neutral or sad context, so they perform worse or poorly in task switching paradigms (Phillips et al., 2002), or in tasks that test proactive control (Dreisbach, 2006). In a similar vein, Braem et al. (2012) recently found that presentation of pictures with positive emotional content lead to less goal shielding of task-sets in a task switching paradigm. In accordance with this, van Steenbergen et al. (2010) found that

pleasure ratings correlated negatively with the amount of conflict-driven control on a flanker task. Zwosta et al. (2013) found greater switch costs for the participants in positive than in the negative mood group. As an alternative, it was proposed that positive mood could provide flexibility to change the focus as instructed: if a narrow focus is required, then positive mood will allow for this (change from global to local) better than the negative mood, as well as for the broad focus (Huntsinger, 2012). Similarly, Marien and colleagues (2012) offered an interesting alternative by showing that positive mood reduced or increased switch costs depending on what focus has been induced previously, in accordance with the idea of hierarchical, instruction-based control over flexibility and rigidity as two processes that can be concurrent, not just sequential (Cools, Lewis, Clark, Barker, & Robbins, 2007).

First hints on the possible neurobiological mechanisms behind the influence of positive mood on cognition came from Ashby et al. (2002) dopamine hypothesis of positive affect. In an attempt to put together the findings that positive mood influences prefrontal-based cognitive control mechanisms, and that at the same time these same functions depend on dopamine projections to and from the ventral tegmentum area (VTA), these authors proposed that the mesencephalic dopamine system moderates the relationship between reward and the liking component, on the one hand, and the relationship between reward and cognitive control, on the other. The support for this hypothesis is indirect, reflected in the idea that increased dopamine and positive affect have somewhat similar effects on cognitive control. Imaging studies provided direct evidence that the presentation of different types of pleasant emotional stimuli elicits activation in brain structures involved in the processing or prediction of reward, such as the VTA, nucleus accumbens, caudate nucleus and medial prefrontal cortex (PFC) (images of beautiful faces, romantic couples, erotic pictures, pictures of subjects' family members, beautiful paintings and pleasant music; see Goschke & Bolte, 2014). Notwithstanding all the challenges remaining, the mesencephalic theory of positive affect

remains influential today in the literature. In a recent review article (Cavanagh & Shackman, 2014), the adaptive control hypothesis was proposed, where the MCC has been given a role of an adaptive control hub, encompassing both cognitive control processes and the affective signals, as they are both concerned with the same issue- supporting goal directed behavior. These authors call on the evidence that subjects with dispositional anxiety show enhanced frontal-midline activation during control tasks, usually measured as ERN, FRN or N2.

Noteworthy, systematic research that tried to associate specific electrophysiological components with positive mood is still scant. If the dorsal ACC (or rostral cingulate zone; see Ullsperger et al., 2014) receives dopamine projections from the VTA and striatum, if it is signaling the change in reward related to dopamine shifts, then ERP components such as ERN and FRN that are thought to arise from this medial PFC region might very well be influenced when (positive) mood is transiently increased, compared to a control mood condition where it is not. There have been findings associating increased FRN component with higher scores on Behavioral Approach System (BAS) questionnaire, thought to measure motivation by positive incentive stimuli (Lange, Leue, & Beauducel, 2012). However, usually, these ERP components have been examined in the literature on affective disorders, such as anxiety (Olvet & Hajcak, 2008), depression (Vanderhasselt et al., 2014), schizophrenia (Kring & Barch, 2014), obsessive-compulsive disorder (Ursu, Stenger, Shear, Jones, & Carter, 2003), substance abuse (Easdon, Izenberg, Armilio, Yu, & Alain, 2005; Ridderinkhof et al., 2002), as opposed to state-dependent effects (such as positive or negative mood) in healthy controls or their alterations in specific pathological conditions. Unipolar major depression is of particular interest here as a model of negative emotionality because of increased levels of Anhedonia or lack of enjoyment for appetitive stimuli or activities that were previously experienced as such (Pizzagalli, 2014). Additionally, it is also characterized by increased levels of negative affectivity, worry, abnormal stress reactivity and low mood (Disner, Beevers, Haigh, & Beck,

2011). In interference tasks such as the Stroop task, an increase in ERN amplitude was found for depressed subjects (Holmes & Pizzagalli, 2008). Similarly, Chiu and Deldin (2007) found increased ERN amplitude related to depression in neutral and punishment conditions of a Flanker task, but not during reward-related conditions, suggesting that depressed individuals are particularly sensitive to punishment, and that changes of the ERN may very well be state dependent. The enhanced sensitivity to loss or punishment is a landmark of depression, anxiety, and negative affect or internalizing disorders more broadly. In the literature on cognitive control and depression, this idea has prevailed and has often been used to explain impairments in this ability as a function of depression/negative affect. For example, it is considered that in switching paradigms, depressed individuals perform poorly or worse (than controls) due to the high salience given to any negative indication about performance, without any estimation of the validity of such information (Elliott et al., 1996; Mies et al., 2011). It was even suggested that this effect could be conceived as an endophenotype, translating the (possibly inherited) exaggerated susceptibility to pronounced, increased reactions to negative feedback about one's performance, that might ultimately play an important role in developing a mood disorder, such as depression (Olvet & Hajcak, 2008). However, at the neurophysiological level, the literature on the ERN and FRN components in depression has provided somewhat mixed results, indicating potential effects of depression's severity (Weinberg, Riesel, & Hajcak, 2011), or the severity of anhedonic symptomatology (Pizzagalli, 2014). Notwithstanding these elements or uncertainties, it should be noted that systematic research focusing on psychopathology (e.g., depression) is needed in order to better understand the complex interplay of mood factors with cognition and learning, and more specifically RL. We adopted this logic in Chapter 4 of this thesis, where changes in RL as a function of unipolar major depression were explored and characterized at multiple levels concurrently (behavioral, computational modeling and ERP).

3.3. Mood and RL

Although indirect and speculative, the goal shielding versus shifting dilemma (or the tension between flexibility and rigidity, as outlined in the previous section) relates to the exploration-exploitation trade-off during RL, as reviewed in the section entitled "Exploration-exploitation trade-off" here above. Accordingly, learning from positive and negative reinforcers would be a logical process to explore for its propensity to be molded by positive (or negative) mood. Exploitation of pre-existing or acquired knowledge about reward or punishment can be seen as goal shielding, or using the best known option to increase gain, while exploration of (new) incentive cues available in the environment can be seen as flexible shifting and updating. Both flexibility (Bless et al., 1996) and exploration (Humphries et al., 2012) have been associated with safe environments and abundance of reward, or internal states that signal safety. Learning new S-R contingencies puts high demands on processes such as cognitive control. However, an additional component in learning is the necessity to keep tracking the changes in these contingencies, not just for the momentary benefit, but because they serve as potent learning signals, they inform about the future possibility to obtain reward (or avoid loss or punishment). That is why in computational learning models, it is possible to extract parameters such as the learning rate- the rate at which the information about reinforces is updated, or exploration- a complementing parameter allowing to see to what extent does the change of the currently known values actually influence the next choice. Possibly, adding the knowledge of the current motivational and emotional state of the actor (i.e., the learning agent) would be a way to increase the predictive capacities of these learning models.

For example, Beeler (2012) suggested that dopamine has a role of energizing behavior, increasing the amount of effort an animal will allocate toward the goal pursuit. In this view, increased dopamine does not enhance the incentive for reward, but rather diminishes sensitivity to cost associated with procuring reward (mice with elevated dopamine showed less pairing between behavioral choices and reward history). So the tonic dopamine regulates the expression of prior reward learning by controlling the degree to which learned reward values will bias the current behavioral choice.

A related example in the context of a different neuromodulator has been given by the "affective keyboard mechanism" of NAc functioning by Berridge and Kringelbach (2013), where they proposed that different environmental settings (stressful vs. calm one) can exert influence on NAc that will produce different combinations of approach/avoid tendencies, and the motivation to move in different direction, depending on the emotional environment.

Bearing in mind the restrictions that previous work has already made adamantly clear, that positive mood is involving more than just a change in dopamine (Ashby et al., 1999) as well as that mesencephalic dopamine alternations are not merely changing the hedonic aspect of the reward, but the encoding of the value of reward (Ashby et al., 1999; Berridge & Kringelbach, 2013), we might nevertheless reasonably expect that "soaking" or immersing the learning system/agent in certain mood state will lead to changes in the quantity and/or quality of learning achieved (especially when it is guided by the processing or reward or punishment), that can eventually be observed by using electrophysiological measures such as the ERN and FRN component, for example. These two ERP components are hallmarks of the RL theory, and are supposed to be sensitive not only to error salience and expectedness (Holroyd & Coles, 2002), but also to the motivational importance of errors or negative feedback at a given moment in time during learning for the agent (Berridge, 2007; Hajcak & Foti, 2008; Salamone et al., 2007). Surprisingly enough, to the best of our knowledge, such systematic

investigation has not been conducted yet in the (psychophysiology oriented) literature. A similar idea has, however, already been visited in relation to negative mood. Shackman & Cavanagh (2014) suggested that negative affect and pain tend to engage the same processes described by theories of cognitive control in order to solve conceptually similar problems, which is what these authors call adaptive control, rather than cognitive control, to underscore its broader contribution to negative affect and nociception. Depression research has been particularly fruitful in gaining new insights into this issue. Earlier studies already showed increased FRN amplitude for negative feedback in patients suffering from Major Depressive Disorder (MDD) (Mueller, Pechtel, Cohen, Douglas, & Pizzagalli, 2015). Moreover, in the same study, a negative relationship between levels of Anhedonia and activation within the anterior portion of the MCC to negative feedback was reported, suggesting that Anhedonia might decrease sensitivity to reward prediction error in this brain region. Additionally, it was found that in the group of MDD patients, there was a basal ganglia dysfunction that led to decrease in consummatory aspect of reward processing (Pizzagalli et al., 2009). Another study found that MDD patients actually displayed better avoidance learning (than controls), accompanied with increased theta-band activity compared to these controls (Cavanagh, Bismark, Frank, & Allen, 2011)

In sum, although positive mood and dopamine increase may share some common variance or (anatomical) ground but are not the same process, we could reasonably expect that inducing positive mood in healthy adult participants (or removing it in the case of depression or Anhedonia in clinical/psychiatric patients) should introduce “visible” changes in the exploration-exploitation trade-off during RL. In accordance with Beeler’s proposition (Beeler, 2012), we could even then formulate the prediction that happy participants would more readily and intensely engage in free exploration during RL, compared to sad or neutral participants, an effect that could translate into an augmentation of beta (i.e., the computational

modeling parameter related indirectly to exploration). We might also surmise that this mood change could lead to changes in the learning rate because it could alter the way these participants could update their knowledge about the contingencies. However, whether such a change would be general or instead valence-specific (i.e., only the learning rate following positive/rewarding feedback could be increased after the induction of positive mood, mimicking somehow a mood-congruency effect; see also Frank et al., 2005) awaits further empirical validation. The present thesis was set out as one of the first attempts to address this question.

4. Outline of the present dissertation

In the previous sections, we have tried to review separate or disparate lines of research (RL, mood/affect and electrophysiology) and emphasized the need to bring them together in order to better understand mechanisms of RL in humans. The goal of this thesis is precisely that one. In doing this, we also hope to contribute to a growing literature or trend in the field that acknowledges the importance of positive emotions (and therefore positive mood), at multiple levels and in different domains. To illustrate this, we could refer to one of the recent United Nations (UN) initiatives that since the year 2011 proposed the usage and quantification of a so-called “happiness index” which is one of the macroscopic measures reflecting the global development of a country, besides what the gross domestic product is for economy, for example. Presumably, a better understanding of how learning is accomplished when the agent is in a specific mood state could help gather new fundamental knowledge about learning and positive emotions, possibly with direct implications for health and well-being.

In Chapter 1, we conducted a first set of studies with the aim to compare different mood induction techniques for their propensity to yield positive changes in the affective state of the participants, but also to assess whether RL was then influenced by the mood elicited and sustained, in a systematic way. To this aim, we used the probabilistic learning task (see Eppinger et al., 2008) described earlier in this chapter. We compared the use of movie clips to guided imagery as mood induction technique, using a between-subjects design. While movie clips may be considered as a standard, and efficient procedure, which are very often used for mood induction purposes (see Gray, Braver, & Raichle, 2002; Nadler, Rabi, & Minda, 2010; Tice, Baumeister, Shmueli, & Muraven, 2007), we were interested in comparing its efficacy with a promising alternative provided by guided imagery. The bioinformational theory of imagery (Lang, 1979) proposes that especially for negatively-valenced events, textual material can serve as a cue to prompt psychophysiological activation similar to the actual experience of this event. It stipulates that the motor control system is active during affective imagery. An fMRI study confirmed this hypothesis and showed activation in the supplementary motor area, prefrontal cortex, and the cerebellum for arousing emotional memories, both positive and negative (Sabatinelli, Lang, Bradley, & Flaisch, 2006). Interestingly, another study showed strong activation in the nucleus accumbens and the medial prefrontal cortex exclusively related to positive imagery, while both pleasant and unpleasant imagery was related to strong activation in the amygdala (Costa, Lang, Sabatinelli, Versace, & Bradley, 2010). We capitalized on a previously developed theory of guided imagery put forward by Emily Holmes (Holmes, Coughtrey, & Connor, 2008; Holmes, Mathews, Dalgleish, & Mackintosh, 2006; Holmes & Mathews, 2010), which is characterized by using the “field perspective”, or imagining from a very specific, personal perspective. The idiosyncrasy of positive mood is what separates it from a more generic kind of positive

feeling produced by a funny movie clip, and what likely gives it strengths to produce more affective (and possibly potent) reactions, and physical sensations (McIsaac & Eich, 2002).

In Chapter 2, we set out to compare mechanisms of RL at multiple levels in participants with either a positive or a neutral mood, using the knowledge gathered in Chapter 1 regarding the use of an efficient and valid mood induction procedure for this aim. We used the computational model of Sutton and Barto with a softmax rule to extract the learning rate and exploration, besides standard measures of learning (accuracy and speed). The learning rate parameter was tracked separately for the positive and the negative feedback, as we were particularly interested in examining whether the potential effect of positive mood on RL could obey a valence-specific principle, such that positive mood would only alter/augment the relationship towards reward (i.e., positive feedback; akin to what was shown previously in the literature by Frank et al., 2005), or instead, it would change the way the learning history is updated in an undifferentiated (i.e., valence-unspecific) manner. Additionally, we were interested in checking if positive mood, compared to a control neutral mood condition (with an active induction phase), could tip the balance between exploration and exploitation. If positive mood can be viewed as a tonic (dopamine-related) change in the system (Aston-Jones & Cohen, 2005; Beeler, 2012; Berridge & Kringelbach, 2013), then we could expect this specific mood state to be associated with more exploration than the neutral mood state. If the learning rate indeed reflects a weighting factor that traces the changes in the environment (Ullsperger et al., 2014), then it should be influenced by mood factors. These measures of RL were used in combination with the ERN and the FRN ERP component, which, as we have reviewed here above in the previous sections, belong to a family of early fronto-central negativities that reflect learning or "alarm signals" that timely inform the organism about the need to change or adapt behavior accordingly in case worse than expected events (such as response errors or negative evaluative feedback) are detected. Although the existing literature

did not allow us to formulate clear, specific or directional predictions for modulatory effects of positive mood on RL at these different levels, we nevertheless did speculate on possible effects that could result from the induction of positive mood during RL. In a nutshell, we reckoned that positive mood could perhaps influence RL in a condition-specific way, increasing reward prediction error or sensitivity in conditions where the prospect of getting a reward was high (i.e., in the deterministic condition selectively), with effects measurable potentially either at the ERN or FRN level.

In the following study (Chapter 3), we tested a new prediction arising from the new findings collected in the previous chapter (Chapter 2). More specifically, we sought to test whether positive mood could foster exploitation of acquired values or knowledge during RL and in turn increase/change learning when it becomes less dependent on external information (feedback). To this aim, we removed the evaluative (visual) feedback provided after each and every decision, once learning was established properly, thereby blocking or neutralizing exploration and favoring exploitation or internalization. Additionally, other than contrasting happy to neutral mood, in this study, we added a third group which received a sad mood induction by using the same guided imagery technique as used and validated in the two previous chapters. By introducing a sad mood condition, we could formally or more directly test the presence of valence effects during RL: would happy and sad mood show changes in RL occurring in different directions (when the learning rate and/or exploration were considered), or alternatively, would they both have a similar impact on RL (when titrated in terms of strength and strategy), relative to the neutral mood group/condition? These questions lay at the heart of Chapter 3.

Finally, in the last study presented in this thesis (Chapter 4), we adopted a different perspective to study effects of mood on RL. There, we studied a large cohort of treatment resistant unipolar MDD patients enrolled in a treatment study, characterized by low positive

affect (and Anhedonia) and high/excessive negative affect. Hence, instead of inducing or injecting a specific affective state (like in the previous chapters), we looked here at a unique condition where mood was abnormal from the outset (i.e. lacking positive affect) to see if RL could be influenced in this specific psychopathological condition, relative to a group of healthy control subjects (matched in age and sex with the MDD patients). Like what was carried out in Chapter 2, changes in RL as a function of MDD were carefully explored at multiple levels in this study (behavioral, computational modeling and EEG/ERP). We tested the prediction that MDD could be accompanied by abnormal ERN components, in keeping with previous models and findings in the literature (Olvet, Klein, & Hajcak, 2010). Moreover, we assessed whether Anhedonia, more than depression per se, could lead to a lower learning rate following positive feedback in these patients (Pizzagalli, 2014).

As all of these empirical chapters were written up as individual studies, the final chapter (General Discussion) of the thesis will be an integration and a thorough discussion of these new findings. There, we will review the implications of these new results regarding modulatory effects of RL mechanisms by (either positive or negative) mood, before outlining new avenues for future research, as well as unsolved issues that remain to be addressed.

5. REFERENCES

Allain, S., Carbonnell, L., Falkenstein, M., Burle, B., & Vidal, F. (2004). The modulation of the Ne-like wave on correct responses foreshadows errors. *Neuroscience Letters*, *372*(1-2), 161–166.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. *Arlington*. American Psychiatric Publishing. doi:10.1176/appi.books.9780890425596.744053

Anokhin, A. P., Golosheykin, S., & Heath, A. C. (2008). Heritability of frontal brain function related to action monitoring. *Psychophysiology*, *45*(4), 524–534.

Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529–550. Retrieved from <Go to ISI>://000082133000004

Ashby, Valentin, & Turken. (2002). The effects of positive affect and arousal on working memory and executive attention Neurobiology and computational models. *Behaviour*, *16*(3), 245–287.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Barrett, L. F. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition & Emotion*.

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology*, *42*(1), 33–42.

Bediou, B., Koban, L., Rosset, S., Pourtois, G., & Sander, D. (2012). Delayed monitoring of accuracy errors compared to commission errors in ACC. *NeuroImage*, *60*(4), 1925–1936.

Beedie, C., Terry, P., & Lane, A. (2005). Distinctions between emotion and mood. *Cognition & Emotion*.

Beeler, J. a. (2012). Thorndike's law 2.0: Dopamine and the regulation of thrift. *Frontiers in Neuroscience*, *6*(AUG), 1–12. doi:10.3389/fnins.2012.00116

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. (1995). "Where did I go wrong?" A psychophysiological analysis of error detection. *Journal of Experimental Psychology. Human Perception and Performance*, *21*(6), 1312–1322.

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology and Behavior*.

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

Berridge, K. C., & Kringelbach, M. L. (2013). Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology*, *23*(3), 294–303. doi:10.1016/j.conb.2013.01.017

Bhanji, J. P., Beer, J. S., & Bunge, S. A. (2010). Taking a gamble or playing by the rules: dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. *NeuroImage*, *49*(2), 1810–9. doi:10.1016/j.neuroimage.2009.09.030

Biederman, I., & Vessel, E. A. (2006). Perceptual Pleasure and the Brain and seeks it through the senses, 247–254.

Bless, H., Schwarz, N., Clore, G. L., Golisano, V., Rabe, C., & Wölk, M. (1996). *Mood and the use of scripts: does a happy mood really lead to mindlessness? Journal of personality and social psychology* (Vol. 71).

Bolte, A., & Goschke, T. (2010). Towards a Theory of Thinking, 261–277. doi:10.1007/978-3-642-03129-8

Botvinick, M., & Braver, T. (2015). Motivation and Cognitive Control : From Behavior to Neural Mechanism. *Annual Review of Psychology*. doi:10.1146/annurev-psych-010814-015044

Braem, S., King, J. a, Korb, F. M., Krebs, R. M., Notebaert, W., & Egner, T. (2013). Affective modulation of cognitive control is determined by performance-contingency and mediated by ventromedial prefrontal and cingulate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(43), 16961–70. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3807025&tool=pmcentrez&rendertype

Braem, S., Verguts, T., Roggeman, C., & Notebaert, W. (2012). Reward modulates adaptations to conflict. *Cognition*, *125*(2), 324–332.

Browning, M., Behrens, T. E., Jocham, G., Reilly, J. X. O., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments Michael Browning, (August 2014), 1–50. doi:10.1038/nn.3961

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, *20*(9), 1637–1655.

Carbonnell, L., & Falkenstein. (2006). Does the error negativity reflect the degree of response conflict? *Brain Research*, *1095*(1), 124–130. doi:10.1016/j.brainres.2006.04.004

Cavanagh, Bismark, Frank, & Allen. (2011). Larger Error Signals in Major Depression are Associated with Better Avoidance Learning. *Front Psychol*, *2*, 331. doi:10.3389/fpsyg.2011.00331

Cavanagh, J. F., & Shackman, A. J. (2014). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology Paris*. doi:10.1016/j.jphysparis.2014.04.003

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

Chase, Frank, Michael, Bullmore, Shakian, & Robbins. (2010). Approach and avoidance learning in patients with major depression and healthy controls : Relations to anhedonia. *Psychol Med*, *40*, 433–440.

Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 509–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24733296

Chiu, P. H., & Deldin, P. J. (2007). Neural evidence for enhanced error detection in major depressive disorder. *The American Journal of Psychiatry*, *164*(4), 608–616. doi:10.1176/appi.ajp.164.4.608

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. a, & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–9. doi:10.1038/sj.npp.1301153

Costa, V. D., Lang, P. J., Sabatinelli, D., Versace, F., & Bradley, M. M. (2010). Emotional imagery: Assessing pleasure and arousal in the brain's reward circuitry. *Human Brain Mapping*, *31*(9), 1446–1457.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*(1-2), 25–62.

Darwin, C., & Ekman, P. (1872). *The expression of the emotions in man and animals (3rd ed.). The expression of the emotions in man and animals 3rd ed*. doi:10.1037/10001-000

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), 876–9. doi:10.1038/nature04766

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303–305. doi:10.1111/j.1467-9280.1994.tb00630.x

Disner, S. G., Beevers, C. G., Haigh, E. a P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews. Neuroscience*, *12*(8), 467–77. doi:10.1038/nrn3027

Dreisbach, G. (2006). How positive affect modulates cognitive control: the costs and benefits of reduced maintenance capability. *Brain and Cognition*, *60*(1), 11–9. doi:10.1016/j.bandc.2005.08.003

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*.

Easdon, C., Izenberg, A., Armilio, M. L., Yu, H., & Alain, C. (2005). Alcohol consumption impairs stimulus- and error-related processing during a Go/No-Go Task. *Brain Research. Cognitive Brain Research*, *25*(3), 873–883. doi:10.1016/j.cogbrainres.2005.09.009

Elliott, R., Sahakian, B. J., McKay, A. P., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1996). Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. *Psychological Medicine*, *26*(5), 975–989.

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*(6), 1714–1720.

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

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*.

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

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, *51*(2-3), 87–107. doi:10.1016/S0301-0511(99)00031-9

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The Processing of Unexpected Positive Response Outcomes in the Mediofrontal Cortex. *Journal of Neuroscience*.

Fiedler, K., Martin, L. L., & Clore, G. L. (2001). Affective states trigger processes of assimilation and accommodation. In *Theories of mood and cognition: A user's guidebook*. (pp. 85–98). Lawrence Erlbaum Associates Publishers. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2001-16267-004&site=ehost-live

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902. doi:10.1126/science.1077349

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*(8), 1062–8. doi:10.1038/nn.2342

Frank, Woroch, & Curran. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. doi:10.1016/j.neuron.2005.06.020

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci*, *359*(1449), 1367–1378. doi:10.1098/rstb.2004.1512

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*.

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, 530–547. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24659000

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: mood and global versus local processing of visual information. *Psychological Science*, *13*(1), 34–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11892776

Gehring, Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error-Detection and Compensation. *Psychological Science*, *4*(6), 385–390. Retrieved from <Go to ISI>://A1993MF61300008

Gehring, Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science : A Journal of the American Psychological Society / APS*, *11*(1), 1–6.

Gehring, W. J., Goss, B., & Coles, M. G. H. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390. doi:10.1111/j.1467-9280.1993.tb00586.x

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

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *NeuroImage*, *47*(4), 2023–2030. doi:10.1016/j.neuroimage.2009.05.064

Goschke, T., & Bolte, A. (2013). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*.

Goschke, T., & Bolte, A. (2014). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*, *62*, 403–423. doi:10.1016/j.neuropsychologia.2014.07.015

Gray, J. R., Braver, T. S., & Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 4115–20. doi:10.1073/pnas.062381899

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*.

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity: Research report. *Psychological Science*, *19*(2), 103–108.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. In *Psychophysiology* (Vol. 40, pp. 895–903).

Hajcak, G., McDonald, N., & Simons, R. F. (2004). *Error-related psychophysiology and negative affect. Brain and cognition* (Vol. 56). doi:10.1016/j.bandc.2003.11.001

Hajcak, G., Nieuwenhuis, S., Ridderinkhof, K. R., & Simons, R. F. (2005). Error-preceding brain activity: Robustness, temporal dynamics, and boundary conditions. *Biological Psychology*, *70*(2), 67–78.

Harmon-Jones, E., Gable, P. a, & Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. *Biological Psychology*, *84*(3), 451–62. doi:10.1016/j.biopsycho.2009.08.010

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2013). Does Negative Affect Always Narrow and Positive Affect Always Broaden the Mind? Considering the Influence of motivational Intensity on Cognitive Scope. *Current Directions in Psychological Science*, *22*(4),301–307.Retrievedfrom http://cdp.sagepub.com/content/22/4/301\nhttp://cdp.sagepub.com/content/22/4/301.short

Hogan, A. M., Vargha-Khadem, F., Saunders, D. E., Kirkham, F. J., & Baldeweg, T. (2006). Impact of frontal white matter lesions on performance monitoring: ERP evidence for cortical disconnection. *Brain*, *129*(8), 2177–2188. doi:10.1093/brain/awl160

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, *1*(4), 304–9. doi:10.1038/1124

Holmes, E. A., Coughtrey, A. E., & Connor, A. (2008). Looking at or through rose-tinted glasses? Imagery perspective and positive mood. *Emotion (Washington, D.C.)*, *8*(6), 875–879. doi:10.1037/a0013617

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). *Positive interpretation training: effects of mental imagery versus verbal training on positive mood. Behavior therapy* (Vol. 37). doi:10.1016/j.beth.2006.02.002

Holmes, & Mathews. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349–362. doi:10.1016/j.cpr.2010.01.001

Holmes, & Pizzagalli, D. (2008). Spatiotemporal dynamics of error processing dysfunctions in major depressive disorder. *Archives of General Psychiatry*, *65*(2), 179–88. doi:10.1001/archgenpsychiatry.2007.19

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Humphries, M. D., Khamassi, M., & Gurney, K. (2012). Dopaminergic Control of the Exploration-Exploitation Trade-Off via the Basal Ganglia. *Frontiers in Neuroscience*, *6*, 9. doi:10.3389/fnins.2012.00009

Huntsinger, J. R. (2012). Does positive affect broaden and negative affect narrow attentional scope? A new answer to an old question. *Journal of Experimental Psychology: General*, *141*(4), 595–600.

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, 1–7. doi:10.1016/j.tics.2013.12.009

Isen, Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*(6), 1122–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3598858

Isen, Johnson, M. M., Mertz, E., & Robinson, G. F. (1985). The influence of positive affect on the unusualness of word associations. *Journal of Personality and Social Psychology*, *48*(6), 1413–1426.

Isen, & Means, B. (1983). The Influence of Positive Affect on Decision-Making Strategy. *Social Cognition*. doi:10.1521/soco.1983.2.1.18

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–96. doi:10.1162/jocn.2010.21548

Kahn, B., & Isen, A. M. (1993). Variety seeking among safe, enjoyable products. *Journal of Consumer Research*, *20*, 257–270.

Kemper, T. D., & Lazarus, R. S. (1992). Emotion and Adaptation. *Contemporary Sociology*.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–947. doi:10.1038/nn1724

Knutson, B., Katovich, K., & Suri, G. (2014). Inferring affect from fMRI data. *Trends in Cognitive Sciences*, *18*(8), 422–428. doi:10.1016/j.tics.2014.04.006

Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: An integrative review. *Neuroscience and Biobehavioral Reviews*, 1–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24681006

Kring, A. M., & Barch, D. M. (2014). The motivation and pleasure dimension of negative symptoms: neural substrates and behavioral outputs. *European Neuropsychopharmacology : The Journal of the European College of Neuropsychopharmacology*, *24*(5), 725–36. doi:10.1016/j.euroneuro.2013.06.007

Lang. (1979). Presidential address, 1978. A bio-informational theory of emotional imagery. *Psychophysiology*, *16*(6), 495–512.

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, *84*(3), 437–450.

Lange, S., Leue, A., & Beauducel, A. (2012). Behavioral approach and reward processing: Results on feedback-related negativity and P3 component. *Biological Psychology*, *89*(2), 416–425.

Larsen, R. J. (2000). Toward a Science of Mood Regulation. *Psychological Inquiry*.

Luck, S. (2005). *An Introduction to the event-related potential technique*. MIT Press.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology. General*, *129*(1), 43–60.

Lyness, W. H., Friedle, N. M., & Moore, K. E. (1979). Destruction of dopaminergic nerve terminals in nucleus accumbens: Effect on d-amphetamine self-administration. *Pharmacology Biochemistry and Behavior*, *11*(5), 553–556.

Madan, C. R. (2013). Toward a common theory for learning from reward, affect, and motivation: the SIMON framework. *Frontiers in Systems Neuroscience*, *7*(October), 59. doi:10.3389/fnsys.2013.00059

Marien, H., Aarts, H., & Custers, R. (2012). Being flexible or rigid in goal-directed behavior: When positive affect implicitly motivates the pursuit of goals or means. *Journal of Experimental Social Psychology*, *48*(1), 277–283. doi:10.1016/j.jesp.2011.08.013

McIsaac, H. K., & Eich, E. (2002). Vantage point in episodic memory. *Psychonomic Bulletin & Review*, *9*(1), 146–150.

Mies, G. W., van der Veen, F. M., Tulen, J. H. M., Birkenhäger, T. K., Hengeveld, M. W., & van der Molen, M. W. (2011). Drug-free patients with major depression show an increased electrophysiological response to valid and invalid feedback. *Psychological Medicine*, *41*(12), 2515–25. doi:10.1017/S0033291711000778

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

Mueller, E. M., Pechtel, P., Cohen, A. L., Douglas, S. R., & Pizzagalli, D. A. (2015). Potentiated processing of negative feedback in depression is attentuated by anhedonia. *Depression and Anxiety*. doi:10.1002/da.22338

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better mood and better performance. Learning rule-described categories is enhanced by positive mood. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(12), 1770–1776. doi:10.1177/0956797610387441

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

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clin Psychol Rev*, *28*(8), 1343–1354. doi:10.1016/j.cpr.2008.07.003

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clinical Psychology Review*, *28*(8), 1343–54. doi:10.1016/j.cpr.2008.07.003

Olvet, D. M., Klein, D. N., & Hajcak, G. (2010). Depression symptom severity and error-related brain activity. *Psychiatry Research*, *179*(1), 30–37.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews. Neuroscience*, *9*(2), 148–158.

Phillips, L. H., Bull, R., Adams, E., & Fraser, L. (2002). Positive mood and executive function: evidence from stroop and fluency tasks. *Emotion (Washington, D.C.)*, *2*(1), 12–22.

Pizzagalli. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, (January), 1–31. doi:10.1146/annurev-clinpsy-050212-185606

Pizzagalli, Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2009). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87.

Rabbit, P. M. A. (1981). Sequential reactions. In H. D (Ed.), *Human skills* (pp. 153–175). New York: Wiley.

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*(2), 264–272.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556. doi:10.1038/nrn2357

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II Current Research and Theory* (Vol. 21, pp. 64–99). Appleton-Century-Crofts. Retrieved from http://homepage.mac.com/sanagnos/rescorlawagner1972.pdf

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. H. (2002). *Alcohol consumption impairs detection of performance errors in mediofrontal cortex. Science (New York, N.Y.)* (Vol. 298).

Ridderinkhof, K. R., Nieuwenhuis, S., & Bashore, T. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*.

Roger, C., Bénar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral Cingulate Zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *NeuroImage*, *51*(1), 391–403.

Rowe, G., Hirsh, J. B., & Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(1), 383–388.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of mind: event-related brain potentials and cognition*. Oxford, England: Oxford University Press.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*.

Sabatinelli, D., Lang, P. J., Bradley, M. M., & Flaisch, T. (2006). Chapter 5 The neural basis of narrative imagery: emotion and action. *Progress in Brain Research*.

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*.

Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(1), 141–151.

Scheffers, M. K., Coles, M. G. H., Bernstein, P., Gehring, W. J., & Donchin, E. (1996). Event-related brain potentials and error-related processing: An analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, *33*(1), 42–53.

Scherg, M. (1990). Fundamentals of dipole source potential analysis. In F. Grandori, M. Hoke, & G. L. Romani (Eds.), *Auditory evoked magnetic fields and electric potentials* (pp. 40–69). Advances in Audiology, Vol.6.Karger, Basel.

Scherg, M., & Picton, T. W. (1991). Separation and identification of event-related potential components by brain electric source analysis. *Electroencephalography and Clinical Neurophysiology. Supplement*, *42*, 24–37.

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

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

Schwager, S., & Rothermund, K. (2013). Counter-regulation triggered by emotions: positive/negative affective states elicit opposite valence biases in affective processing. *Cognition & Emotion*, *27*(5), 839–855.

Shackman, A. J., Salomons, T. V, Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews. Neuroscience*, *12*(3), 154–167.

Sutton, & Barto. (1990). Time-Derivative Models of Pavlovian Reinforcement. In M. Gabriel & J. Moore (Eds.), Learning and Computational Neuroscience: Foundations of Adaptive Networks (pp. 497–537). MIT Press.

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: an introduction.* Cambridge, MA: MIT Press.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies.* New York: Macmillan.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology*, *43*(3), 379–384. doi:DOI 10.1016/j.jesp.2006.05.007

Tooby, J., & Cosmides, L. (1990). The Past Explains the Present Emotional Adaptations and the Structure of Ancestral Environments. *Ethology and Sociobiology*, *11*, 375–424.

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

Unger, K., Heintz, S., & Kray, J. (2012). Punishment sensitivity modulates the processing of negative feedback but not error-induced learning. *Frontiers in Human Neuroscience*, *6*(June), 1–16. doi:10.3389/fnhum.2012.00186

Unger, K., Kray, J., & Mecklinger, A. (2012). Worse than feared? Failure induction modulates the electrophysiological signature of error monitoring during subsequent learning. *Cognitive, Affective, & Behavioral Neuroscience*.

Ursu, S., Stenger, V. A., Shear, M. K., Jones, M. R., & Carter, C. S. (2003). Overactive Action Monitoring in Obsessive-Compulsive Disorder: Evidence From Functional Magnetic Resonance Imaging. *Psychological Science*. doi:10.1111/1467-9280.24411

Van Steenbergen, H., Band, G. P. H., & Hommel, B. (2010). In the mood for adaptation: how affect regulates conflict-driven control. *Psychological Science*, *21*(11), 1629–34. doi:10.1177/0956797610385951

Vanderhasselt, M.-A., De Raedt, R., De Paepe, A., Aarts, K., Otte, G., Van Dorpe, J., & Pourtois, G. (2014). Abnormal proactive and reactive cognitive control during conflict processing in major depression. *Journal of Abnormal Psychology*, *123*(1), 68–80. doi:10.1037/a0035816

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2013). Positive emotion broadens attention focus through decreased position-specific spatial encoding in early visual cortex: Evidence from ERPs. *Cognitive, Affective, & Behavioral Neuroscience*. doi:10.3758/s13415-012-0130-x

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2014). Feeling happy enhances early spatial encoding of peripheral information automatically: electrophysiological time-course and neural sources. *Cognitive, Affective & Behavioral Neuroscience*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24570275

Vassena, E., Silvetti, M., Boehler, C. N., Achten, E., Fias, W., & Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS ONE*, *9*(3).

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: A reexamination of available data. *Biological Psychology*, *64*(3), 265–282.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the "error negativity" specific to errors? *Biological Psychology*, *51*(2-3), 109–128.

Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, *46*(10), 2545–2555. doi:10.1016/j.neuropsychologia.2008.04.006

Weinberg, A., Riesel, A., & Hajcak, G. (2011). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. *Motivation and Emotion*, *36*(1), 84–100. doi:10.1007/s11031-011-9269-y

Wise, R. A., Spindler, J., deWit, H., & Gerberg, G. J. (1978). Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. *Science (New York, N.Y.)*, *201*(4352), 262–264.

Woergoetter, F., & Porr, B. (2008). Reinforcement learning. *Scholarpedia*. doi:10.4249/scholarpedia.1448

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The Neural Basis of Error Detection: Conflict Monitoring and the Error-Related Negativity. *Psychological Review*, *111*(4), 931–959. doi:10.1037/0033-295X.111.4.931

Yue, X., Vessel, E. A., & Biederman, I. (2007). The neural basis of scene preferences. Neuroreport, 18(6), 525–529.

Zwosta, K., Hommel, B., Goschke, T., & Fischer, R. (2013). Mood states determine the degree of task shielding in dual-task performance. *Cognition & Emotion*, *27*(6), 1142–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23438389

CHAPTER 2

Ways of being happy: differential effects of guided imagery and movie clips on happiness levels and reinforcement learning[1]

Unlike negative mood, positive affect is thought to yield a wide range of beneficial effects for health and well-being, including a higher flexibility. This (cognitive) gain might in turn improve learning, especially when this process requires to switch (readily and frequently) back and forth between exploration and exploitation, as evidenced during probabilistic learning for example. In this study, we tested this prediction. Participants received a mood induction procedure (MIP) to elicit either happy or neutral mood. This MIP was based on either guided imagery or movie clips. Then, they completed a standard probabilistic learning task. Results showed that while the MIP was successful to increase and maintain levels of happiness or pleasantness (equally so with these two different MIPs), positive mood was not associated with a specific gain or change during probabilistic learning compared to neutral mood. Additional analyses confirmed that not only the amount of learning (quantitative effects), but also its specific unfolding over time (qualitative effects) remained immune to the changes in the affective state of the participant brought about by the MIP. We discuss these new findings in light of the affective science literature that posits general gains in flexibility or cognitive control following the induction or experience of positive affect/mood.

[1]Bakic, J., & Pourtois, G., (2012). Ways of being happy: differential effects of guided imagery and movie clips on happiness levels and reinforcement learning.

1. INTRODUCTION

Without doubt, the way we feel profoundly influences the way we think and reason. In their extensive review, Lyubomirsky, King and Diener (2005) reported significant associations between the experience of (mild) long lasting positive affect and better work performance, higher job satisfaction, increased marital happiness, less loneliness, and better health, amongst others. Even though these associations did not inform about the causal (beneficial) role of positive affect on cognition, independent evidence and models available in the affective sciences literature argue for such a direct link, and the enhanced protective function of positive emotion for health and well-being (and hence better life outcomes). For example, positive mood might increase sensitivity for rewards in the environment, potentially increasing the chance to approach or discover new opportunities (Gray, 1994). In their influential theory, Ashby, Isen and Turken (1999) proposed that positive affect enhances cognitive flexibility, thus facilitating creativity and switching between different cognitive mind sets. This association between positive emotion and an enhanced cognitive flexibility or creativity also lies at the heart of the "broaden and build" theory (Fredrickson, 2001, 2004; Garland et al., 2010), where different emotions/valences are thought to serve distinctive adaptive functions. While negative emotions signal imminent threat and the need to mobilize (mental) resources ("narrowing" effect), positive emotions have the opposite effect, they broaden attention. More specifically, they are signaling safety, thus creating an opportunity to flexibly consider different response options, and build additional resources that can later be used to fight against deleterious effects of stress, and in turn increase resilience. In their functional classification of positive emotion, Shiota, et al. (2014) postulate that positive emotions foster play, a creative process that enables to practice complex skills and boost cognitive flexibility (in a safe situation). However, few studies have actually confirmed at the empirical level clear links between positive emotion and an enhanced cognitive flexibility or

creativity (Isen et al., 1987; Isen, Johnson, Mertz, & Robinson, 1985; Isen, 1984), or a putative broader attention focus (Bruyneel et al., 2013; Rowe, Hirsh, & Anderson, 2007; Vanlessen, Rossi, De Raedt, & Pourtois, 2013, 2014). Moreover, other theoretical accounts posit that (positive) emotions actually compete with cognition for limited resources or attention capacities, and thus are not always or necessarily beneficial to (cognitive) flexibility (Ellis & Ashbrook, 1988; Pessoa, 2008).

Presumably, (cognitive) flexibility is multifaceted and not only at play during complex reasoning or creativity tasks (such as the AUT or RAT; see Guilford, 1967; Mednick, 1962), but also during learning where new associations have to be formed, updated and sometimes later altered depending on their actual validity or accuracy. Accordingly, it might be that positive emotion is beneficial to specific aspects or components of flexibility (and thus learning), instead of creating a sort of uniform and unspecific change or boost in creativity or attention broadly speaking (Ashby et al., 1999; Fredrickson, 2001), even though very few studies have actually corroborated this conjecture. Consistent with this assumption, Nadler, Rabi and Minda, (2010) recently found that positive, compared to neutral or negative mood, increased cognitive flexibility if and only if the task required hypothesis testing and rule selection.

The main goal of our study was to extend these earlier findings and test whether positive mood (when compared to a control mood condition dominated by neutral emotional content) might influence learning, when it is probabilistic in nature and requires therefore a high amount of (cognitive) flexibility regarding hypothesis testing and rule selection. Probabilistic learning offers a valuable means to test our prediction given that it requires to explore new associations (usually linking specific visual stimuli to manual responses), as well as exploit knowledge already acquired during the preceding encounters, such as to maximize reward (Cohen, McClure, & Yu, 2007). The accuracy of the (current) behavior is typically

signaled by means of a visual feedback (immediately following the action) in probabilistic learning tasks. Crucially, optimal behavior requires to flexibly switch over time between these two opposing modes of learning, or forces, exploration vs. exploitation (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Jepma & Nieuwenhuis, 2011). For this purpose, we used a probabilistic learning task previously validated in the literature (Eppinger et al., 2008). We chose this specific task setting (see also Frank, Woroch, & Curran, 2005; (Bakic, Jepma, De Raedt, & Pourtois, 2014) because it allowed us to explore and characterize mood-related changes during learning across different conditions varying in reward probability and shown in random order. In this context, the use of a flexible cognitive mode or style is therefore desirable and hence we surmised that if positive mood influences cognitive flexibility, it should also be associated with possible gains or benefits at the behavioral level during probabilistic learning.

A main challenge or caveat in research on positive emotion is about what is actually entailed by positive emotion, as well as how it is typically induced in healthy adult participants and hence operationalized at the methodological level. Usually, positive mood induction is achieved by means of watching (funny) movie clips, or receiving (unexpected) rewards, gifts, presents or positive feedback. However, more recently, alternative mood induction procedures (MIPs) have been devised (including guided/mental imagery), which enable to fine-tune and individualize/personalize the change in the current affective state of the participant (see Holmes, Mathews, Dalgleish, & Mackintosh 2006; Holmes, Coughtrey, & Connor, 2008). According to Westermann, Stahl and Hesse (1996) who compared different MIPs, watching (funny) movie clips is not only the most commonly used method in the literature, but also the most efficient. On the other hand, more recent research (e.g., Nadler et al., 2010) actually suggests that movie clips mostly induce amusement, a (positive) emotion characterized as low in intensity and motivation, and lacking personal effort or investment.

However, Csikszentmihalyi (1999) cogently showed that an important aspect of the happy feeling might be this specific component of personal effort or investment. By definition, the level of involvement and engagement (or more generally "immersion") may vary substantially across participants when watching the same movie clips. Rottenberg, Ray and Gross (2007) analyzed several qualitative aspects of movies as mood induction and concluded that, in case experimenters want to achieve higher ecological validity and increase demand it is recommended to opt for relived emotion or scripted/unscripted social interaction procedures because they are more engaging than movies. Accordingly, the second aim of our study was to compare, using a between-subject design, two different MIPs (i.e., watching move clips vs. guided imagery) to assess if they could lead to a comparable increase at the subjective level in positive affect in healthy adult participants and in turn influence differently probabilistic learning. Presumably, if personal effort or investment in the elicitation of the positive emotion is an important factor accounting for the amount and quality of the experienced (positive) emotion at the subjective level (Csikszentmihalyi, 1999), then we hypothesized that the use of guided imagery might lead to stronger mood changes than watching move clips, and perhaps then, larger changes in probabilistic learning too. In this study, we directly put to the test theses specific predictions.

2. METHODS

We conducted two different experiments that had the same general structure (i.e., probabilistic learning task preceded by positive vs. neutral mood induction), but differed regarding the way positive (or neutral) mood was eventually induced in healthy adult participants. In Experiment 1, guided imagery was used for this purpose, while (funny) movie clips served as MIP in Experiment 2. None of the participants of Experiment 1 took part in Experiment 2, and vice versa.

2.1. Participants

*Experiment 1.*Forty four undergraduate students from Ghent University participated in this experiment in exchange for course credits or a financial compensation of 10 euros. They were randomly allocated to one of two groups: positive vs. neutral mood. They were all right-handed, with no past or current neurological or psychiatric problems, they had normal or corrected to normal vision, and all gave written informed consent prior to the start of the experiment. We used several exclusion criteria (see also Bakic et al., 2014). First, participants in the positive group had to show a marked increase in positive mood (a reliable increase in positive affect had to be present after the MIP, relative to the mood state measured at baseline, when entering the laboratory). Based on this criterion, the data of five participants were eventually removed from the analyses. In the neutral group, no substantial change in mood (increase or decrease) between baseline and post MIP had to be observed. The data of five participants were excluded to comply with this criterion. Second, participants showing no clear learning during the main task (i.e., their learning curves remained flat and they did not differ from chance level) were excluded as well (n=3). The final sample consisted of 31 participants (Mean age= 21.13 years, SD=1.80, 10 males), 16 in the positive and 15 in the neutral mood group. *Experiment 2.* Forty five undergraduate students from Gent University were recruited. Exclusion criteria were the same as in Experiment 1. In the positive mood group, the data of six participants were excluded due to a lack of change/increase in positive affect after the MIP relative to the baseline. Five other participants were excluded from the neutral group because their (positive) mood increased after the MIP. Finally, the data of five additional participants were excluded since they showed no learning during the main task. The final sample consisted of 29 participants (Mean age= 19.90 years, SD =2.09, 7 males), 16 in the positive, and 13 in the neutral mood group. These two experiments were approved by the local ethics committee.

2.2. Probabilistic learning task

We used a probabilistic learning task previously devised by Eppinger et al. (2008). In this task, participants were asked to decipher and learn, by trial and error, several hidden stimulus (S) – response (R) associations. For each trial, participants were asked to decide, with a constant speed limit, whether the stimulus shown on the screen was associated with either response 1 or 2 (corresponding to two pre-defined keys of a response box). Performance feedback (F) was given following each response made.

Participants were presented with 6 different visual stimuli (A-B-C-D-E-F), belonging to three conditions (unknown to the participants) that differed regarding the actual (reward) probability of the S-R mapping. In one condition, each stimulus of the pair was exclusively associated with one of the two response keys, corresponding to a "deterministic" S-R mapping (i.e., response 1 was always correct for stimulus A, while response 2 for stimulus B). In the probabilistic condition, stimulus C was associated 80% of the time with response 1 (and thus 20% of the time with the concurrent response 2), while stimulus D had a reverse S-R mapping. Finally, in the random condition, each stimulus of the pair was associated equally often to each of the two response keys (i.e., stimuli E and F were associated 50% of the time with response 1 and 50% of the time with response 2).

The visual stimuli were colorful line drawings (Rossion & Pourtois, 2004), presented against a white homogenous background on a 17-inch computer screen. These stimuli were visual objects belonging to different semantic categories (artefacts, buildings, musical instruments, clothes, vehicles, furniture).

The trial structure was as follows: it began with a fixation cross lasting 250 ms. Then, the visual stimulus was presented for 500 ms, followed by a blank screen of 300 ms. Response deadline was set to 800 ms following stimulus onset. Then, following the response,

500 ms elapsed (blank screen) before performance feedback was presented for 500 ms. This feedback was provided in the form of a written word (in Dutch) shown in black against a white homogenous background. This word was “goed” (correct), “fout” (incorrect), or “te traag” (too late). The inter-trial interval was constant (500 ms) and it corresponded to a blank screen, after which a new trial ensued. Responses (i.e., key presses) were recorded using a response box.

Each participant completed two blocks of 240 trials. Each block had a new set of six different stimuli, each repeated forty times, so that participants had to learn new S-R mappings in each block. Trial order within a block as well as the order of the two blocks were alternated across participants.

2.3. Procedure

The two experiments differed only regarding the type of the MIP used (either guided imagery – Experiment 1, or watching movie clips – Experiment 2). In order to get acquainted with the task, participants first completed a short practice session of 20 trials. Next, either a positive or neutral mood was induced by means of the MIP before the beginning of the first block. The same MIP was repeated (5 minutes) halfway (after 120 trials) during the first block. The same procedure was repeated for the following blocks (i.e., the MIP was used each time at the beginning of the block and then again after the first half of trials was completed) in order to sustain the targeted mood (change) throughout the whole experimental session. Hence, in total, participants encountered the MIP five times.

In order to strengthen the effect of positive mood, an evaluative feedback was added (rewarding in the positive mood group, and neutral in the neutral group) at the end of each block (see also Bakic et al., 2014, for a similar procedure). This (bogus) feedback consisted of a small text shown on the screen, informing participants that they had to wait briefly until the

computer had calculated online their learning performance up to that trial number. After a few seconds, an Excel-like scatter plot appeared on the screen, showing them their alleged performance level relative to a group of peers. Their score was indicated by means of a color dot. This dot was positioned systematically either higher up in the distribution of scores for participants in the positive mood group, or somewhere in the middle of the distribution for those belonging to the neutral mood group. Next to this scatter plot, a specific written message was included. It informed them to try to keep the same level of performance. Manipulation checks of mood based on visual analog scales or VAS's (see results here below) confirmed that this procedure (combined with the MIP) actually produced the desired effects: an increase of positive affect in the positive mood group, with no change in affect (neither positive, nor negative) in the neutral mood group. However, we have good reasons to believe that the change in positive mood (in the positive mood group) was mainly due to the MIP and the use of either guided imagery or movie clips (see also Vanlessen et al., 2013, 2014), and not so much to this (infrequent) feedback manipulation that occurred only three times during the course of the experiment. After each block, participants were asked to indicate, for each of the 6 stimuli, the clarity and certainty of each of the six stimulus-response (S-R) associations, by means of a horizontal 10-cm VAS.

Finally, participants were asked to fill out several trait-related questionnaires: the Beck Depression Inventory (Beck, Steer, Ball, & Ranieri, 1996), and the Resilience scale translated in Dutch (Portzky, De Bacquer, Audenaert, & Wagnild, 2010). The whole experiment lasted about 2 hours.

2.4. Mood induction

Experiment 1. We used a previously validated MIP (see Vanlessen, et al., 2013, 2014; Bakic et al., 2014). Mood was induced by means of a guided imagery procedure, where participants

were instructed to vividly imagine reliving either a happy or neutral (depending on the group they were assigned to) autobiographical memory (Holmes et al., 2008, 2006). First, the participants were trained in taking a field perspective (i.e., imagining from one's own perspective) during mental imagery. Then they had to choose an appropriate happy/neutral event, an episodic memory that happened at least a week before, and report about it. For the recall that would ensue, they were instructed to keep their eyes closed and visualize all the specificities of the memory, and to use the field perspective (Watkins & Moberly, 2009, based on Holmes et al., 2008). The actual recall session was divided into two parts of 30 seconds, and in between participants were asked questions about different aspects of the happy/neutral memory they were imagining. Participants were blind to the real purpose of the procedure, believing that it was about remembering an event from the past as vividly as possible (and not about re-experiencing the emotion of the event). After each mood induction, participants marked on 10-cm horizontal VAS their current level of happiness, pleasantness, and sadness, with "neutral" on one end/anchor to "as happy/pleasant/sad as I can imagine" on the other.

Experiment2. Positive or neutral mood was induced by watching short (either positive or neutral) movie clips. These movie clips were selected from a larger set of excerpts that were all rated by an independent sample of participants. For this pilot experiment, the movies were gathered from a variety of internet sources, and matched for length (M=4.53 minutes, SD=0.79), use of English language and music, number and type of characters (human, animated, animal). Thirty seven psychology students at Ghent University (Mean age = 20.8 years, SD = 20.80; 8 males) first watched and then rated twenty different movie clips shown in random order (10 in each condition), along several dimensions each time: happiness, sadness, relaxation, tension, interest, and level of attention/immersion using 15-points Likert scales. Based on these ratings, we eventually selected 8 movies clips (4 in the positive and 4 in the neutral condition). The 4 "happy" movie clips had a mean happiness rating of 12.01

(SD = 0.26). By contrast, the 4 "neutral" movie clips had a mean happiness rating of 9.04 (SD = 0.83)[2], ($t(6)=6.83$, $p<.01$). Importantly, these two sets of movie clips were matched for interest (M=10.85, SD= 0.90 for positive and M= 11.07, SD= 0.19 for neutral clips) ($t(6)=-0.49$, $p=.64$) and relaxation (M=12.11, SD= 1.64 for positive and M= 12.01, SD= 0.46 for neutral clips , $t(6)=0.12$, $p=.91$).

2.5. Data analysis

Mood manipulation. The efficiency of the change in positive (relative to neutral) mood following the MIP was assessed by mixed model ANOVAs with Group (n=2) and Experiment (n=2) as between subject factors and Time as within subject factor (n=5), followed by independent t-tests. In order to control for potential baseline differences, baseline measure was subtracted from every subsequent time point and these difference scores were then subjected to the same mixed model ANOVA.

Accuracy. Accuracy during the probabilistic learning task was expressed in proportion of correct responses, separately for each Condition (n=3). Moreover, for each Condition separately (2 stimuli x 40 repetitions), changes of learning performance as a function of time were captured by grouping the data into four bins of equal size (i.e., 20 trials/condition) (see Eppinger et al., 2008 for a similar approach). These data were then submitted to a mixed model ANOVA with Group and Experiment as between subject factors, and Condition and Bin as within subject factors.

Reaction time. Reaction times data were analyzed using the same statistical model as used for accuracy data.

[2] In the positive group, selected movie clips were: "Mr. Bean, heart attack and first aid", "Gopher broke", excerpt from "Britain's got talent", "Wrecking Orchestra's Amazing Tron Dance". In the neutral group, selected movie clips were : "BBC earth's life of penguins", "How it's made: chocolate", "What's good for you: Afternoon apathy syndrome", "How it's made: French fries".

Switch after negative feedback. We also extracted and analyzed the proportion of switches after negative feedback, as this metric has previously been related to exploration indirectly (e.g. Hills, Todd, & Goldstone, 2010). This measure was submitted to a mixed model ANOVA with Group and Experiment as between subject factor and Condition as within subject factor.

Final cumulative accuracy. This measure represents the cumulative accuracy at the final trial, where participants have seen the stimulus for the last time (trial 40). These data were submitted to a mixed-model ANOVA with Group and Experiment as between subject factor and Condition as within subject factor.

Change point analysis. A potential limitation of the previous analysis carried out on the accuracy data is that arbitrary bins of equal sizes (consisting of 10 trials per stimulus, 20 per condition) are formed and learning is thereby assessed as a function of change in accuracy across these 4 consecutive bins (Eppinger et al., 2008). However, such a grouping method may potentially overlook interesting (but more subtle) differences occurring between the two mood groups when the trial-by-trial evolution in learning (across the 40 repetitions of a given stimulus) is taken into account or modeled. To this end we first calculated the cumulative accuracy at each stimulus presentation for all subjects. Next, the subject-specific cumulative accuracy under the random condition was subtracted at each trial. This new endpoint, calculated under the deterministic and probabilistic condition for each subject, was modeled as a function of the repetition number (excluding the first repetitions) for each subject using a multiphase model (Cudeck and Klebe, 2002). More precisely, a two-phase quadratic-linear model with continuity between segments was fitted using proc nlin in SAS 9.2 for each subject. The repetition number where the model switches from the quadratic increase to a flat line is the change point and determines where learning leveled off after the initial rise. The

estimated change points for each subject were then submitted to a mixed-model ANOVA with Group and Experiment as between subject factor and Condition as within subject factor.

Post-experiment ratings. The clarity and certainty ratings were subjected to a mixed model ANOVA with Group and Experiment as between subject factor and Condition as within subject factor.

Questionnaires. Scores on BDI and RS-nl were subjected to independent t-test analyses with Group and Experiment as factors.

3. RESULTS

Mood/Happiness. The ANOVA revealed a significant two-way Group*Time interaction, ($F(4,224)=14.74$, $p<.01$, $\eta^2=.21$), indicating that positive mood increased happiness ratings compared to neutral mood (see Figure 1). The three-way interaction approached significance, $F(4,224)=2.30$, $p=.06$, $\eta^2=.04$), suggesting possible differential effects between the two MIPs. Additionally, main effects of Experiment ($F(1,56)=6.19$, $p<.05$, $\eta^2=.10$), Group ($F(1,56)=11.78$, $p<.01$, $\eta^2=.17$) and Time ($F(4,224)=5.22$, $p<.01$, $\eta^2=.09$) were significant.

Because the three-way interaction approached significance in this omnibus ANOVA, we performed separate ANOVAs for the two experiments. For guided imagery (Experiment 1), the analysis showed a significant Group*Time interaction ($F(4,116)=9.00$, $p<.01$, $\eta^2=.24$), and main effect of Group ($F(1,29)=6.78$, $p<.05$, $\eta^2=.19$). For movie clips (Experiment 2), the analysis also revealed significant Group*Time interaction ($F(4,108)=8.24$, $p<.01$, $\eta^2=.23$), and a main effect of Group ($F(1,27)=5.48$, $p<.05$, $\eta^2=.17$). Moreover the main effect of Time was also significant in this analysis and experiment ($F(4,108)=4.94$, $p<.01$, $\eta^2=.16$). Unlike what was found in Experiment 1, the significant main effect of time found in Experiment 2 indicated that mood was not stable across the different measurement points. When comparing

levels of happiness between time point 2 and 5 with paired t-tests within each group, we found that the mood was actually stable in all groups (all p's >.05), except for the positive mood group of Experiment 2 (movie clips). There, levels of happiness increased between the second (M= 55.56, SD=17.69) and last (M=63.73, SD=14.62) measurement moment (t(15)=-2.66, p<.05).

Table 1. Happiness ratings. Means (±1 standard deviations) are provided for the positive and neutral groups, for each experiment separately. Results of direct statistical comparisons (by means of independent samples t-tests) are also provided.

Measure point	Experiment 1			Experiment 2			Comparison	
	Positive	Neutral	t-Test	Positive	Neutral	t-Test	Positive T-test	Neutral T-test
1	59.81 (30.38)	57.53 (23.35)	0.23	41.74 (16.90)	41.39 (22.31)	0.50	2.08*	1.86
2	77.63 (23.55)	49.00 (32.30)	2.83**	55.56 (17.69)	43.61 (19.85)	1.71	3.00**	0.52
3	75.38 (22.52)	52.00 (31.64)	2.38*	60.37 (15.59)	37.12 (23.22)	3.22**	2.19*	1.40
4	78.25 (21.54)	43.07 (35.58)	3.36**	57.62 (21.04)	41.15 (21.55)	2.07*	2.74**	0.17
5	81.69 (21.31)	50.27 (33.86)	3.11**	63.73 (14.62)	39.08 (21.31)	3.69**	2.78**	1.03

*p< .05, **p< .01.

Interestingly, direct independent T-tests (Table 1) showed that guided imagery (Experiment 1) led to significantly higher levels of happiness than movie clips (Experiment 2), while no such differences were evidenced between the two neutral groups. However, the

fact that this increase in happiness for the positive mood group in Experiment 1 compared to Experiment 2 was already significant at baseline suggests that the two groups actually differed already from one another prior to the MIP and hence this result has to be interpreted with caution.

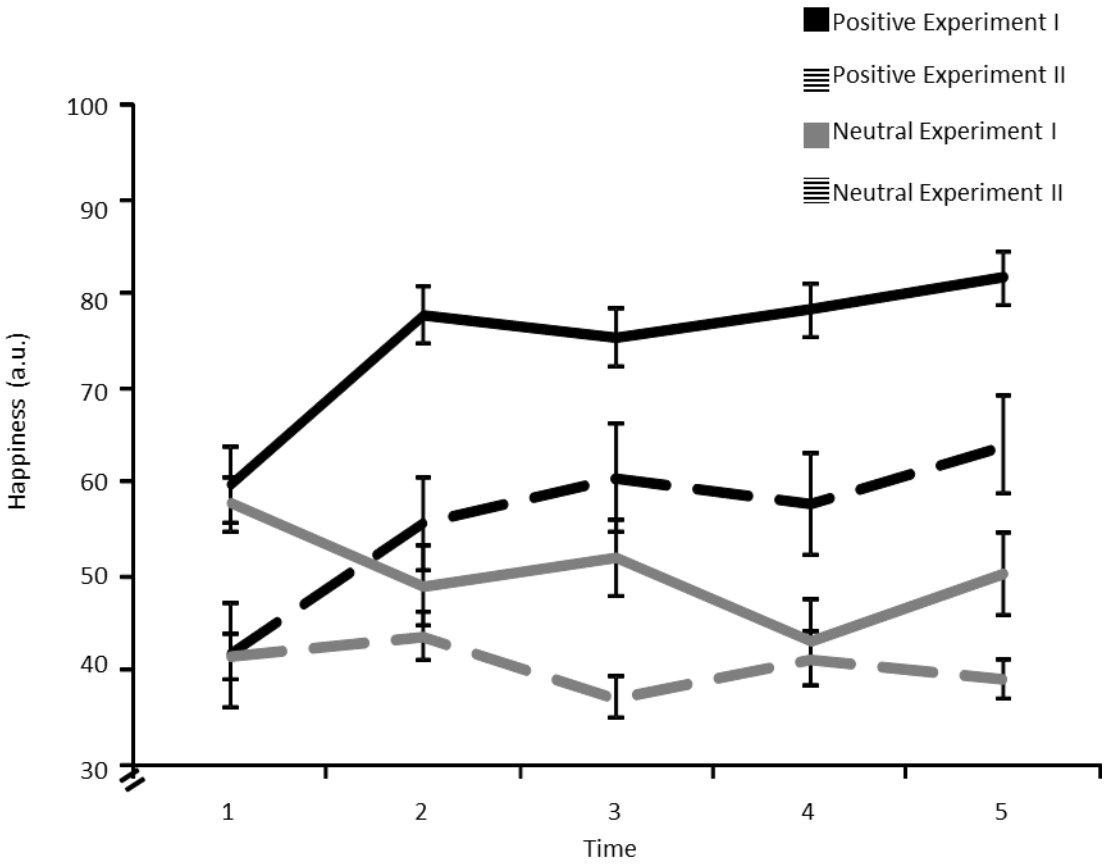

Figure 1. Happiness ratings shown separately for the two groups (Positive vs. Neutral) and the two MIPs (Guided Imagery vs. Movie clips).

Pleasantness. The analysis showed a significant two-way Group*Time interaction ($F(4,224)=7.29$, $p<.01$, $\eta^2=.12$) indicating that, akin to happiness ratings, the two positive mood groups showed a steep increase in levels of pleasantness following the MIP compared to the two neutral mood groups that did not show this effect. The main effects of Experiment ($F(1,56)=7.58$, $p<.01$, $\eta^2=.12$), Group ($F(1,56)=13.37$, $p<.01$, $\eta^2=.19$) and Time ($F(4,224)=5.20$, $p<.01$, $\eta^2=.09$) were also significant (see Table 2).

Noteworthy, direct independent t-tests (Table 2) showed that pleasantness was higher following MIP based on guided imagery than movie clips. No such effect was found when comparing the two neutral groups to each other.

Table 2. Results of the Pleasantness MIP. Means (±1 standard deviations) for positive and neutral groups of two experiments. Results of independent samples t-tests within experiment 1 (df=29), experiment 2(df=27), and direct comparison for positive groups (df=30) and neutral groups (df=26).

Measure point	Experiment1			Experiment 2			Comparison	
	Positive	Neutral	t-test	Positive	Neutral	t-test	Positive t-test	Neutral t-test
1	63.00 (28.41)	54.53 (25.94)	0.87	47.64 (16.31)	39.20 (21.99)	1.19	1.88	1.67
2	77.94 (20.48)	50.73 (31.22)	2.89**	55.72 (18.46)	43.83 (19.81)	1.67	3.22**	0.69
3	76.38 (20.54)	55.13 (29.41)	2.34*	60.66 (16.48)	41.38 (19.37)	2.90**	2.39*	1.44
4	79.00 (18.28)	47.13 (35.01)	3.21**	58.52 (20.83)	39.59 (23.03)	2.32*	2.96**	0.66
5	81.56 (18.40)	53.00 (32.47)	3.04**	62.87 (13.32)	37.37 (20.58)	4.03**	3.29**	4.49

* p< .05, **p< .01.

Sadness. The analysis revealed a significant Experiment*Group interaction ($F(1,55)=3.97$, $p=.05$, $\eta 2=.07$) as well as Group*Time interaction ($F(4,220)=3.31$, $p<.05$, $\eta 2=.06$) and a main effect of Time ($F(4,220)=2.66$, $p<.05$, $\eta 2=.05$). As can be seen from Table 3, levels of sadness were numerically higher in the positive mood group of Experiment 2 compared to Experiment

1, with an opposite trend when comparing the two neutral groups. However, these effects were non-significant (all p's>.05).

Table 3. Results of the Sadness MIP. Means (±1 standard deviations) for positive and neutral groups of two experiments. Results of independent samples t-tests within experiment 1 (df=29), experiment 2(df=27), and direct comparison for positive groups (df=30) and neutral groups (df=26).

Measure point	Experiment 1			Experiment 2			Comparison	
	Positive	Neutral	t-Test	Positive	Neutral	t-Test	Positive t-Test	Neutral t-Test
1	6.75 (5.30)	5.67 (7.35)	-0.47	12.68 (11.50)	4.20 (8.52)	2.21*	-1.87	0.49
2	2.40 (2.50)	10.93 (17.76)	-1.84	9.05 (11.19)	5.28 (5.71)	1.01	-2.25*	1.10
3	5.44 (9.58)	8.13 (12.74)	-0.67	7.21 (6.99)	4.48 (4.27)	1.23	-0.60	0.98
4	3.31 (4.11)	6.93 (13.29)	-1.04	7.98 (8.79)	3.58 (2.09)	1.76	-1.92	0.90
5	2.87 (2.75)	5.40 (8.89)	-1.08	5.11 (4.05)	4.11 (2.49)	0.78	-1.83	0.50

*p< .05, **p< .01.

Too late responses. The number of too late responses was modest (M=1.70, SD=1.07). However, it did vary depending on Experiment (F(1,56)=17.77, p<.01, η2=.24). Participants in Experiment 1 (M=1.35, SD=0.82) had overall a smaller number of too late responses than participants in Experiment 2 (M=2.14, SD=1.0), (t(58)=-3.35, p<.01

Accuracy. As expected (see Eppinger et al., 2008), the analysis confirmed a significant Condition*Bin interaction ($F(6,336)=21.47$, $p<.01$, $\eta^2=.28$). This interaction showed that learning was the steepest and highest for the deterministic condition, intermediate for the probabilistic condition, and flat for the random condition (Figure 2). Main effects of Condition ($F(2,112)=233.24$, $p<.01$, $\eta^2=.81$) and Bin ($F(3,168)=49.68$, $p<.01$, $\eta^2=.47$) were also significant.

Moreover, the Experiment*Bin interaction ($F(3,168)=4.45$, $p<.05$, $\eta^2=.07$) was also significant, suggesting that regardless of the MIP used, learning was somehow different between the two experiments. Followed up t-tests showed a better learning at bin 1 (i.e., the first quarter of the trials) for participants in Experiment 2 (M=0.60, SD=0.08) than Experiment 1 (M=0.56, SD=0.05), ($t(58)=-2.22$, $p<.05$). No such difference between the two experiments was found for the three subsequent (three) bins (all p's>.05).

Reaction times (RTs) for correct responses. Analysis of RTs for correct responses showed a significant three way Experiment*Condition*Bin interaction ($F(6,336)=2.31$, $p<.05$,$\eta^2=.04$), a significant Experiment*Bin interaction ($F(3,168)=3.28$, $p<.05$,$\eta^2=.06$), a significant Condition*Bin interaction ($F(6,336)=16.39$, $p<.01$,$\eta^2=.23$), a significant main effect of Bin ($F(3,168)=24.48$, $p<.01$, $\eta^2=.30$) and a significant main effect of Condition ($F(2,112)=230.20$, $p<.01$,$\eta^2=.80$). However, subsequent t-tests failed to reveal reliable differences between the two experiments and/or the two MIPs.

Switches after the negative feedback. The Condition*Bin interaction ($F(6,336)=2.18$, $p<.05$,$\eta2=.06$), as well as the main effects of Condition ($F(2,112)=12.61$, $p<.01$, $\eta2= .18$) and Bin ($F(3,168)=34.43$, $p<.01$, $\eta2= .37$) were significant. These effects indicated that participants switched the most after receiving negative feedback in the random condition and the least in deterministic condition, with the probabilistic condition being intermediate

between these two. Moreover, this change in the rate of switches after negative feedback as a function of condition grew steadily from bin 1 to bin 4. However, neither Experiment nor Group did influence these effects.

Final cumulative accuracy. Final cumulative accuracy was the highest in the deterministic condition, lower in the probabilistic condition and the lowest in the random condition, as evidenced by a significant main effect of Condition ($F(1,50)=282.95$, $p<.01$, $\eta^2=.85$). However, none of the other experimental factors influenced this condition-dependent effect(all p's>.05).

Change point. The data of six participants (one in the positive mood group of Experiment 1,three in the positive mood group of Experiment 2 and two in the neutral mood group of Experiment 2) had to be excluded from this analysis because no clear change point was found in these data. The analysis revealed a significant two-way Experiment*Group interaction ($F(1,50)=4.57$, $p<.05$, $\eta2= .08$), as well as significant main effect of Group ($F(1,50)=4.29$, $p<.05$, $\eta2= .08$). Follow-up t-tests showed that change point occurred earlier for the neutral mood group of Experiment 2 (movie clips; M=18.22, SD=2.73) than Experiment 1 (guided imagery; M=21.37, SD=3.68), ($t(24)=2.40$, $p<.05$). This difference was not found when comparing directly the two positive mood groups to each other (guided imagery; M= 21.26, SD=6.91; movie clips; M=25.26, SD=8.79), ($t(26)=-1.35$, $p=.19$). Additionally, change point did not differ between the two groups in Experiment 1, ($t(28)=-0.06$, $p=.96$), while it did in Experiment 2, ($t(22)=2.55$, $p<.05$).

Post-experiment ratings. The mixed-model ANOVA carried out on the clarity ratings showed a significant Experiment*Condition interaction ($F(2,110)=13.08$, $p<.01$, $\eta^2=.19$), and significant main effects of Experiment ($F(1,55)=13.47$, $p<.01$, $\eta^2=.20$) and Condition ($F(2,110)=63.75$, $p<.01$, $\eta^2=.54$). The analysis performed on the certainty ratings revealed

significant main effects of Experiment ($F(1,55)=10.84$, $p< .05$, $\eta^2= .17$) and Condition ($F(2,110)=206.6$, $p< .01$, $\eta^2= .80$). These results showed that perceived clarity (and response certainty) regarding the S-R associations went down monotonically when S-R probabilities decreased, but that this decrease was not balanced between the two experiments. Direct comparisons between the two experiments showed that they did not differ in clarity for the deterministic condition, but they did for the probabilistic and random conditions (Table 4).

Certainty was the highest for the deterministic condition ($M=57.51$, $SD=10.06$) and intermediate for the probabilistic condition ($M=36.86$, $SD=9.84$; $t(58) = 13.22$, $p<.01$). Certainty was also significantly higher in the probabilistic compared to the random condition ($M=24.47$, $SD=11.87$; $t(58)=8.33$, $p<.01$). Additionally, lower certainty ratings were found in Experiment 1 than Experiment 2 for the deterministic condition selectively (Table 4).

Questionnaires. Independent t-tests revealed significant group difference for the BDI in Experiment 1. Moreover, when compared directly, we found that resilience scores of both positive and neutral mood groups of Experiment 1 were significantly higher than in Experiment 2 (see Table 5).

Table 4. Results of the Clarity and Certainty comparisons for the two experiments, means (±1 standard deviations) and independent samples t-tests (df=57).

		Experiment 1	Experiment 2	t-Test
Clarity	Deterministic	80.65 (28.90)	84.22 (6.95)	-0.64
	Probabilistic	82.26 (18.72)	70.88 (14.47)	2.59*
	Random	62.37 (27.21)	32.13 (7.87)	5.67**
Certainty	Deterministic	52.74 (8.97)	62.79 (8.53)	-4.40**
	Probabilistic	35.22 (8.27)	38.68 (11.20)	-1.36
	Random	22.15 (7.48)	27.04 (15.08)	-1.60

*p< .05, **p< .01.

Table 5. BDI and RS questionnaires. Means (±1 standard deviations) for positive and neutral group of two experiments, and independent samples t-tests between groups in experiment 1 (df=29), experiment 2(df=27), and direct comparison for positive groups (df=30) and neutral groups (df=24).

Condition	Experiment 1			Experiment 2			Comparison	
	Positive	Neutral	t-Test	Positive	Neutral	t-Test	Positive	Neutral
BDI	10.38 (6.20)	5.60 (4.60)	2.42 *	7.38 (5.23)	4.31 (3.07)	1.86	1.50	0.86
RS-nl	76.00 (8.12)	76.33 (8.72)	-0.11	53.00 (8.59)	49.31 (9.53)	1.10	7.78 **	7.84 **

*p< .05, **p< .01.

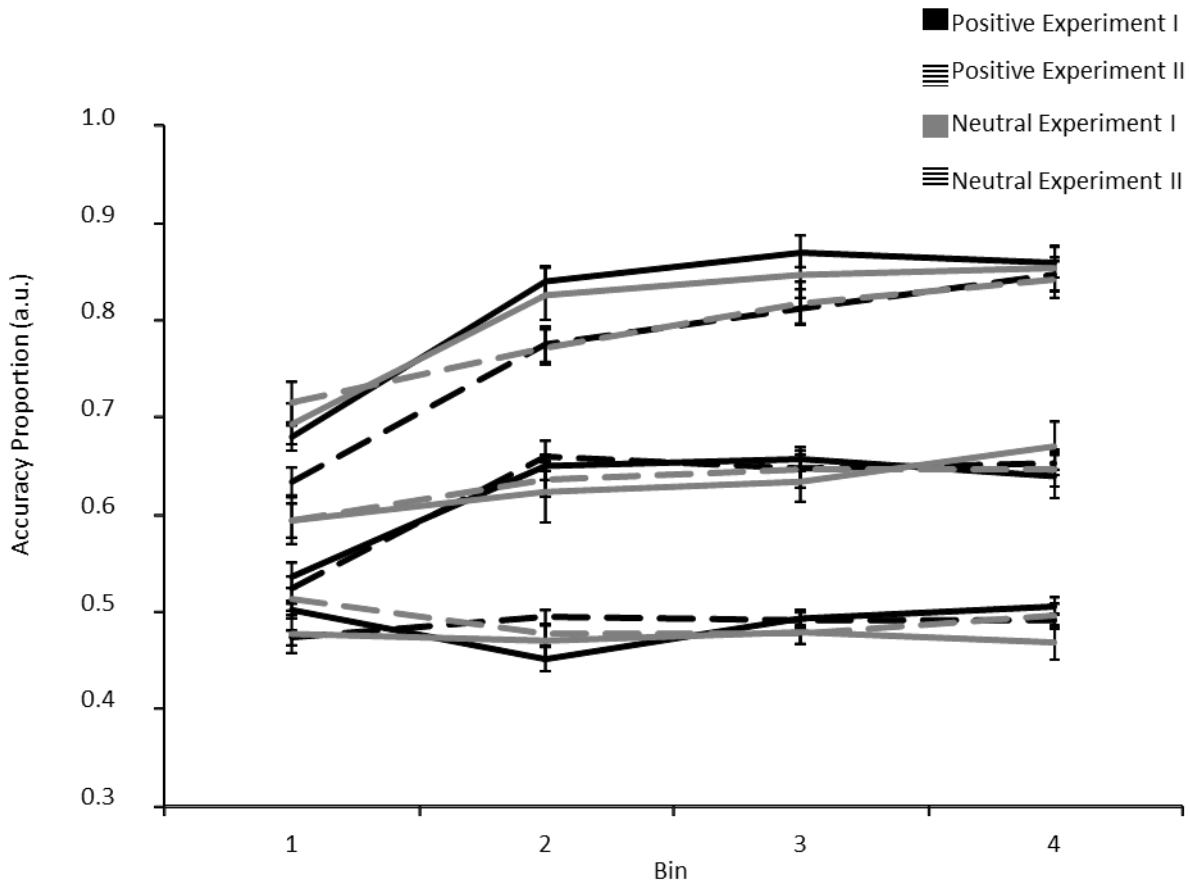

Figure 2. Accuracy data (i.e., proportion of correct responses) decomposed as a function of bin, condition and group.

4. DISCUSSION

The goal of this study was twofold. First we sought to test whether the transient experience of positive mood, such as induced experimentally in healthy adult participants by means of a specific MIP, might influence probabilistic learning given that this process requires to flexibly alternate between two dominant modes or styles, namely exploration vs. exploitation (Cohen et al., 2007; Eppinger et al., 2008; Frank, Doll, Oas-Terpstra, & Moreno, 2009; Goschke & Bolte, 2013). Learning was quantified using standard accuracy and RT data, as well as the rate of switches after negative feedback (which provides an estimate or proxy of exploration, see Hills et al., 2010) and change/inflection point (which likely informs about a qualitative change during learning). Second, we contrasted guided imagery to movie clips as MIP, to

evaluate whether the former might yield a stronger change in positive affect/mood than the latter, in line with previous studies and models put forward in the literature (Chiew & Braver, 2014; Csikszentmihalyi, 1999; Tice, Baumeister, Shmueli, & Muraven, 2007). Our results showed that while the task chosen (see Eppinger et al., 2008) was successful to spark off probabilistic learning, positive mood did not influence this process in a straightforward way however (e.g., learning performance was not higher or better after the induction of positive compared to neutral mood). Moreover, although our MIP turned out to be efficient to induce and sustain positive mood (at the subjective level) throughout the whole experimental session, we failed to observe systematic differential effects between the two MIPs (guided imagery vs. movie clips). Here below, we discuss some of the reasons likely accounting for these null findings in greater detail.

We replicated the behavioral results obtained earlier by Eppinger et al. (2008) using the same probabilistic learning task. Accuracy/performance was clearly modulated by the S-R probabilities, such that learning was the highest in the deterministic, intermediate in the probabilistic and low or at chance level in the random condition. Post-experiment ratings also corroborated this conclusion. Likewise, when using a more refined measure of exploration, namely the number of switches after negative feedback, we found a monotonic decrease of this rate depending on the S-R probability: participants switched less after negative feedback in the deterministic than in the probabilistic condition, which in turn had a lower rate than the random condition. Exploration behavior (Hills et al., 2010) was therefore clearly varying as a function of the probabilistic structure of the task. By contrast, the analysis carried out on change points did not reveal a clear outcome depending on probabilities of S-R associations.

However and crucially, we failed to evidence significant mood-dependent modulations of probabilistic learning. Hence, positive mood did not increase or decrease (probabilistic) learning, nor did it influence exploration, relative to a control condition with an active MIP

(but neutral content). Importantly, this lack of modulation of learning by positive mood could not be explained by a suboptimal or inefficient MIP. In both experiments, we found a sharp increase in positive affect at the subjective level (both for happiness as well as a milder positive emotion or feeling, as captured by pleasantness) following the MIP, for the positive mood group selectively. In addition, our results also clearly showed that this change in positive affect following the MIP in the positive mood group did not simply fade away, but instead, it remained high and stable throughout the entire duration of the experimental session, and hence, during the probabilistic learning task. Accordingly, when treated in isolation, each of the two experimental phases (MIP on the one hand and probabilistic learning on the other) yielded the expected results, but we did not find evidence for a clear transfer of the beneficial effects brought about by the MIP to the probabilistic learning task. Contrary to our predictions, feeling happy did not lead to a change in probabilistic learning performance (likely resulting from a higher flexibility). Therefore, our findings do not confirm earlier models or theories put forward in the literature, postulating a more flexible or explorative processing style after the induction or (temporary) experience of positive affect (Ashby et al., 1999; Fredrickson, 2001, 2004; Shiota et al., 2014). Along the same line, our results also do not support the possibility that positive mood might create distraction or merely brings (more) noise in the system (as opposed to enhancing flexibility, attention or creativity), and thereby might impact (learning) behavior negatively (Ellis & Ashbrook, 1988; Pessoa, 2008). Further, while Nadler et al. (2010) previously found that positive mood fostered learning when it was based on hypothesis testing (hence sharing perhaps some similarities with the probabilistic learning task as used here), there are major methodological differences between this earlier study and our study, including regarding the way positive affect is induced, as well as how learning is eventually conceived and operationalized, which might account for this discrepancy. Although speculative at this stage, several reasons, either conceptual or

methodological in nature, might potentially explain the lack of mood effects during probabilistic learning, as found here. At the theoretical level, it is conceivable that our MIPs (either guided-imagery or movie clips) did not improve flexibility or a higher-order (perhaps prefrontal and dopaminergic-based) cognitive function (see Ashby et al., 1999), but instead other (simpler) processes or mental operations, such as selective attention for example (see Vanlessen et al., 2013, 2014). In this scenario, no clear effect of positive mood on probabilistic learning is evidenced because this complex process does not encompass gains in cognition/attention usually accompanying the experience of positive affect. Indeed, a broadening of attention is probably not needed in order to perform better (or worse) during probabilistic learning. Alternatively, it may be the case that the link between (cognitive) flexibility and probabilistic learning is more complex or less transparent than initially thought. More specifically, switching back and forth between exploration and exploitation, as required during probabilistic learning, might be supported by another sort or type of flexibility (see Frank et al., 2005) than what is usually entailed in earlier dominant models in the literature (Ashby et al., 1999; Fredrickson, 2001, 2004; Huntsinger, 2012). At the methodological level, we cannot formally exclude the possibility that (beneficial) effects of positive mood during probabilistic learning were actually very well present, but our measures of learning were not sensitive enough to capture these mood-related changes or gains. For example binning or grouping accuracy data (see also Eppinger et al., 2008) might undermine the possibility to reveal subtle changes in learning as a function of positive mood. We note however that when we used alternative measures of learning (such as the switch rate after negative feedback or change point, which both seek to exploit the evolution of learning across the single trial data and hence capitalize on qualitative changes during learning, as opposed to quantitative changes in the case of accuracy for example), we still failed to disclose significant group differences, which casts doubt on the assumption that we actually used suboptimal or invalid

measures of learning in our study. In this context, the use of complementing EEG/ERP (see Frank et al., 2005; Eppinger et al., 2008) and/or computational modeling methods (Jepma & Nieuwenhuis, 2011) might turn out to reveal important mood-related differences during probabilistic learning, which might go unnoticed when using "standard" behavioral measures or indices.

A second important observation in our study is that positive mood was equally strongly induced following guided imagery (Experiment 1) or watching (funny) movie clips (Experiment 2). Even though guided imagery was found to produce here and there sometimes stronger changes in levels of positive mood than movie clips (see pleasantness ratings and Table 2), these differences were rather modest and not systematic. When focusing on levels of happiness, we found a larger increase for guided imagery than movie clips but this effect was mainly explained by pre-MIP, baseline differences between the two experiments. Because Experiment 1 (guided imagery) was performed during the spring/summer period while Experiment 2 (move clips) took place in the autumn season, it may well be that participants were spontaneously or intrinsically happier in Experiment 1 than 2 (Keller et al., 2005), regardless of the MIP (see also Table 1). This seasonal effect might also explain some of the other and largely unexpected differences found between the two experiments, including the number of too late responses, accuracy results for bin 1 and post-experiment ratings. Accordingly, we cannot formally exclude the possibility that some of the significant differences found between the two MIPs in their propensity to elicit happy mood or positive affect (see Tables 1-3) are actually conflated by "noise" or nuisance in our between-subjects experimental design, such as seasonal fluctuations in baseline levels of happiness.

Although movie clips are often considered in the literature as the best means to induce (positive) mood (Martin, 1990), probably because of obvious advantages related to its "easy" application or administration and the existence of specific data bases, the results of our study

show that guided imagery is able to produce equally strong (if not stronger) changes in the mood state of the participants than movie clips. As such, our findings concur on the idea that mentally imagining certain actions or activities vividly (here with a positive content) has clear consequences for the current mood state and presumably well-being more generally. A possible advantage in using guided imagery over movie clips is that ecological validity and authenticity (Rottenberg et al., 2007) of the feeling of happiness are by definition maximized or optimized: while watching movie clips, many (uncontrolled or hardly controllable) socio-cultural factors (as well as prior exposure) may actually play a role in the elicitation of the positive affect or mood triggered by these visual stimuli. Obviously, good or famous movies are perceived or remembered as such because they likely promote immersion (maybe even a sense or impression of flow even, see Csikszentmihalyi (1999) in most of the spectators. However, this level of engagement appears pretty hard to titrate in laboratory conditions. In comparison, when using guided imagery, each participant is asked to relive vividly personal or idiosyncratic episodic memories, an element which inherently fosters engagement and individualization of the experienced feeling. Moreover, it should also be emphasized that both movie clips and guided imagery were MIPs free of changes in motivation, besides positive affect. When reward, positive (performance) feedback or unexpected gifts are provided to elicit positive mood, a potential problem is that changes in mood are somewhat confounded by changes in motivation (Shiota et al. 2014). In our study, the two MIPs used did not create changes in motivation that might have influenced performance or behavior during probabilistic learning.

A few limitations have to be noted. First, our sample size was modest. Notwithstanding this element, we found clear and significant effects of the two MIPs, as well as reward probability during learning. Hence, it seems unlikely that sample size (or power) alone accounts for the present null findings regarding effects of positive mood on probabilistic

learning, as well as possible differential effects of the two MIPs to elicit positive mood. Second, because Experiment 1 and 2 took place at different periods of the year, seasonal effects may have played a role and brought additional noise that might have masked or obscured important differences between the two groups. Future studies should eliminate this confound, as well as include MIP as a within-subject factor probably. Third, the lack of clear differentiation between the two MIPs in the elicitation of positive mood might be imputed to the use of VASes (targeting subjective levels of happiness, pleasantness and sadness) that were perhaps not entirely appropriate to do so. Additional scales or measures focusing on the level of interest, amusement, or the strength of internal/external focus might be added to help better tease apart the specific contribution of guided imagery vs. movie clips to the elicitation of positive mood. Finally, it appears essential to include electrophysiological or psychophysiological measurements which correlate "directly" with (the size and extent) learning at a given moment in time (such as the ERN and FRN components; see Frank et al., 2005), and hence which eventually provide more objective estimates of the underlying cognitive and emotional processes under scrutiny than standard behavioral measures (accuracy and RTs) or subjective ratings. Ideally, behavioral and psychophysiological should be used concurrently in order to unravel mood-related changes during probabilistic learning.

Altogether, the results of this study show that positive mood can reliably be elicited by either guided imagery or movie clips, equally strongly in these two cases. However, this change in the current affective state of the participant does not produce clear quantitative or qualitative changes in the way probabilistic learning unfolds over time (in healthy adult participants). As such, these new findings do not lend support to the assumption of a (general) change in (cognitive) flexibility brought about by the MIP that could in turn bias or improve learning, when this process is based on the direct processing of positive and negative reinforcers available in the environment.

5. REFERENCES

Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529–550. Retrieved from <Go to ISI>://000082133000004

Bakic, J., Jepma, M., De Raedt, R., & Pourtois, G. (2014). Effects of positive mood on probabilistic learning: Behavioral and electrophysiological correlates. *Biological Psychology*, *103*, 223–232. doi:10.1016/j.biopsycho.2014.09.012

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588–597. doi:10.1207/s15327752jpa6703_13

Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 509–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24733296

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 933–42. doi:10.1098/rstb.2007.2098

Csikszentmihalyi, M. (1999). If we are so rich, why aren't we happy? *American Psychologist*. doi:10.1037/0003-066X.54.10.821

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), 876–9. doi:10.1038/nature04766

Ellis, H. A., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fielder & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 25–43). Lewiston, NY: Hogrefe.

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

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*(8), 1062–1068. doi:10.1038/nn.2342

Frank, Woroch, & Curran. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. doi:10.1016/j.neuron.2005.06.020

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *Am Psychol*, *56*(3), 218–226. Retrieved from doi:10.1037/0003-066x.56.3.218

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci*, *359*(1449), 1367–1378. doi:10.1098/rstb.2004.1512

Garland, E. L., Fredrickson, B., Kring, A. M., Johnson, D. P., Meyer, P. S., & Penn, D. L. (2010). Upward spirals of positive emotions counter downward spirals of negativity: insights from the broaden-and-build theory and affective neuroscience on the treatment of emotion dysfunctions and deficits in psychopathology. *Clinical Psychology Review*, *30*(7), 849–64. doi:10.1016/j.cpr.2010.03.002

Goschke, T., & Bolte, A. (2013). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*.

Gray, J. A. (1994). Personality dimensions and emotion systems. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp. 329–331). New York: Oxford University Press.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: priming exploration and exploitation across domains. *Journal of Experimental Psychology. General*, *139*(4), 590–609. doi:10.1037/a0020666

Holmes, E. A., Coughtrey, A. E., & Connor, A. (2008). Looking at or through rose-tinted glasses? Imagery perspective and positive mood. *Emotion (Washington, D.C.)*, *8*(6), 875–879. doi:10.1037/a0013617

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). *Positive interpretation training: effects of mental imagery versus verbal training on positive mood. Behavior therapy* (Vol. 37). doi:10.1016/j.beth.2006.02.002

Huntsinger, J. R. (2012). Does positive affect broaden and negative affect narrow attentional scope? A new answer to an old question. *Journal of Experimental Psychology: General*, *141*(4), 595–600.

Jorm, A. ., Christensen, H., Henderson, A. ., Jacomb, P. ., Korten, A. ., & Rodgers, B. (1998). Using the BIS/BAS scales to measure behavioural inhibition and behavioural activation: Factor structure, validity and norms in a large community sample. *Personality and Individual Differences*, *26*(1), 49–58. doi:10.1016/S0191-8869(98)00143-3

Keller, M. C., Fredrickson, B. L., Ybarra, O., Côté, S., Johnson, K., Mikels, J., … Wager, T. (2005). A warm heart and a clear head: The contingent effects of weather on mood and cognition. *Psychological Science*, *16*(9), 724–731.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: does happiness lead to success? *Psychological Bulletin*, *131*(6), 803–855. doi:10.1037/0033-2909.131.6.803

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review*, *10*(6), 669–697.

Moore, S. G., Ferguson, M. J., & Chartrand, T. L. (2011). Affect in the aftermath: how goal pursuit influences implicit evaluations. *Cognition & Emotion*, *25*(3), 453–465. doi:10.1080/02699931.2010.538598

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better mood and better performance. Learning rule-described categories is enhanced by positive mood. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(12), 1770–1776. doi:10.1177/0956797610387441

Portzky, M., Wagnild, G., De Bacquer, D., & Audenaert, K. (2010). Psychometric evaluation of the Dutch Resilience Scale RS-nl on 3265 healthy participants: a confirmation of the association between age and resilience found with the Swedish version. *Scandinavian Journal of Caring Sciences*, *24 Suppl 1*, 86–92. doi:10.1111/j.1471-6712.2010.00841.x

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15109163

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: a basic mechanism that warrants flexibility in emotion and motivation. *Emotion (Washington, D.C.)*, *8*(1), 34–46. doi:10.1037/1528-3542.8.1.34

Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. Handbook of emotion elicitation and assessment. In *Handbook of emotion elicitation and assessment Series in affective science* (pp. 2–28). doi:10.1080/02699939508408966

Schwager, S., & Rothermund, K. (2013). Counter-regulation triggered by emotions: positive/negative affective states elicit opposite valence biases in affective processing. *Cognition & Emotion*, *27*(5), 839–855.

Shiota, M. N., Neufeld, S. L., Danvers, A. F., Osborne, E. A., Sng, O., & Yee, C. I. (2014). Positive emotion differentiation: A functional approach. *Social and Personality Psychology Compas*, *8*(3), 104–117.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology*, *43*(3), 379–384. doi:DOI 10.1016/j.jesp.2006.05.007

Van Steenbergen, H., Band, G. P., & Hommel, B. (2010). In the mood for adaptation: how affect regulates conflict-driven control. *Psychological Science*, *21*(11), 1629–1634. doi:10.1177/0956797610385951

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2013). Positive emotion broadens attention focus through decreased position-specific spatial encoding in early visual cortex: Evidence from ERPs. *Cognitive, Affective, & Behavioral Neuroscience*. doi:10.3758/s13415-012-0130-x

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2014). Feeling happy enhances early spatial encoding of peripheral information automatically: electrophysiological time-

course and neural sources. *Cognitive, Affective & Behavioral Neuroscience*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24570275

Westermann, R., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity o f mood induction procedures : a meta- analysis, *26*(January 1995), 557–580.

CHAPTER 3

Effects of positive mood on probabilistic learning: behavioral and electrophysiological correlates[1]

Whether positive mood can change reinforcement learning or not remains an open question. In this study, we used a probabilistic learning task and explored whether positive mood could alter the way positive versus negative feedback was used to guide learning. This process was characterized both at the behavioral and electro-encephalographic levels. Thirty two participants were randomly allocated either to a positive or a neutral (control) mood condition. Behavioral results showed that while learning performance was balanced between the two groups, participants in the positive mood group had a higher learning rate than participants in the neutral mood group. At the electrophysiological level, we found that positive mood increased the error-related negativity when the stimulus-response associations were deterministic, selectively (as opposed to random or probabilistic). However, it did not influence the feedback-related negativity. These new findings are discussed in terms of an enhanced internal reward prediction error signal after the induction of positive mood when the probability of getting a reward is high.

[1] Bakic, J., Jepma, M., De Raedt, R., & Pourtois, G. (2014). Effects of positive mood on probabilistic learning: Behavioral and electrophysiological correlates. *Biological Psychology*, *103*, 223–232. doi:10.1016/j.biopsycho.2014.09.012

1. INTRODUCTION

The influential "broaden and build" theory put forward by Fredrickson (2004) stipulates that positive emotion is associated with a variety of beneficial changes regarding emotional and cognitive functioning. Among them, positive mood might shield the organism from experiencing high levels of negative affect (or anxiety) during the encounter of punishment or worse than expected outcomes (Fredrickson, 2004). Conversely, it might foster reward processing in the face of positive or favorable outcomes. Previous research already showed that depression (characterized by a lack of positive affect or Anhedonia) influences these two processes concurrently (Eshel & Roiser, 2010). Moreover, earlier studies and models already emphasized that positive affect fuels problem solving, creativity as well as cognitive control, suggesting that its effect on high-level cognition might be domain general, as opposed to valence specific (for a review, see Ashby, Isen, & Turken, 1999; Garland et al., 2010). In other words, positive mood might influence specific cognitive processes, with or without producing a selective change in the way either reward or punishment is eventually processed. Furthermore, positive mood might play an important role in well-being and contribute to tip the balance of emotional processing by downplaying negative affect, while concurrently up-regulating the weight of positive affect. However, experimental evidence corroborating this assumption is scant. An elegant means to assess how positive mood might dynamically influence emotional processing (when positive and negative reinforcers are in direct competition in the environment and are therefore deemed equally important events to guide behavior) is provided by probabilistic learning tasks.

During probabilistic learning, the agent usually takes actions, and through trial and error, he/she tries to decipher associations linking stimuli to rewards (Frank, Woroch, & Curran, 2005). Feedback (being either a reward or a punishment) is provided following each action to signal

whether this action was correct or not. This typically results in an increase in rewarded actions and a decrease in punished actions over time, which is known as the law of effect in reinforcement learning (RL) model (Walsh & Anderson, 2011). Updating of the response pattern based on the reward and punishment can be traced by the learning rate parameter (α), which reflects current levels of uncertainty about one's choice strategies, and reliance on the history of previous responses (Behrens, Woolrich, Walton, & Rushworth, 2007). Whether or not positive emotion can reliably influence these processes remains an open issue. Earlier research already showed that positive mood could be beneficial (compared to neutral or negative mood) to cognitive flexibility and in turn learning performance (when it is primarily rule-based or rule described; see Nadler et al., 2010). However, it is still unclear whether these changes in learning following the induction of positive mood may be explained by a differential sensitivity to cues signaling (violation of) reward.

To address this question, in this study, we used a previously validated mood induction procedure (Vanlessen, Rossi, De Raedt, & Pourtois, 2013) to induce, using a between-subjects design, either a positive or neutral (control) mood, which was then maintained successfully throughout the entire experimental session by means of several rehearsals. Participants performed a variant of the probabilistic learning task previously devised by Eppinger, Kray, Mock, and Mecklinger (2008), while electroencephalogram (EEG) was recorded continuously in order to explore changes in the error and feedback related negativity (ERN and FRN) components depending on learning and mood. We chose this specific task setting (see also Frank et al., 2005) because it allows us to explore and characterize mood-related (neurophysiological) changes occurring either at the response or feedback level across different conditions varying in reward probability and shown in random order (Eppinger et al., 2008). These two ERP components are

traditionally related to error processing based on internal (motor) and external cues, respectively. The ERN/Ne is a negative deflection over fronto-central electrodes peaking ~0-100 ms after the onset of an incorrect response, hence is based on an internal (motor) representation (Falkenstein et al., 1991; Falkenstein et al., 2000; Carbonnell & Falkenstein, 2006; Gehring, Goss, Coles, Meyer, & Donchin, 1993). By contrast, the FRN is thought to be its feedback-related counterpart, appearing roughly 250-400 ms after the presentation of negative feedback (Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997).

In our study, we examined mood effects on learning in this task using (i) standard behavioral data (i.e., speed and accuracy of the responses, as in Eppinger et al., 2008), and (ii) estimated learning-rate parameters extracted from a reinforcement-learning (RL) model (Sutton & Barto, 1998) with separate learning rates for positive and negative reinforcement (e.g., Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). Two separate sets of predictions were made. At the behavioral level, we expected that positive mood might alter learning during this task. More specifically, because positive mood might fuel reward processing (Garland et al., 2010) and it shares common neurobiological ground with it (Ashby et al., 1999), we surmised that participants in the positive mood group would be more sensitive to positive feedback than participants in the neutral mood group. Accordingly, learning rate following positive feedback could be higher in the positive, compared to the neutral mood group. At the electrophysiological level, because both the ERN and FRN components have been linked to reward prediction error (Holyroyd & Coles, 2002, Nieuwenhuis, Slagter, von Geusau, Heslenfeld, & Holroyd, 2005), we predicted that positive mood might influence the size of these ERP components (see also Lange, Leue, & Beauducel, 2012), in opposite situations however. More specifically, we predicted that positive mood could increase the magnitude of the ERN component (coding reward prediction error at the

response level) when learning was deemed feasible during the probabilistic learning task (Eppinger et al., 2008), that is when (motor) cues signaling violation of reward were actually informative to guide learning (i.e., they conveyed relevant information about the current goal conduciveness of the action performed). By contrast, when learning was made impossible (because feedback cues signaling reward or punishment appeared to be random and/or unrelated to the actual action performed), we surmised that positive mood could increase the size of the FRN component (coding reward prediction error at the feedback level). Translated to the task previously devised by Eppinger et al. (2008) and used here, we therefore hypothesized that positive mood (compared to neutral mood) could increase the ERN when the clarity and certainty regarding the stimulus-response associations were high ("deterministic" or rule-based learning; see also Nalder et al., 2010). Conversely, we expected positive mood to augment the FRN component when the clarity and certainty regarding the stimulus-response associations was low ("probabilistic" learning or random condition; see also Lange, Leue, & Beauducel, 2012). Hence, following standard practice (see Frank et al., 2005), we conceived the ERN and FRN components as valid electrophysiological markers of violation of reward. If confirmed, these results might lend support to the assumption that positive mood could influence RL by changing the sensitivity to both motor (ERN) and feedback (FRN) cues signaling violation of reward (Garland et al., 2010).

2. METHODS

2.1. Participants

Thirty eight participants (undergraduate psychology students) took part in the study in exchange of 30 Euro compensation. They were randomly assigned to one of two groups: positive vs. neutral mood. They were all right-handed, with no past or current neurological or psychiatric problems, they had normal or corrected-to normal vision, and all gave written informed consent prior to the start of the experiment. The data of six participants were removed according to the following exclusion criteria. First, participants in the positive group had to show a marked increase in positive mood following the mood induction compared to the baseline (i.e., the average increase had to be above the baseline value, and there should be no decrease), while no change in positive mood was expected to take place in the neutral group. In this group, two participants were excluded since their average happiness level was higher than the mean of the positive group. Second, participants showing no learning during the main task (i.e., their learning curves did not differ from chance level) were excluded (n=3; two in the positive and one in the neutral mood group). The behavioral results obtained for the accuracy, RT and learning rate data still remained significant when including them in the statistical analyses. However, because they did not show learning, their data were deemed noisy and were therefore removed from the subsequent statistical analyses. Finally, the data of one participant was excluded because of technical problems during the recording of the EEG. The final sample consisted of 32 participants (mean age= 22.3 years, S.D. =2.4, 25 females), 16 in each mood group. The study was approved by the local ethics committee.

2.2. Mood induction

We used a previously validated mood induction procedure (MIP; see Vanlessen, et al., 2013, 2014). Mood was induced by means of a guided imagery procedure, where participants were instructed to vividly imagine reliving either a happy or neutral (depending on the group they were assigned to) autobiographical memory (Holmes, Coughtrey, & Connor, 2008; Holmes, Mathews, Dalgleish, & Mackintosh, 2006). First, the participants were trained in taking a field perspective (i.e., imagining from one's own perspective) during mental imagery. Then they had to choose an appropriate happy/neutral event, an episodic memory that happened at least a week before, and report about it. For the recall that would ensue, they were instructed to keep their eyes closed and visualize all the specificities of the memory, and to use the field perspective (Watkins & Moberly 2009, based on Holmes et al., 2008). The actual recall session was divided into two parts of 30 s, and in between participants were asked questions about different aspects of the happy/neutral memory they were imagining. Participants were blind to the real purpose of the procedure, believing that it was about remembering an event from the past as vividly as possible (and not about re-experiencing the emotion of the event). After each mood induction, participants marked on 10-cm horizontal visual analogue scales (VAS) their current level of happiness, pleasantness, and sadness, with "neutral" on one end/anchor to "as happy/pleasant/sad as I can imagine" on the other.

2.3. Probabilistic learning task

We used a probabilistic learning task previously validated by Eppinger et al. (2008). In this task, participants were asked to decipher and learn, by trial and error, several hidden stimulus-response (S-R) mappings. For each trial, participants were asked to decide, with a time limit, whether the

stimulus shown on the screen was associated with response 1 or 2. Feedback on the choice made by the participant was given following every response made.

Participants were presented with 6 visual stimuli (A-B-C-D-E-F), belonging to three conditions (unknown to the participants) that differed regarding the actual probability of the S-R mapping (100%, 80% or 50%). In the condition 100%, each stimulus of the pair was always associated with one of the two response keys, corresponding to a "deterministic" S-R mapping (i.e., response 1 was always correct for stimulus A, and response 2 for the stimulus B). In the condition 80 %, the S-R mapping was "probabilistic", given that stimulus C was associated 80% of the time with response 1 (and thus 20% of the time with the concurrent response 2), while stimulus D had a symmetric probability for the S-R mapping. Finally, in the condition 50% ("random" S-R mapping), each stimulus of the pair was associated equally often to each of the two response keys (i.e., stimuli E and F were associated 50% of the time with response 1 and 50% of the time with response 2).

Colorful line drawings (Rossion & Pourtois, 2004) were used as visual stimuli, presented against a white homogenous background on a 17-inch computer screen. These stimuli were visual objects belonging to different semantic categories (artifacts, buildings, musical instruments, clothes, vehicles, furniture). Their mean size was 7 cm width x 5 cm height, corresponding to 5 x 3,6 degrees of visual angle at 80 cm viewing distance.

The trial structure was as follows: it began with a fixation cross of 250 ms duration. Then, the stimulus was presented for 500 ms, followed by a blank screen of 300 ms. Response deadline was set to 800 ms following the onset of the visual stimulus on the screen. Then, performance feedback was presented for 500 ms. The feedback was provided in the form of a written word (in

Dutch) shown in black against a white homogenous background. This word was “goed” (correct), “fout” (incorrect), or “te traag” (too late). The inter-trial interval was constant (500 ms) and it corresponded to a blank screen, after which a new trial would ensue. Manual responses (i.e., key presses) were recorded using the Cedrus response box.

Each participant completed three blocks of 240 trials. Each block had six different stimuli, each repeated forty times. Accordingly, participants had to learn new S-R mappings in each block. Trial order within a block as well as the order of the three blocks were alternated across participants.

2.4. Procedure

First, participants were prepared for EEG recording. In order to get acquainted with the task, they completed a short practice session of 20 trials. Next, either a positive or neutral mood was induced by means of the MIP before the beginning of the first block. The same MIP was briefly rehearsed (5 minutes) halfway (after 120 trials) during the first block. The same procedure was repeated for the following two experimental blocks (i.e., the MIP was used each time at the beginning of the block and then rehearsed after the first half of trials was completed) in order to sustain the targeted mood throughout the whole experimental session. Hence, in total, participants encountered six times the MIP.

In order to strengthen the effect of positive mood, an evaluative feedback was added (rewarding in the positive mood group, and neutral in the neutral group) at the end of each block. This (bogus) feedback consisted of a small text shown on the screen, informing participants that they had to wait briefly until the computer had calculated online their learning performance up to that trial number. After a few seconds, an Excel-like scatter plot appeared on the screen, showing

them their performance level allegedly relative to a group of peers. Their score was indicated by means of a color dot. This dot was positioned systematically either higher up in the distribution of scores for participants in the positive mood group, or somewhere in the middle of the distribution for those belonging to the neutral mood group. Next to this scatter plot, a specific written message was included. It informed them to try to keep the same level of performance. Manipulation checks of mood based on visual analog scales (see results here below) confirmed that this procedure (combined with the MIP) actually produced the desired effects: an increase of positive affect in the positive mood group, with no change in affect (neither positive, nor negative) in the neutral mood group. However, we have good reasons to believe that the change in positive mood (in the positive mood group) was mainly due to the MIP and the use of guided imagery (see also Vanlessen et al., 2013, 2014), and not so much to this (infrequent) feedback manipulation that occurred only three times during the course of the experiment. After each block, participants were asked to indicate, for each of the 6 stimuli, the clarity and certainty of each of the six stimulus-response (S-R) associations, by means of a horizontal 10-cm VAS. Furthermore, they were asked to rate the amount of positive vs. negative feedback they thought they received during this last block (using a 10 cm VAS going from "exclusively negative" to "exclusively positive"), as well as how much they liked or disliked these positive vs. negative feedback when receiving them (using a Likert scale spanning from 0 to 100).

Finally, participants were asked to fill out several trait-related questionnaires: the Beck Depression Inventory (Beck, Steer, Ball, & Ranieri, 1996), the Resilience scale translated in Dutch (Portzky, De Bacquer, Audenaert, & Wagnild, 2010), and the Frederickson's list of emotions (Fredrickson, 2001). The whole experiment lasted about 2 hours.

2.5. EEG recording

EEG was recorded continuously using 64-channels by means of a Biosemi Active Two system (www. Biosemi.com). The EEG was sampled at 512 Hz, with CMS-DRL serving as the reference-ground. The EEG signal was filtered off line, using a 0.016 to 70 Hz filter (12db/oct), with a 50 Hz notch and re-referenced using the linked (average) mastoids. Individual epochs were segmented using a -/+ 500 ms interval around the response (see Aarts & Pourtois, 2010; Aarts et al., 2013; Pourtois, 2011), and -200/+1000ms around the onset of the visual feedback (see Aarts & Pourtois, 2012). Eye blinks were removed automatically via vertical ocular correction (Gratton, Coles, & Donchin, 1983), using two electrodes, placed above and below the right eye. Individual epochs were baseline corrected using the entire pre-stimulus time interval for the FRN (i.e., 200 ms) and the first 200 ms of the pre-response time-interval for the ERN (i.e., from -500 to -300 ms prior to response onset).

Artifact rejection was based on a ± 100 μV amplitude cutoff. For response locked segments, 87% of the individual segments were kept and eventually included in the averages. No significant group difference [positive: M=87.68, SEM=4.8; neutral M=86.42, SEM=4.82, $t(1,30)= 0.73$, $p>.05$] was found for this metric. For feedback–locked segments, 90% of the individual epochs were kept. No group difference was found for this metric either [positive: M= 97.57, SEM= 2.93; neutral : M= 83.42, SEM= 14.44, $F(1,30)= 1.74$, $p>.05$]. Finally, individuals epochs were averaged separately for the different conditions and subjects, and an additional low pass filter set to 30 Hz was applied on the individual averages before grand-averaging.

2.6. Data analysis

Mood manipulation. The efficiency of the change in positive (relative to neutral) mood following the MIP was assessed by means of mixed model ANOVAs with Group as between subject factor and Time (n=7) as the within subject factor.

Accuracy analyses. Accuracy data were expressed in proportions of correct responses from the total number of trials, separately for each Condition (n=3). Moreover, for each Condition separately (2 stimuli x 40 repetitions), changes of learning performance as a function of time were captured by grouping the data into four bins of equal size (i.e., 20 trials/condition) (see Eppinger et al., 2008 for a similar approach). These data were then submitted to a mixed model ANOVA with Group as between subject factor, and Condition and Bin as within subject factors.

Reinforcement-learning model. We fitted a basic Q-learning algorithm to each participant's choice data, using separate learning-rate parameters ($\propto_+$ and $\propto_-$) for positive ("correct"; modeled as 1) and negative ("incorrect"; modeled as 0) feedback, respectively. These parameter estimates complement the accuracy data by informing about the participants' tendency to learn from positive vs. negative feedback.

On each trial, the model estimates the value *Q* (i.e., probability of receiving "correct" feedback) associated with each of the 2 possible responses, separately for each stimulus. At the beginning of each block, the values of both responses for all stimuli were initialized to 0.5. Then, each time that a participant selected response *r* for stimulus *s* on trial *t*, and received positive feedback (1), the estimated value of the chosen response for that stimulus was updated according to:

$$Q_{s,r,t+1} = Q_{s,r,t} + \propto_{+}[1 - Q_{s,r,t}]$$

Similarly, when the response resulted in negative feedback (0), the estimated value of the chosen response for that stimulus was updated according to:

$$Q_{s,r,t+1} = Q_{s,r,t} + \propto_{-}[0 - Q_{s,r,t}]$$

The value of the other (unchosen) response associated with that stimulus was also updated (cf. Matsumoto, Matsumoto, Abe, & Tanaka, 2007), such that the expected values of response 1 and response 2 associated with a given stimulus sum to 1. Thus, if the value of the chosen response increases, the value of the unchosen response decreases with the same amount, and vice versa.

Choice behavior was modeled using the 'softmax' rule, with inverse-gain parameter β. With lower values of β response selection is determined more by the relative estimated values of the two responses, whereas with higher values of β response selection is more evenly distributed across the 2 responses (i.e., more random). The probability of choosing response 1 over response 2 for stimulus *s* on trial *t*, ($P_{s,1,t}$), is computed as:

$$P_{s,1,t} = \frac{e^{\frac{Q_{s,1,t}}{\beta}}}{e^{\frac{Q_{s,1,t}}{\beta}} + e^{\frac{Q_{s,2,t}}{\beta}}}$$

We fitted the model to each participant's trial-by-trial choice sequences by maximizing the log-likelihood of the observed choices. We estimated the $\propto_{+}$, $\propto_{-}$ and β parameters for the 3 probability conditions together (deterministic, probabilistic and random, see above; see Walsh & Anderson, 2014), resulting in 3 free parameters. We also fitted three other versions of the model and compared their goodness of fit using the Bayesian Information Criterion (BIC). First, we

modeled the data using a shared learning rate for positive and negative feedback. Moreover, for each of these two models (using either separate or shared learning rates for positive and negative feedback), we modeled the updating of the two response options as either dependent or independent (see also results section below).

To optimize the parameter fits we used Matlab's fmincon function (Coleman et al., 1996), a constrained nonlinear optimization algorithm, with twenty randomized starting parameter estimates. The trials in which no response was made within the 800-ms time limit were omitted.

We submitted the best fitting $\propto_+$ and $\propto_-$ parameters to a mixed model ANOVA with group (neutral vs. positive mood) as between-subject factor and feedback valence ($\mathrm{correct}/\propto_+$ vs. $\mathrm{incorrect}/\propto_-$) as within-subject factors. We also performed a group comparison on the best fitting β parameter, using an independent samples t-test.

EEG analyses. We analyzed two well-documented error-related ERP components (Eppinger et al., 2008; Frank, et al., 2005; Mies et al., 2011). For the ERN, the mean amplitude was calculated in an interval spanning 100 ms after response onset at electrode FCz. For the FRN, we used a similar 100 ms time interval (centered around the peak; 50 ms prior and 50 ms after it) and calculated the mean amplitude of this component at the same fronto-central electrode (see Eppinger et al., 2008). The FRN peak was defined as the most negative deflection arising at electrode FCz in the 230-350 ms time window following feedback onset. Importantly, for the probabilistic condition, we only included trials corresponding to the dominant S-R mapping (80%). Hence, epochs corresponding to the non-dominant S-R mapping (20%) were removed from the averages giving rise to the FRN component in this condition (probabilistic condition). For each of these two deflections separately, a mixed-model ANOVA was performed on the

average mean amplitudes with group as between subject and condition and response accuracy as within subject factors. In a second step, we computed difference waveforms by subtracting the ERP activity of incorrect from correct trials, separately for the ERN and FRN components, following standard practice (Eppinger et al., 2008).

3. RESULTS

3.1. Behavioral Results

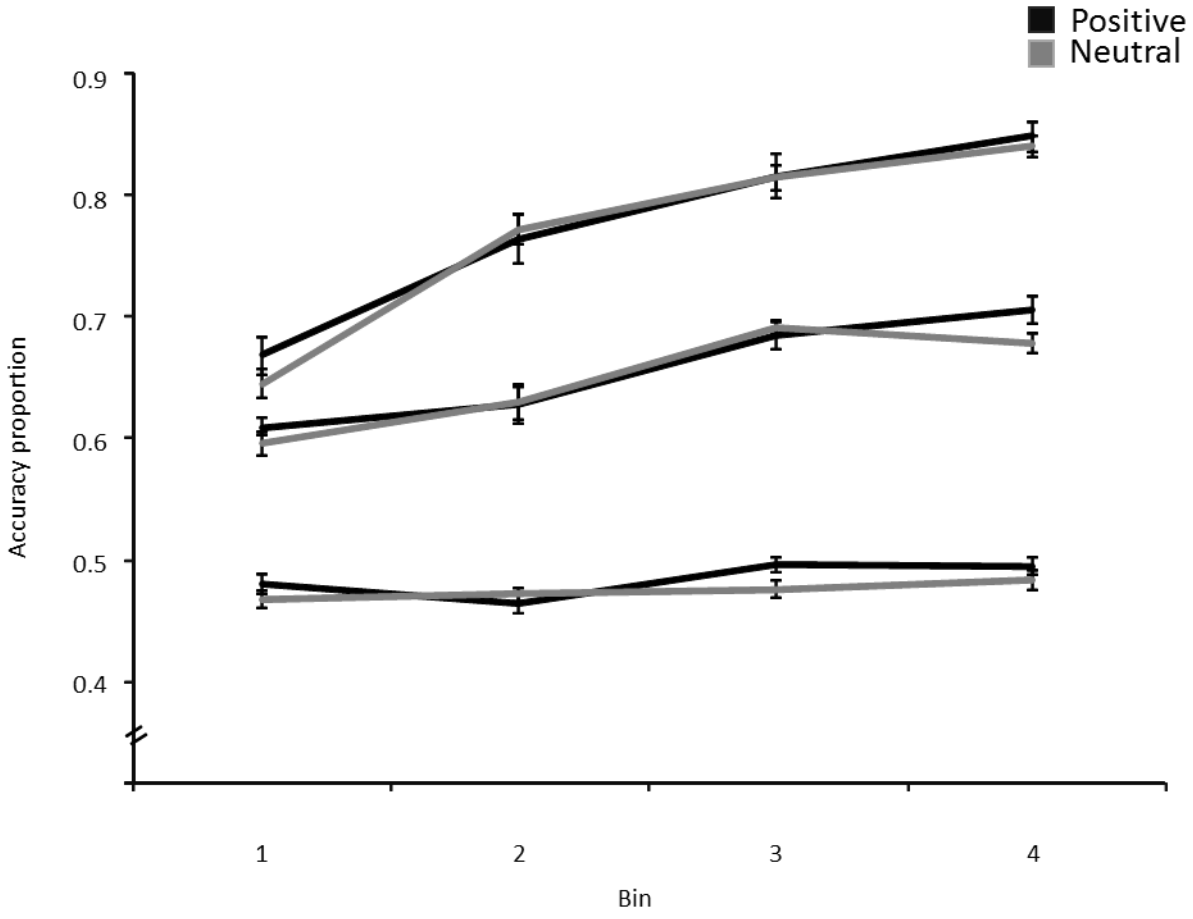

Figure 1. Accuracy data (i.e., proportion of correct responses) decomposed as a function of bin, condition and group.

Table 1. Results of the MIP. Means (± standard deviation) and results of the group comparison (based on independent-samples t-tests) between the positive (n=16) and the neutral (n=16) mood group.

Measure point	Happiness			Pleasantness			Sadness			Arousal		
	Positive	Neutral	t-test	Positive	Neutral	t-test	Positive	Neutral	t-test	Positive	Neutral	t-test
Baseline	39.9 (16.9)	44.3 (17.7)	-0.7	44.7 (14.3)	39.4 (22.1)	0.8	11.2 (15.0)	8.8 (11.2)	0.5	5.5 (2.2)	5.0 (2.7)	0.6
1	59.6 (14.8)	43.7 (19.2)	2.6*	58.3 (5.8)	37.2 (21.6)	3.1**	7.7 (12.0)	8.3 (8.3)	-0.2	4.2 (2.0)	5.9 (2.4)	-2.2*
2	61.0 (13.3)	36.4 (22.8)	3.7**	62.9 (13.1)	40.7 (23.7)	3.3**	6.9 (6.1)	7.6 (10.0)	-0.3	3.5 (1.2)	6.0 (2.0)	-4.3*
3	58.7 (13.9)	35.7 (25.5)	3.2**	59.3 (16.0)	36.9 (26.7)	2.9**	8.4 (14.5)	4.1 (8.5)	1.0	4.1 (1.7)	5.1 (2.9)	-1.2
4	61.8 (15.0)	36.9 (27.8)	3.2**	63.6 (14.8)	38.0 (29.1)	3.4**	6.6 (9.8)	5.7 (12.5)	0.2	3.7 (1.7)	5.9 (2.7)	-2.8**
5	63.9 (13.4)	32.8 (31.0)	3.7**	64.1 (15.9)	32.8 (33.3)	3.4**	6.9 (9.3)	5.8 (16.9)	0.2	3.4 (1.5)	5.4 (2.9)	-2.4*
6	63.7 (13.7)	42.9 (25.3)	2.9**	63.3 (16.3)	42.0 (26.2)	2.8**	11.9 (22.8)	7.2 (8.2)	0.8	4.3 (1.5)	5.7 (2.6)	-2.0

*p< .05, **p< .01.

Mood. The analysis of the MIP ratings (see Table 1) showed a significant Time*Group interaction for pleasantness ($F(6,180)= 2.9$, $p<.01$, $\eta^2= .1$), happiness ($F(6,180)= 6.6$, $p<.01$, $\eta^2= .2$) and arousal ($F(6,180)= 3.6$, $p<.01$, $\eta^2= .1$). With the exception of the baseline measurement where no group difference arose, the positive group had a significantly higher level of pleasantness, happiness and arousal compared to the neutral mood group for almost all subsequent measurement points following the MIP. The two groups did not differ with regards to subjective levels of sadness ($p>.8$).

Accuracy. The number of too late responses was modest (mean proportion for the positive mood group: 2.68, SD: 0.27; and for the neutral group: 2.85, SD=0.27) and did not differ between the two Groups, nor did it vary depending on condition or Bin (Bin 1: M=2.7, SD=0.3; Bin 2 : M=2.6, SD=0.3; Bin 3 : M=2.6, SD=0.2; Bin 4 : M=2.7, SD=0.3; all p's> 0.9). The analysis carried out on the proportions of correct responses (Figure 1) revealed significant main effects of Condition ($F(2,60)=230.6$, $p<.01$, $\eta^2= .9$) and Bin ($F(3,90)=26.7$, $p<.01$, $\eta^2= .5$), as well as a significant Condition* Bin interaction ($F(6,180)=15.3$, $p<.01$, $\eta^2= .8$) showing that learning curves differed across the three conditions. Learning was the quickest and accuracy the highest in the deterministic Condition, intermediate in the probabilistic Condition and merely absent in the random Condition. However, the factor Group did not interact with the main experimental factors (all p's>.1). The main effect of Group was not significant either ($p= .3$). Hence these standard accuracy data suggest a balanced learning across the two groups (Eppinger et al., 2008). The analysis performed on the RT data for correct responses revealed a significant main effect of Condition ($F(2,60)=6.8$, $p<.01$, $\eta^2= .2$), while all the other effects were non-significant (all p's>.05). Follow up t-test showed faster RTs for the deterministic (Mean RT: 390.5, SD: 26.9; $t(31)= -2.8$, $p<.01$,) and probabilistic Condition (Mean RT: 388.1, SD: 24.5, $t(31)= -3.5$, $p<.01$),

compared to the random Condition (Mean RT: 399.5, SD: 30.0). RT between the deterministic and probabilistic Condition did not differ significantly (t (31)= .7, p>.5).

We also extracted and analyzed the proportion of switches after negative feedback, as this metric has previously been related to exploration indirectly (e.g., Hills et al., 2010). This analysis confirmed that this rate varied significantly depending on reward probability/condition ($F(2,60) = 23.1$, $p<.01$, $\eta^2=.4$). Switch rate after negative feedback was higher for the random (M=6.5, SD=1.6) than either the deterministic (M=4.7, SD=1.6; t(31)=5.8, p<.001) or the probabilistic condition (M=4.7, SD=1.7; t(31)= 6.12, p<.001). The difference between the deterministic and the probabilistic condition was not significant (t(31)=.07, p=.94). This effect did not interact with group/mood (F(2,60)=1.4, p=.25).

Learning rate. The ANOVA performed on the estimated learning-rate (α) parameters revealed significant main effects of Feedback Valence ($F(1,30) = 116.65$, $p< .01$, $\eta^2= .7$), and group ($F(1,30)= 6.2$, $p<.05$, $\eta^2= .2$), indicating that the learning rate was higher for positive than for negative feedback, and that the positive mood group had overall higher learning rates than the neutral mood group. Follow-up contrasts (using independent samples T-tests) showed a higher learning rate in the positive (M=0.31, SD=0.15) than in the neutral mood group (M=0.21, SD=0.10) following positive feedback (t(30)= 2.20, p<.05). Likewise, a higher learning rate in the positive (M=0.03, SD=0.03) than in the neutral mood group (M=0.00, SD=0.01) was also evidenced following negative feedback (t(30)=2.61, p<.05).

The group comparison performed on the inverse-gain parameter (β) revealed no significant effect (p>.05; positive mood group: M=0.58, SD=0.50; neutral mood group: M=0.057, SD=0.018).

We also computed the Bayesian Information Criterion (BIC) for this model (Model I; see Table 2). We also compared the BIC obtained for this model (allowing different learning rates for positive and negative feedback) to another model (Model III in Table 2) where the learning rate was the same for positive and negative feedback. However, as can be seen from Table 2 we obtained a better fit (i.e., significantly lower BIC) with the former compared to the latter model.

Table 2. BIC (mean ± standard deviation in parenthesis) obtained for the four different models, separately (see text for details). Model I: separate learning rates for positive and negative feedback, and inter-dependent response alternatives; Model II: separate learning rates for positive and negative feedback, and independent response alternatives, Model III: one learning rate for positive and negative feedback, and inter-dependent response alternatives; Model IV: one learning rate for positive and negative feedback, and independent response alternatives.

	BIC	Pairwise comparisons	t-Test
Model I	712.26 (119.86)	Model I vs. Model II	-4.78**
Model II	721.68 (115.68)	Model I vs. Model III	-8.50**
Model III	773.43 (94.80)	Model I vs. Model IV	-8.35**
Model IV	774.96 (95.78)	Model II vs. Model III	-0.72

*p< .05, **p< .01.

In our model, the values of both the chosen and the unchosen response options were updated after each feedback, in such a way that their values always summed to 1. So after each choice, the value of the non-selected response increased/decreased with the same amount as the value of the chosen response decreased/increased. We felt is was appropriate to use this 'double-update' model (cf. Matsumoto et al., 2007) as participants were to learn which of the 2 responses was most often associated ("correct") with a given stimulus, and these stimulus-response associations remained stable during the experiment (i.e., evidence in favor of one response automatically also provides evidence against the other response, and v.v.). However, this way of

modeling may potentially make the interpretation of positive vs. negative learning rates more difficult. Accordingly, we also computed two alternative 'single-update' models (in which only the value of the chosen response was updated; see Models II and IV in Table 2). However, for these two models, the BIC was significantly larger (reflecting a worse model fit) than for our main model.

Ratings. The mixed-model ANOVA carried out on the clarity ratings showed a significant main effect of Condition ($F(2,60)=730.3$, $p< .01$, $\eta^2= 1\ .0$). The analysis performed on the certainty ratings revealed a main effect of Condition alike ($F(2,60)=141.5$, $p< .01$, $\eta^2= .9$). These results showed that perceived clarity and response certainty went down monotonically when S-R probability decreased. However, positive mood did not influence these subjective ratings (p's> .9). Clarity was higher for the deterministic condition (M=83.0, SD=5.0) compared to the probabilistic condition (M=75.7, SD=6.7; $t(31)=6.8$, $p<.01$), while clarity for the latter was also significantly higher than for the random condition (M=37.4, SD=5.6; $t(31) =24.9$, $p<.01$). Likewise, certainty was the highest for the deterministic condition (M=65.3, SD=8.4) and intermediate for the probabilistic condition (M=48.8., SD=9.0; $t(31) = 9.5$, $p<.01$). Certainty was also significantly higher in the probabilistic compared to the random condition (M=36.3, SD=8.5; $t(31)=8.5$, $p<.01$).

In contrast, positive mood did reliably affect the subjective experience of participants regarding the amount of positive vs. negative feedback they thought they had received during the experiment, as well as their like-dislike reactions to them (see Table 3). Participants in the positive mood group reported having received significantly more positive than negative feedback, compared to the neutral group ($t(30)=3.7$, $p< .01$). Moreover, the former participants felt happier than the latter when receiving positive feedback ($t(30)=3.4$, $p< .01$), while they were also more

unhappy when receiving negative feedback when compared to the neutral mood group ($t(30)=3.9$, $p< .01$).

Table 3. Subjective ratings (regarding the number of positive feedback received throughout the experiment, as well as the like-dislike reaction to them) and BDI scores. For each measure, the mean (± standard deviation in parenthesis), as well as the direct group comparison (using unpaired t-test) are provided.

Estimate of the positive feedback received			Liking positive feedback			Disliking negative feedback			BDI		
Positive	Neutral	t-Test	Positive	Neutral	t-Test	Positive	Neutral	t-Test	Positive	Neutral	t-Test
45.4 (9.7)	30.2 (13.6)	3.7**	47.6 (8.7)	36.8 (9.2)	3.4**	50.7 (7.7)	40.1 (7.8)	3.9**	7.8 (7.0)	6.1 (5.7)	0.7

*p< .05, **p< .01.

Table 4. Mean ERP activity (± 1 standard deviation) for each condition and accuracy level, separately for each ERP component and group. Results of the direct pairwise comparisons (degrees of freedom:15) between the two accuracy levels (correct vs. incorrect), using post hoc t-tests.

ERP component	Condition	Mood					
		Positive			Neutral		
ERN		Correct	Incorrect	t-test	Correct	Incorrect	t-test
	Deterministic	-0.1 (2.6)	-4.5 (2.7)	8.9**	0.5 (3.2)	-2.4 (3.3)	5.9**
	Probabilistic	-0.8 (2.6)	-1.6 (1.8)	2.1*	0.6 (3.3)	-0.7 (3.2)	3.1*
	Random	-1.3 (2.8)	-1.6 (2.2)	0.9	-0.3 (2.8)	-0.2 (2.7)	-0.4
FRN							
	Deterministic	-0.4 (2.0)	-0.1 (4.2)	-0.4	-0.9 (2.1)	-2.6 (2.2)	2.7*
	Probabilistic	0.7 (3.2)	-0.5 (3.2)	4.3**	-0.2 (2.2)	-1.1 (2.6)	2.1*
	Random	1.7 (3.3)	0.2 (3.7)	3.2**	0.2 (1.7)	-1.1 (2.0)	2.8**

*p< .05, **p< .01.

Questionnaires. Independent t-tests failed to reveal significant group differences for the BDI (Table 3) or resilience scale (all p's>.6). The two groups also had a balanced performance regarding the way they filled out the Frederickson's list of emotions (i.e., no significant group differences for the frequencies of the emotional words chosen, all p's >.8).

3.2. ERP Results

The analysis carried out on the ERN mean amplitudes showed a significant main effect of Condition ($F(2,60)= 10.8$, $p< .01$, $\eta^2= .3$) and Accuracy ($F(1,30)= F(1,30)=79.6$, $p< .01$, $\eta^2= .7$). The two way Condition* Accuracy interaction ($F(2,60)= 45.2$, $p< .01$, $\eta^2= .6$) and the three way Group*Condition*Accuracy interaction ($F(2,60)= 3.8$, $p< .05$, $\eta^2= .1$) were significant, indicating that positive mood modulated error processing at the response level in a condition and accuracy

specific way (see Figures 2, 3 and 5, and Table 4). To disentangle this interaction effect, we computed difference scores for the ERN (incorrect – correct) per condition, and analyzed how this measure varied depending on condition. Importantly, this ANOVA yielded a significant Group x Condition interaction ($F(2,60)=3.8$, $p<.05$, $\eta^2= .1$). A direct group comparison showed that the ERN was larger in the positive ($M=-4.5$, $SD=2.0$) compared to the neutral mood group ($M=-2.9$, $SD=1.9$) in deterministic condition ($t(30)=-2.3$, $p<.05$) only (Figure 2). The two groups did not differ from each other in the two other conditions (all ps>.2).

The analysis performed on the mean amplitudes of the FRN (Figure 4) yielded a significant main effect of Condition ($F(2,60)=13.6$, $p<.01$, $\eta^2= .3$) and Accuracy ($F(1,30)=11.6$, $p<.01$, $\eta^2= .3$). The interaction between these two factors did not reach significance ($F(2,60)=1.1$, $p=.34$). Moreover, positive mood did not reliably influence these effects (all ps >.1).

Finally, to verify that positive mood influenced the ERN component selectively (while leaving the FRN component unaffected), we run an omnibus ANOVA with ERP Component (either ERN or FRN), Condition and Group as factors. This analysis showed a significant three-way interaction ($F(2,60)=6.6$, $p<.01$, $\eta^2= .2$), confirming that positive mood selectively influenced the ERN component in a condition-specific way (i.e., for the deterministic condition).

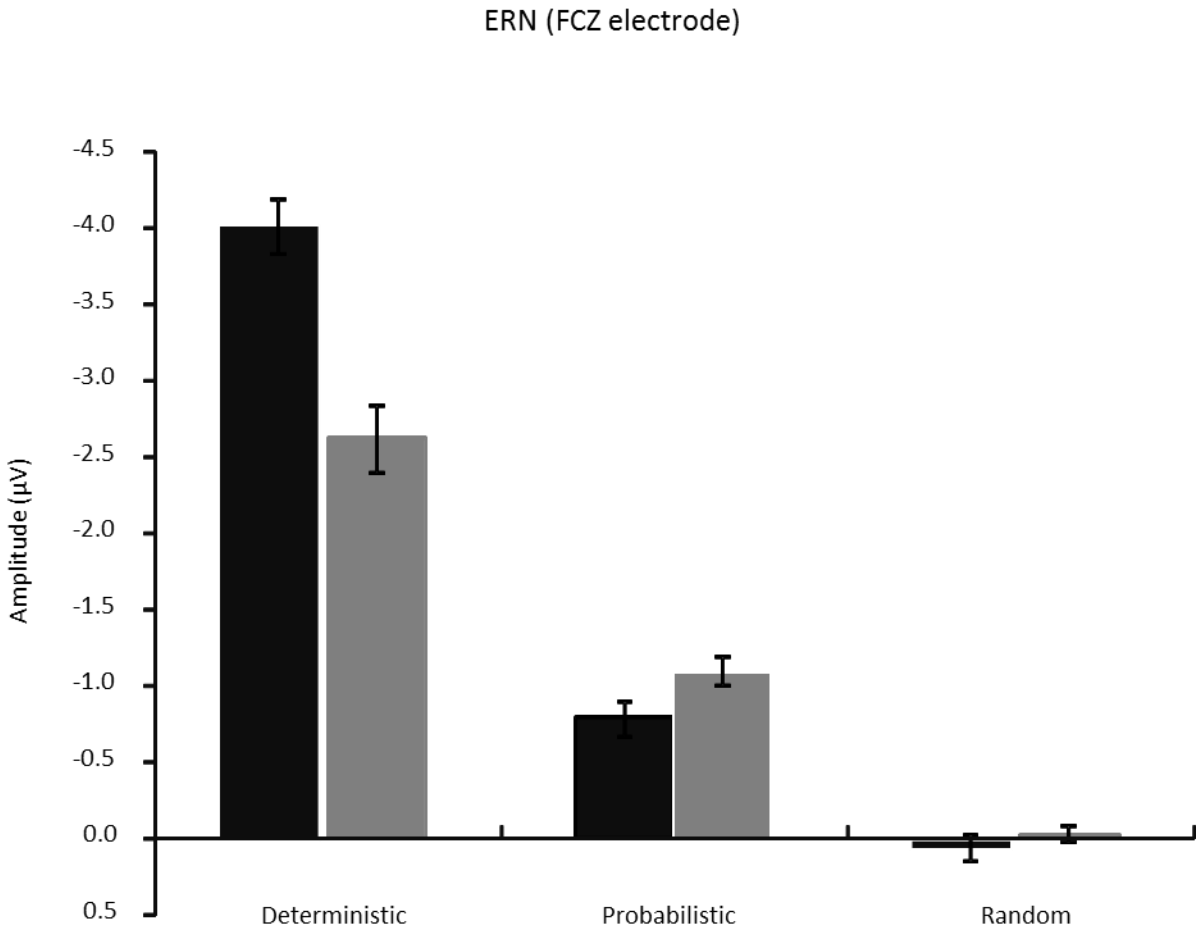

Figure 2. Mean amplitude in μV (+ 1 S.E.M) of the difference score (incorrect minus correct) for the ERN component at the electrode FCz, shown separately for each group and condition.

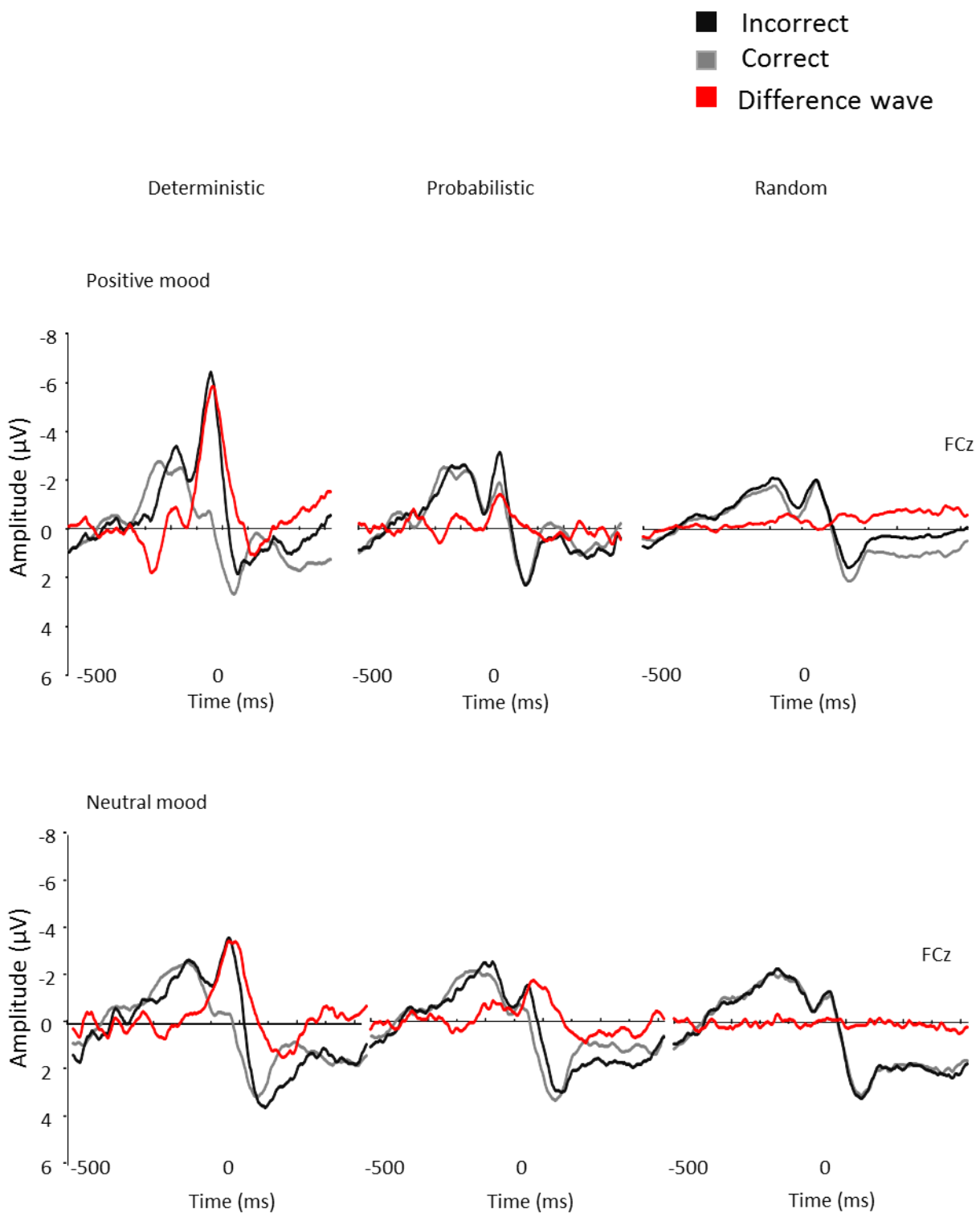

Figure 3. Grand average ERP waveforms for the response-locked ERP data (electrode FCZ), separately for each group, condition and accuracy level.

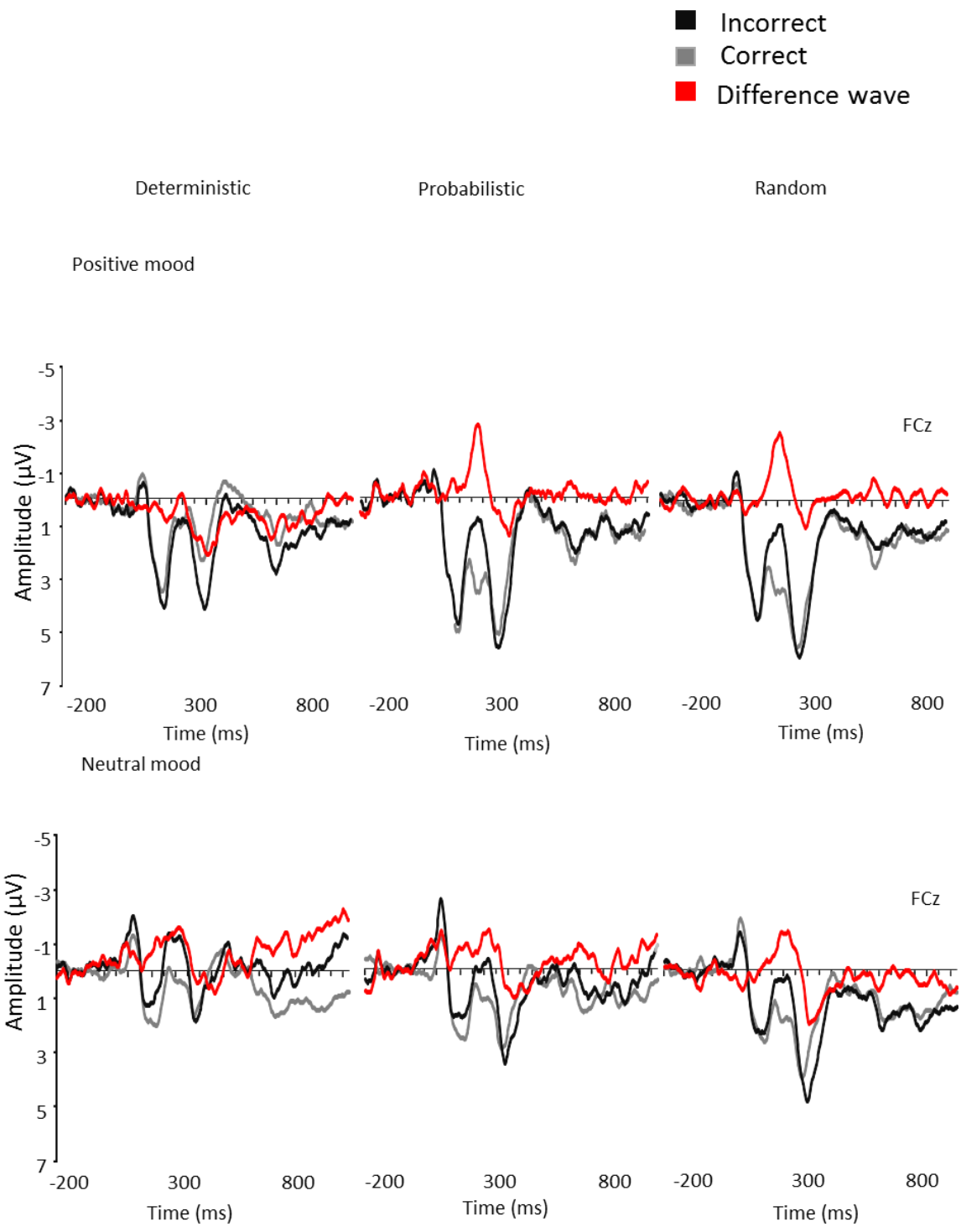

Figure 4. Grand average ERP waveforms for the feedback-locked ERP data (electrode FCZ), separately for each group, condition and accuracy level.

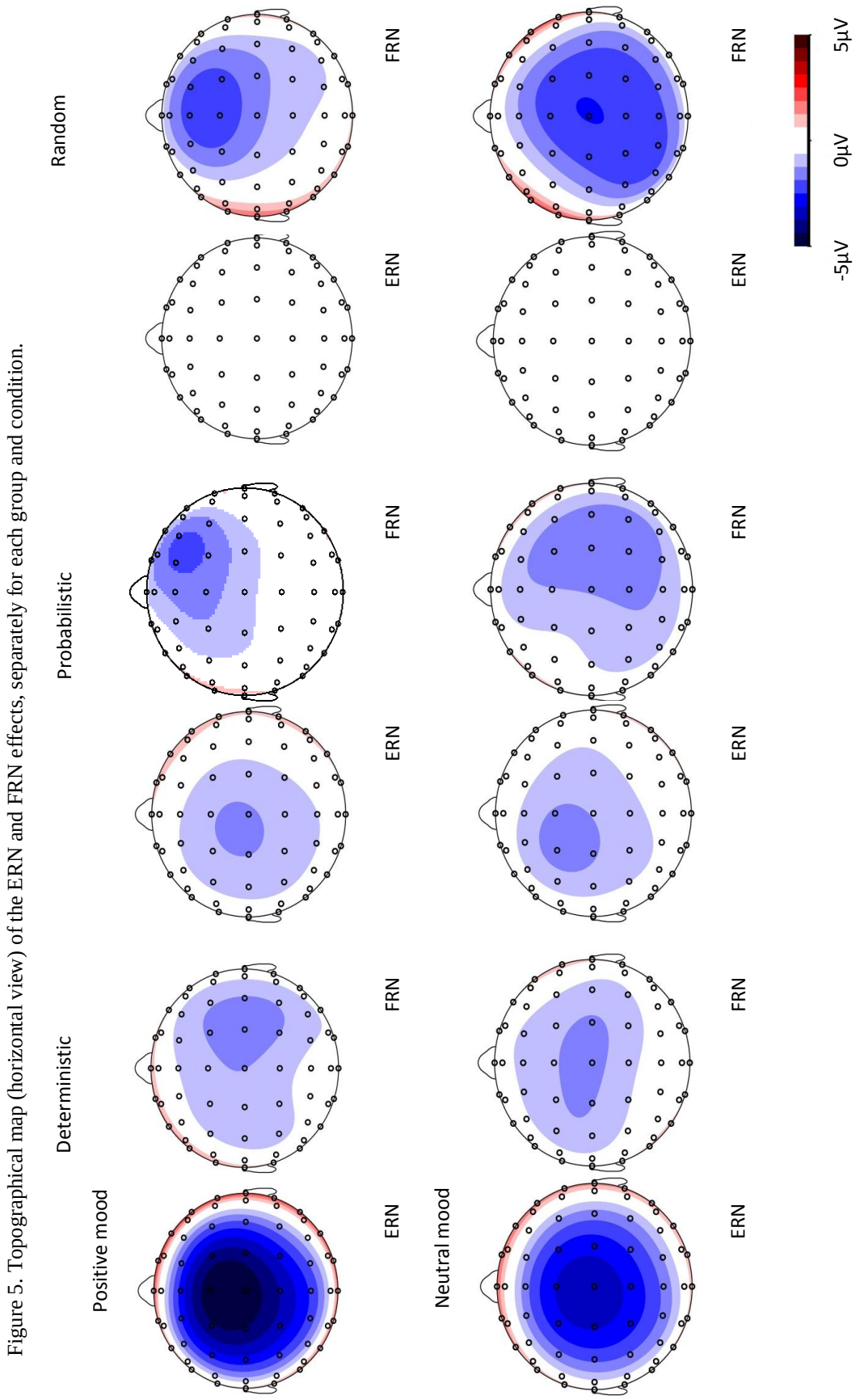

Figure 5. Topographical map (horizontal view) of the ERN and FRN effects, separately for each group and condition.

4. DISCUSSION

In this study we put to the test the assumption that the transient experience of positive mood might alter the way positive vs. negative feedback is used to guide behavior during a standard probabilistic learning task. More specifically, we sought to assess whether positive mood might be accompanied by a change in the way positive reinforcers are processed (Eshel & Roiser, 2010). To this aim, we used a previously validated probabilistic learning task (Eppinger et al., 2008) enabling us to titrate changes in learning both at the behavioral and electrophysiological (ERP) levels. Learning was not only characterized in terms of accuracy and speed, but also using formal learning parameters extracted from a RL model. A reliable and sustained change in positive affect was brought about in one group of participants using a previously validated MIP (Vanlessen et al., 2013). Their behavioral performance and ERP components were compared to another group of participants, for which the mood was kept neutral throughout the task. The results of this study confirm that the current internal state of the participant (here positive mood) can substantially alter the way feedback information is used during probabilistic learning.

Whereas positive mood did not merely alter the accuracy and speed of learning at the behavioral level (when compared to a control neutral mood condition), a strength of our study was that we characterized changes in RL as a function of this specific mood state using formal learning parameters, besides these standard behavioral measures. Using this procedure, we found that being in a happy mood increased the learning rate, irrespective of the valence of the (preceding) feedback triggering this updating effect. Interestingly, Frank, Moustafa, Haughey, Curran and Hutchinson (2007) previously argued that an increased learning rate could reflect a

rapid and enhanced adaptation to changing outcomes, which proved to be an important component of the internalization process at play during learning. Because this enhancement in the learning rate as a function of positive mood was found for both positive and negative feedback, our new results indirectly suggest that positive mood might foster the (rapid or optimal) internalization of the task rule during probabilistic learning, via modulations of both reward (positive/correct feedback) and punishment (negative/incorrect feedback) processing. By contrast the randomness of participants' choices[2] (reflected in the inverse-gain, or β, parameter) was not influenced by positive mood. More generally, these results emphasize the added value of model-based analysis (relative to the standard accuracy or RT data) to reveal subtle group differences related to the encounter and transient experience of positive mood in our study. As a matter of fact, this dissociation suggests that learning might be qualitatively different between the two mood groups, even though a rough quantitative estimate of learning (based on accuracy aggregated across several trials) fails to reveal reliable group differences (see Figure 1).

Positive mood was also associated with a larger ERN effect than neutral mood, for the deterministic condition selectively. Classically, the ERN component is thought to reflect prediction error based on the processing of internal (perhaps motor-related) cues (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000), as opposed to external ones in the case of the FRN (Eppinger et al. 2008). Accordingly, our ERN results are compatible with the notion of a larger RL signal present at the response level (Holroyd & Coles, 2002), for this specific condition (deterministic) and group (positive mood). With this scenario, positive mood would therefore heighten prediction error encoding (likely within the dorsal ACC and interconnected midbrain

[2] Previous studies have linked this parameter (beta) to exploration (Jepma & Nieuwenhuis, 2011). However, the current experimental paradigm (a two-alternative forced-choice task with stable S-R probabilities) was not designed to measure the exploration-exploitation tradeoff. Hence, in the current study this parameter mainly reflects the randomness of participants' choices.

dopaminergic brain structures; see Holroyd & Coles, 2002), if and only if reward probability (acquired through learning) is actually high. These results suggest therefore that positive mood does not simply create an unspecific shift in motivation, or a uniform boost in reward processing (across all conditions in an undifferentiated manner). Instead, the significant interaction effect between mood/group and condition found at the level of the ERN is compatible with the idea that positive mood may boost reward prediction error during learning when the probability of getting a reward (as computed based on an internal/motor representation – ERN component) is high.

Interestingly, earlier ERP studies consistently found a larger ERN component in participants characterized by internalizing traits (sub clinical) or disorders (clinical), or negative affect broadly defined (Vaidyanathan, Nelsonm & Patrick, 2012; Olvet & Hajcak, 2008; Aarts & Pourtois 2010; Moser, Moran, Schroder, Donnellan, & Yeung, 2013). By comparison, our new electrophysiological results show that the transient experience of positive affect can be associated with an augmented ERN effect alike, casting doubt on the idea that an overactive ERN necessarily reflects a stable endophenotype for internalizing traits or disorders along the main psychopathology continuum (Olvet & Hajcak, 2008). Therefore, this neurophysiological effect may reflect a transient change and increase following the induction of positive mood in the dopaminergic-dependent neural system (connecting the basal ganglia to the rostral cingulate zone) thought to underlie reward prediction error and in turn the generation of the ERN component (see Holroyd & Coles, 2002). Noteworthy, this conjecture was already made by Ashby et al. (1999) about 15 years ago, but never or rarely put to the test directly. This theory assumes that during periods of mild positive affect, there is a concomitant increased dopamine release in the mesocorticolimbic system, and also the nigrostriatal system, which can in turn

influences performance on a variety of cognitive tasks. Accordingly, our new neurophysiological results are important because they lend empirical support to this dominant model, while they also show that the ERN component is not only sensitive to negative affect or punishment (trait characteristics), but also changes in positive mood (state effect).

It should be noted however that this somewhat subtle change in the way positive feedback (or reward) is processed after the induction of positive mood did not seem to be related to (or explained by) obvious modifications in the way participants actually perceived and experienced "retrospectively" this probabilistic learning task. In both groups, participants' subjective visibility and certainty of the S-R associations varied according to the experimental conditions in the predicted direction. Notwithstanding a lack of group difference for these two variables, an interesting finding was that participants in the positive mood group reported to have received more positive feedback than in the neutral mood group, as well as to like them more, even though we could not assess whether this effect was condition specific. Future studies are needed to test whether these learning (rate) and neurophysiological effects (ERN and FRN components) might predict reliable changes in the way probabilistic learning is eventually experienced by the participants (after the induction of positive mood).

For the FRN component, we found that this feedback-locked ERP activity showed a clear amplitude modulation depending on the varying reward probabilities (i.e., the FRN component was large for the random and probabilistic conditions, but small for the deterministic condition; see also Eppinger et al., 2008), irrespective of the mood change however. Hence, positive mood appears to exert selective influences on reward prediction error during probabilistic learning, namely at the motor (internal) level, while leaving the feedback (external) level unchanged. Although these ERP results reveal some specificity in the way positive mood may shape reward

processing during learning, we note nonetheless that the results obtained for the learning parameters (RL) point to a general increase in learning rate following positive feedback (after the induction of positive mood). Moreover, we did not find any significant correlation between amplitude variations at the level of the ERN and this learning parameter (nor for the inverse-gain parameter). This discrepancy between the ERP and computational modeling results might stem from methodological reasons (e.g., averaging of several trials regardless of trial history in the case of FRN vs. updating of beliefs depending on the most recent feedback in the case of learning rate).

A few limitations have to be noted. As is often the case with research on positive mood, it remains difficult to disentangle effects of positive valence from arousal per se. Our results clearly show that participants in the positive mood group not only experienced more happiness and pleasantness relative to the neutral mood group, but also more arousal (at a subjective level though). However, earlier studies found that valence, but not arousal, did contribute to adaptive goal-directed behavior in front of conflicting situations or events (van Steenbergen, Band, & Hommel, 2010). On the other hand, arousal per se appears unlikely to account for the complex interaction effects on the FRN and ERN components in our study. Another caveat concerns the apparent dissociation found between clear cut effects of positive mood on learning rate, the FRN as well as ERN component on the one hand, but on the other hand, the lack of obvious group differences regarding response accuracy during the probabilistic learning task. Three elements are here noteworthy. First, accuracy data were computed in line with Eppinger et al. (2008), where bins (i.e., average of performance across several trials) were eventually used. Although this method is probably suited to grasp “gross” changes in learning depending on time and condition (see Figure 1), it is probably not sensitive enough to detect finer alterations occurring on a trial by

trial basis during probabilistic learning, as a function of positive mood, as our modeling and ERP data reveal. Second, our results also show that participants did not perform at ceiling at the end of the experimental session. It is possible that group differences on accuracy would appear when using more trials. Third, the absence of any group difference regarding accuracy (bearing also in mind that participants in the two groups showed normal learning curves depending on the experimental condition; see Figure 1) was actually an asset in the present case, because we could then compare learning parameters and ERP data when the amount of response errors (and hence negative feedback) was actually balanced between the two groups. In other words, the present results cannot be explained by a general asymmetry or imbalance in accuracy during the probabilistic learning task across the two groups.

To conclude, the results of this study provide evidence for a selective modulatory effect exerted by positive mood on probabilistic learning. At the behavioral level, positive mood heightens the learning rate (as opposed to merely increasing "exploration" for example), in particular when positive feedback can be used to guide learning directly. At the electrophysiological level, this effect is primarily expressed in the way (reward) prediction error is processed (at the response levels) when the certainty and visibility regarding the S-R associations are (relatively) high (i.e., in the deterministic condition). An open question for future research concerns the potential costs and benefits for the organism associated with this change in reward prediction error following the induction of positive mood. For example, whether or not an enhanced or sharper reward prediction error mechanism during the encounter or experience of positive mood may foster learning (compared to a neutral or negative mood), remains to be established at the empirical level. In this study, we failed to observe improvements in learning at the behavioral level following the induction of positive mood, although positive mood influenced

the learning rate and shaped a well-known electrophysiological marker of reward prediction error (i.e., the ERN).

5. ACKNOWLEDGEMENTS

This worked is supported by a Concerted Research Action Grant from Ghent University awarded to RDR (#BOF10/GOA/014) and by the Belgian Science Policy, Interuniversity Attraction Poles program (P7/11). GP is funded by the European Research Council (Starting Grant #200758) and Ghent University (BOF Grant #05Z01708). The authors have no conflict of interest to declare.

6. REFERENCES

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective & Behavioral Neuroscience*, 10(4), 479–92. doi:10.3758/CABN.10.4.479.

Aarts, K., & Pourtois, G. (2012). Anxiety disrupts the evaluative component of performance monitoring: An ERP study. Neuropsychologia, 50, 1286– 1296.

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, 13(5):960-73.

Ashby, F.G., Isen, A.M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence in cognition. *Psychological Review*, *106 (3)*, 529-550.

Beck, A.T., Steer, R.A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories- IA and II in psychiatric outpatients. *Journal of Personality Assessment*, *67(3)*, 588-597.

Bediou, B., Koban, L., Rosset, S., Pourtois, G., & Sander, D. (2012). Delayed monitoring of accuracy errors compared to commission errors in ACC. *NeuroImage*, 60(4), 1925–36. doi:10.1016/j.neuroimage.2012.02.010.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–21. doi:10.1038/nn1954.

Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. *The British Journal of Clinical Psychology / the British Psychological Society*, 22 (Pt 3), 173–181. doi:10.1111/j.2044-8260.1983.tb00598.x.

Carbonnell, L., & Falkenstein, M. (2006). Does the error negativity reflect the degree of response conflict? *Brain research*, 1095(1), 124–130. doi:10.1016/j.brainres.2006.04.004.

Coleman, T. F., & Li, Y. (1996). An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*. doi:10.1137/0806023.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., & Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876-879. doi: 10.1038/nature04766.

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46(2)*, 521-539.

Eshel, N., & Roiser, J.P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*, 188-124.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78, 447–455. doi:10.1016/0013-4694(91)90062-9.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51(2-3), 87–107. doi:10.1016/S0301-0511(99)00031-9.

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 of the United States of America*, 104(41), 16311–16316.

Frank, M. J., Woroch, B.S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47,* 495-501.doi: 10.1016/j.neuron.2005.06.020.

Fredrickson, B.L. (2001). The role of positive emotions in positive psychology. The broaden-and-build-theory of positive emotions. *American Psychologist*, *56(3)*, 218-226.

Fredrickson, B.L. (2004). The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci, 359(1449),* 1367–1378. doi:10.1098/rstb.2004.1512.

Garland, E. L., Fredrickson, B., Kring, A.M., Johnson, D.P., Meyer, P.S., & Penn, D.L. (2010). Upward spirals of positive emotions counter downward spirals of negativity: Insights from the broaden-and-build theory and affective neuroscience on the treatment of emotion dysfunctions and deficits in psychopathology. *Clinical Psychology Review*, *30*, 849-864.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error-Detection and Compensation. *Psychological Science*, 4(6), 385–390. Retrieved from <Go to ISI>://A1993MF61300008.

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

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *NeuroImage*, 47(4), 2023–2030. doi:10.1016/j.neuroimage.2009.05.064.

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

Hills, T.T., Todd, P.M., & Goldstone, R.L. (2010). The central executive as a search process: priming exploration and exploitation across domains. *J Exp Psychol Gen*., 139(4), 590-609. doi: 10.1037/a0020666.

Holmes, E.A., Coughtrey, A.E., & Connor, A. (2008). Looking at or through rose-tinted glasses? Imagery perspective and positive mood. *Emotion, 8(6)*, 875-879.

Holmes, E.A., Mathews, A., Dalgleish, T., & Mackintosch, B. (2006). Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behavioral Therapy*, *37(3)*, 237-247.

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

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off : Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23(7)*, 1587-96. doi: 10.1162/jocn.2010.21548.

Lange, S., Leue, A., & Beauducel, A. (2012). Behavioral approach and reward processing: Results on feedback-related negativity and P3 component. *Biological Psychology*, 89(2), 416–425.

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci*., 10(5):647-56.

Mies, G.W., van der Veen, F.M., Tulen, J.H.M., Birkenhäger, T.K., Hengeveld, M.W., & van der Molen, M.W. (2011). Drug-free patients with major depression show an increased electrophysiological response to valid and invalid feedback. *Psychological Medicine 41*, 2515-2525. doi:10.1017/S0033291711000778.

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

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, 7(August), 466. doi:10.3389/fnhum.2013.00466.

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better mood and better performance. Learning rule-described categories is enhanced by positive mood. Psychological

Science : *A Journal of the American Psychological Society* / APS, 21(12), 1770–1776. doi:10.1177/0956797610387441

Nieuwenhuis, S., Slagter, H. A., von Geusau, N. J. A., Heslenfeld, D. J., & Holroyd, C. B. (2005). Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *The European Journal of Neuroscience*, *21*(11), 3161–3168. doi:10.1111/j.1460-9568.2005.04152.x.

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clin Psychol Rev*, 28(8), 1343–1354. doi:10.1016/j.cpr.2008.07.003.

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

Portzky, M., Wagnild, G., De Bacquer D., & Audenaert, K. (2010). Psychometric evaluation of the Dutch Resilience Scale RS-nl on 3265 healthy participants: A confirmation of the association between age and resilience found with the Swedish version. *Scandinavian Journal of Caring Sciences*, *1*, 86-92. doi:10.1111/j.1471-6712.2010.00841.x.

Pourtois, G. (2011). Early Error Detection Predicted by Reduced Pre-response Control Process: An ERP Topographic Mapping Study. *Brain Topography*, 23(4),403-22.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33(2)*, 217-236.

Sallet, J., Camille, N., & Procyk, E. (2013). Modulation of feedback-related negativity during trial-and-error exploration and encoding of behavioral shifts. *Frontiers in Neuroscience, 7*, 209. doi:10.3389/fnins.2013.00209.

Sutton, R., & Barto, A. G. (1998). Reinforcement learning: an introduction. Cambridge, MA: MIT Press.

Vaidyanathan, U., Nelson, L. D., & Patrick, C. J. (2012). Clarifying domains of internalizing psychopathology using neurophysiology. *Psychological Medicine*. doi:10.1017/S0033291711001528.

Van Steenbergen, H., Band, G. P., & Hommel, B. (2010). In the mood for adaptation: how affect regulates conflict-driven control. *Psychological Science*, *21*(11), 1629–1634. doi:10.1177/0956797610385951.

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2013). Positive emotion broadens attention focus through decreased position-specific spatial encoding in early visual

cortex: Evidence from ERPs. *Cognitive, Affective & Behavioral Neuroscience, 13(1),* 60-79. doi: 10.3758/s13415-012-0130-x.

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2014). Feeling happy enhances early spatial encoding of peripheral information automatically: electrophysiological time-course and neural sources. *Cognitive Affective and Behavioral Neuroscience*,14 (3), 951-969. doi: 10.3758/s13415-014-0262-2.

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America*, 108(47), 19048–19053. doi:10.1073/pnas.1117189108

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

Walsh, M.M., & Anderson, J.R. (2014). Navigating complex decision spaces: Problems and paradigms in sequential choice. *Psychological Bulletin*, 140(2):466-86. doi: 10.1037/a0033455

Watkins, E.R., & Moberly, N.J. (2009). Concreteness training reduces dysphoria: A pilot proof-of-principle-study. *Behavioral Research and Therapy, 47(1),* 48-53. doi: 10.1016/j.brat.2008.10.014.

CHAPTER 4

What is in the feedback? Effect of induced happiness vs. sadness on probabilistic learning with vs. without exploration[1]

According to dominant neuropsychological theories of affect, emotions signal salience of events and facilitate a wide spectrum of response options or action tendencies. Valence of an emotional experience is pivotal here, as it alters reward and punishment processing, as well as the balance between safety and risk taking, which can be translated into changes in the exploration-exploitation tradeoff during reinforcement learning (RL). To test this idea, we compared the behavioral performance of three groups of participants that all completed a variant of a standard probabilistic learning task, but who differed regarding which mood state was actually induced and maintained (happy, sad or neutral). To foster a change from an exploration to an exploitation-based mode, we removed feedback information once learning was reliably established. Although changes in mood were successful, learning performance was balanced between the three groups. Critically, when focusing on exploitation-driven learning only, they did not differ either. Moreover, mood valence did not alter the learning rate or exploration per se, when titrated using complementing computational modeling. These results challenge the assumption that mood valence alone is enough to create strong shifts in the way exploitation or exploration is eventually carried out during (probabilistic) learning. In this context, we discuss the possibility that both valence and arousal are actually necessary components of the emotional mood state to yield changes in the use and exploration of incentives cues during RL.

[1] Bakic, J., De Raedt, R., Jepma, M., & Pourtois, G. (2014). What is in the feedback? Effect of induced happiness vs. sadness on probabilistic learning with vs. without exploration. Manuscript in preparation

1.INTRODUCTION

Many students would agree that studying for an exam after a heartbreak is a particularly hard thing to do. On the other hand, some professors would argue that, if one wants to excel, also the happy, falling in love phase is best avoided altogether. Even if it was possible, would it really be best if emotions were somehow hushed and kept at bay in order to learn? Or is it possible to turn both, sadness and happiness, into learning fuel? If so, which one works better?

Emotions are complex, multi-faceted phenomena that signal importance of events and guide actions to maximize benefits and minimize damage. From an evolutionary perspective, development of such variety and richness of emotions as we know today enabled more flexible, more adaptive functioning, and ultimately, a wider spectrum of response options (Lang & Bradley, 2010). In that sense, valence of emotional experiences plays a pivotal role: positively valenced emotions (such as happiness, joy, amusement, pleasantness) are hypothesized to signal safety and instigate creativity, exploration, playfulness, and risk-taking. In contrast, negatively valenced emotions, such as fear, sadness, anger, disgust, or frustration, signal threat and the need to recruit additional resources to deal with potential harm or loss (Ashby, Isen, & Turken, 1999; Fredrickson, 2004; Isen, 1993).

From this initial premise, different expectations about the effects of positive and negative emotions can be derived. Most research on the topic was performed by inducing mood, using different strategies (including movie clips, images, music, autobiographical pieces, guided imagery; see Martin, 1990; Westermann, Stahl, & Hesse, 1996). Moods are considered to be more enduring and milder than emotions, and are not directed towards a certain entity, but are rather “non-focal” (Bolte & Goschke, 2010). Mood effects have been

examined in the area of creative thinking (Isen, Daubman, & Nowicki, 1987; Isen, Johnson, Mertz, & Robinson, 1985; Isen, 1984), attention (Huntsinger, 2012; Vanlessen, Rossi, De Raedt, & Pourtois, 2014), or cognitive control (Fröber & Dreisbach, 2014; van Steenbergen, Band, & Hommel, 2010). Effects of mood on performance are rather mixed though, with some studies showing that positive mood does not necessarily translate into improved (behavioral) performance, or that negative mood does not automatically lead to detrimental effects for cognition and behavior.

As a matter of fact, reinforcement learning (RL) is a particularly good candidate as a process to be modulated by mood, because by definition, it relies directly on the processing of positive vs. negative information or incentives to achieve a goal at hand. Stimulus-response associations (S-R) are being formed in a trial-and-error fashion, based on externally provided feedback, reward or punishment, about one's own actions (Sutton & Barto, 1998). If current mood provides an emotional context for the learning situation, then it could change the salience of error and reward, or how threatening and appetitive they are eventually perceived, and in turn processed. A performance monitoring system in charge of learning will value opportunities and threats in surroundings differently depending on the current state and the needs of the organism. Most theories of RL argue that performance optimization is based on the right amount of exploitation of rewarding options, and exploration of less known, but potentially even more beneficial alternatives (Aston-Jones & Cohen, 2005; Behrens, Woolrich, Walton, & Rushworth, 2007; Cohen, McClure, & Yu, 2007; Jepma & Nieuwenhuis, 2011). These two concurrent processes, exploration and exploitation, have complementary benefits or functions: while it is important to keep current goals in mind and not allow for distractions (i.e., favor exploitation), it is at the same time important to keep the environment in check for potential changes that might reliably influence performance (i.e., foster exploration). We are hypothesizing that this trade-off between exploration and

exploitation might be susceptible to changes in the current mood state of the participant. If positive mood really does lead to more exploration of less known options, while negative mood leads to more stringent focus (Bolte & Goschke, 2010), oriented towards negative information, then this effect should be visible in the exploration-exploitation trade-off during RL. Moreover, we can expect that mood manipulation will also influence the usage of positive and negative feedback for response updating, such that positive groups update more based on the positive feedback, while in the sad group avoiding negative feedback might be more salient determinant of future behavior.

Along these lines, Unger, Kray and Mecklinger (2012) have shown that, in a learning paradigm, inducing a feeling of performance-related failure changes the strategy towards more error-driven behavioral control, while it concurrently increases electrophysiological measure of error monitoring (at the level of the error related negativity, ERN). In a previous study (Bakic, Jepma, De Raedt, & Pourtois, 2014), using a similar probabilistic learning paradigm, we also showed that, compared to neutral, happy mood led to increase in ERN when learning was deterministic, and was associated with an increase in the learning rate (but not exploration). Here, we used a probabilistic learning task (Eppinger, Kray, Mock, & Mecklinger, 2008), in which some of the S-R associations that had to be learned were probabilistic in nature, meaning that the feedback was probabilistic in a pre-determined, probabilistic way. This situation usually creates a certain amount of uncontrollability/ uncertainty that learning agents have to overcome in order to optimize their learning performance. This uncertainty might be dealt with differently depending on the current emotional state of the participant. Even though in this study we showed that, by modulating the current mood of the participant, we were able to modulate the learning rate (accompanied by change on the electrophysiological level as well, more precisely for the ERN component; see Bakic et al., 2014), we failed however to show that positive mood led to clear benefits or

impairments in the actual learning performance during this probabilistic task (i.e., happy participants did not perform better or worse than neutral participants during RL).

Accordingly, in the present study, we sought to adapt this experimental paradigm (see Bakic et al., 2014) in a way that would allow us to maximize the chance to capture such a difference at the behavioral level between the two groups. For this purpose, in addition to a standard initial learning phase (consisting of trials made each time of stimulus-response-feedback associations) that is identical as in previous research (Bakic et al., 2014; Eppinger et al., 2008), we added a second phase, where feedback on task performance was now omitted. In other words, during this second phase (when learning was already established), we changed the trial structure in such a way that a standard stimulus-response-feedback sequence was changed to stimulus-response one, preventing participants from using feedback information (and thereby exploration) to guide learning. At this point, participants can only use stored value estimates, and are no longer able to track state transitions of value. Based on the results obtained in our previous study, we know that S-R associations have already been formed in the first phase when participants move to the second one. More specifically, internalization of task rules took place and externally provided feedback is no longer necessary to perform the task accurately. This is consistent with the assumption that exploration of different response alternatives is no longer needed, and the (direct) exploitation of the acquired knowledge can be carried out. Using this specific manipulation, we wanted to examine whether creating such clear-cut difference between the exploration and exploitation stage of the task could eventually lead to a clearer difference at the behavioral level between the two groups than in our previous study (Bakic et al., 2014). Additionally, other than comparing only happy and neutral mood, in this study, we added a third group of participants who received a similar mood induction procedure (MIP) but with a sad content. This way, we could assess whether

sadness might perhaps produce opposite effects on RL compared to happiness, suggesting that mood valence plays a critical role in triggering these changes during RL.

To summarize, the goal of this study is to test the effects of inducing happy, neutral or sad mood on RL (operationalized using a probabilistic learning task; see Bakic et al., 2014), when this process is broken down into two consecutive phases: an initial learning phase relying on the use of external feedback information to guide learning (where both exploration and exploitation are used in synergy), followed by a second phase where feedback is omitted (and exploitation alone is encouraged). Our experimental design involves comparisons of three groups of participants differing from one another regarding the actual mood state induced (happy, neutral, or sad), but using the same guided imagery procedure (Holmes & Mathews, 2010). Based on our first study (Bakic et al., 2014), during the first part of the task, we do not expect to find group differences in rough measures of learning (e.g., accuracy). We surmise, however, that the happy group will show a higher learning rate (with no change in exploration), compared to the neutral (and/or sad) group. If sad mood influences learning performance in an opposite manner compared to the positive mood group, then we can expect a lower learning rate in this group compared to the two other ones (happy and neutral). Additionally, we predict that during the second phase of the task where feedback information is no longer available, happy mood could be associated with a better learning performance than either neutral or sad mood given that this specific mood state could bolster internalization of the task rules and in turn exploitation (Nadler, Rabi, & Minda, 2010).

2.METHODS

2.1. Participants

Fifty two participants (undergraduate psychology students) took part in the study in exchange for course credits. They were randomly assigned to one of the three mood groups: happy, neutral or sad mood. They were all right-handed, with no past or current neurological or psychiatric problems, they had normal or corrected-to normal vision, and all gave written informed consent prior to the start of the experiment. The data of seven participants were removed according to the following exclusion criteria (see also Bakic et al., 2014). First, one participant was excluded in the happy group because of the lack of a marked increase in happy mood following the MIP compared to the baseline (i.e., the average increase was not above the baseline value). Likewise, one participant was excluded from the sad group due to the lack of a marked increase in sadness relative to the baseline mood measurement. Finally, two participants were excluded from the neutral mood group because their average happiness level was higher than the mean of the happy group, whereas no change in mood was expected to take place in this control group. Second, participants showing no learning during the main task (i.e., their learning curves did not differ from chance level) were excluded as well (n=3; one in each group). Note that the behavioral results obtained for the accuracy, RT and learning rate data remained unchanged when including them in the statistical analyses. However, because they did not show learning, their data were deemed noisy and were therefore removed from the subsequent statistical analyses. The final sample consisted of 45 participants (mean age= 20.62 years, S.D. =2.29, 29 females), 14 in the happy, 14 in the neutral, and 17 in the sad mood group. The study was approved by the local ethics committee.

2.2. Mood induction

We used a previously validated MIP (Vanlessen, Rossi, De Raedt, & Pourtois, 2013; Vanlessen et al., 2014; Bakic et al., 2014). Mood was induced by means of a guided imagery procedure, where participants were instructed to vividly imagine reliving either a happy, neutral, or sad (depending on the group they were assigned to) autobiographical memory (Holmes, Coughtrey, & Connor, 2008; Holmes, Mathews, Dalgleish, & Mackintosh, 2006). First, the participants were trained in taking a field perspective (i.e., imagining from one's own perspective) during mental imagery. Then they had to choose an appropriate happy/neutral/sad event, an episodic memory that happened at least a week before, and to report explicitly about it. For the recall that would ensue, they were instructed to keep their eyes closed and visualize all the specificities of the memory, and to use the field perspective (Watkins & Moberly 2009, based on Holmes et al., 2008). The actual recall session was divided into two parts of 30s each, and in between participants were asked questions about different aspects of the happy/neutral/sad memory they were imagining. Participants were blind to the real purpose of the procedure, believing that it was about remembering an event from the past as vividly as possible (and not about re-experiencing the actual emotion of this specific event). After each mood induction, participants marked on 10-cm horizontal visual analogue scales (VAS) their current level of happiness, pleasantness, and sadness, with "neutral" on one end/anchor to "as happy/pleasant/sad as I can imagine" on the other. Arousal was measured on a 9-point Likert scale.

2.3. Probabilistic learning task

A modified version of the probabilistic learning task previously validated by Eppinger et al. (2008) was used in this study, with the first phase of the experiment being the same as in previous studies (Bakic et al., 2014; Eppinger et al., 2008). During this phase, participants were asked to decipher and learn, by trial and error, several hidden stimulus-response (S-R)

mappings. For each trial, participants were asked to decide, with a time limit, whether the stimulus shown on the screen was associated with response 1 or 2. Visual feedback regarding the actual choice made by the participant was given following each and every response made during this first phase. Upon completion of this first phase, participants move to the second phase of the experiment, where feedback information was now omitted but task instructions remained unchanged.

Participants were presented with 6 visual stimuli (A-B-C-D-E-F), belonging to three conditions (unknown to the participants) that differed regarding the actual probability of the S-R mapping (100%, 80% or 50%). In the condition 100%, each stimulus of the pair was always associated with one of the two response keys, corresponding to a "deterministic" S-R mapping (i.e., response 1 was always correct for stimulus A, and response 2 for the stimulus B). In the condition 80 %, the S-R mapping was "probabilistic", given that stimulus C was associated 80% of the time with response 1 (and thus 20% of the time with the concurrent response 2), while stimulus D had a symmetric probability for the S-R mapping. Finally, in the condition 50% ("random" S-R mapping), each stimulus of the pair was associated equally often to each of the two response keys (i.e., stimuli E and F were associated 50% of the time with response 1 and 50% of the time with response 2).

Colorful line drawings (Rossion & Pourtois, 2004) were used as visual stimuli, presented against a white homogenous background on a 17-inch computer screen. These stimuli were visual objects belonging to different semantic categories (artifacts, buildings, musical instruments, clothes, vehicles, furniture). Their mean size was 7 cm width x 5 cm height, corresponding to 5 x 3,6 degrees of visual angle at 80 cm viewing distance.

For the first phase of the experiment, the trial structure was as follows: it began with a fixation cross of 250 ms duration. Then, the stimulus was presented for 500 ms, followed by a

blank screen lasting 300 ms. Response deadline was set to 800 ms following stimulus onset. Then, performance feedback was presented for 500 ms. The feedback was provided in the form of a written word (in Dutch) shown in black against a white homogenous background. This word was "goed" (correct), "fout" (incorrect), or "te traag" (too late). The inter-trial interval was constant (500 ms) and it corresponded to a blank screen, after which a new trial ensued. Manual responses (i.e., key presses) were recorded using the Cedrus response box. After participants completed 240 trials (6 stimuli x 40 repetitions), trial structure changed. During this second phase, trial structure was the same as for the first phase of the experiment, with the following notable exception: instead of a meaningful feedback response (informing the participant about his actual accuracy) appearing on the screen after each manual, there were three "x" signs shown as visual feedback (in the same location and for the same duration as in the previous phase of the experiment). In this way, trial structure remained identical, the only difference being the lack of informative external feedback. Participants performed 120 trials (6 stimuli x 20 repetitions) with this uninformative feedback ("no feedback" condition here after).

Each participant completed two blocks of 360 trials (240 with and 120 without feedback). Each block had six different stimuli. Accordingly, participants had to learn six new S-R associations in each block. Trial order within a block as well as the order of the two blocks were alternated across participants.

2.4. Procedure

In order to get acquainted with the task, participants first completed a short practice session of 20 trials. Next, either a happy, neutral, or sad mood was induced by means of the MIP before the beginning of the first block. In order to sustain the targeted mood throughout the whole experimental session, the same MIP was briefly rehearsed (5 minutes) during both blocks, every 120 trials (corresponding to two bins; see data analysis here below). The same

procedure was also repeated during the break between two blocks. Hence, in total, participants encountered the MIP 7 times. Additionally, participants assigned to the sad group received one more MIP at the very end of the experiment. This MIP consisted of actively reliving a happy past memory episode (very much like what was made in the happy mood group) in order to make sure that these participants (sad mood group) would not leave the lab with a lingering sad mood. Self-ratings after this MIP showed that happiness ratings for the sad mood group went back to a neutral mood baseline, being in turn comparable to those of the neutral mood group after the experiment (see Figure 1).

In order to strengthen the effect of mood, an evaluative feedback was added (rewarding in the happy mood group, neutral in the neutral group, and mildly negative in the sad mood group) at the end of each block. This (bogus) feedback consisted of a small text fragment shown on the screen, informing participants that they had to wait briefly until the computer had calculated online their learning performance up to that trial number. After a few seconds, an Excel-like scatter plot appeared on the screen, showing them their performance level allegedly relative to a group of peers. Their score was indicated by means of a color dot. This dot was positioned systematically either higher up in the distribution of scores for participants in the happy mood group, somewhere in the middle of the distribution for those belonging to the neutral mood group, and slightly lower for the participants in the sad mood group. Next to this scatter plot, a specific written message was included. It informed them to try to keep the same level of performance or perform better if possible. Manipulation checks based on VASes (see results below) confirmed that this procedure (combined with the MIP) actually produced the desired effects: an increase of happiness in the happy mood group, with no change in affect (neither happy, nor sad) in the neutral mood group, and a decrease of happiness in the sad mood group. However, we have good reasons to believe that the change in happiness (or sadness in the sad mood group) was mainly due to the MIP and the use of

guided imagery (see also Vanlessen et al., 2013, 2014), and not so much to this (infrequent) feedback manipulation that occurred only four times during the course of the experiment. Moreover, after each block, participants were asked to indicate, for each of the 6 stimuli, the clarity and certainty of each of the six stimulus-response (S-R) associations, by means of a horizontal 10-cm VAS. Furthermore, they were asked to rate the amount of positive vs. negative feedback they thought they received during this last block (using a 10 cm VAS going from "exclusively negative" to "exclusively positive"), as well as how much they liked or disliked these positive vs. negative feedback when receiving them (using a Likert scale spanning from 0 to 100).

Finally, participants were asked to fill out two trait-related questionnaires: the Beck Depression Inventory (Beck, Steer, Ball, & Ranieri, 1996), and the Resilience scale translated in Dutch (Portzky, De Bacquer, Audenaert, & Wagnild, 2010). The whole experiment lasted for about 2 hours.

2.5. Data analysis

Mood manipulation. The efficiency of the increase/decrease in happy mood (relative to the neutral group) following the MIP was assessed by means of mixed model ANOVAs with group (n=3) as between subject factor and time (n=7) as within subject factor.

Accuracy analyses. Accuracy data were expressed in proportions of correct responses from the total number of trials, separately for each condition (n=3). Moreover, for each condition separately (2 stimuli x 40 repetitions), changes of learning performance as a function of time were captured by grouping the data into bins of equal sizes (i.e., 20 trials/condition), (see Eppinger et al., 2008; Bakic et al., 2014 for a similar approach). These data were then submitted to a mixed model ANOVA with group (n=3) as between subject factor, and condition (n=3) and bin (n=4) as within subject factors. Additionally, in order to compare

possible differences between learning with vs. without feedback, we averaged the scores of bins 3 and 4 together (exploitation and exploration), and bin 5 and 6 (exploitation only), and submitted these mean values to a mixed model ANOVA with group as between subject factor, and condition and phase as within subject factors. Where necessary, Greenhouse-Geisser correction for sphericity was performed, and corrected p values were reported, together with uncorrected eta square measure of effect size.

RL model. For the first phase of the experiment (with feedback information provided to the participants), we used two complementary measures based on the modeling procedure described previously in Bakic et al. (2014). We computed first the learning rate parameter (α), which determines the impact of the most recent feedback on the expected S-R-F associations , so that higher learning rates result in more fluctuations in response behavior from trial to trial, and lower learning rates result in more stable S-R-F associations and response behavior. We also calculated a second parameter, β, which is a "noise" parameter that reflects how random choices are (and hence provides an indirect measure of exploration).

3. RESULTS

Mood. The analysis of the MIP ratings showed a significant Time*Group interaction for pleasantness ($F(7.25, 152.22)= 8.53$, $p<.01$, $\eta2= 0.29$), happiness ($F(9.45, 198.40)= 11.18$, $p<.01$, $\eta2= 0..35$), sadness ($F(9.84, 81.19)= 6.98$, $p<.01$, $\eta2= 0.25$), and arousal ($F(8.22, 172.53)= 2.36$, $p<.05$, $\eta2= 0.10$). The main effect of group was also significant for pleasantness ($F(2, 42)= 42.02$, $p<.01$, $\eta2= 0.67$), happiness ($F(2,42)= 43.71$, $p<.01$, $\eta2= 0.68$), sadness ($F(2,42)= 24.71$, $p<.01$, $\eta2= 0.54$), and arousal ($F(2,42)= 5.79$, $p<.05$, $\eta2= 0.22$).

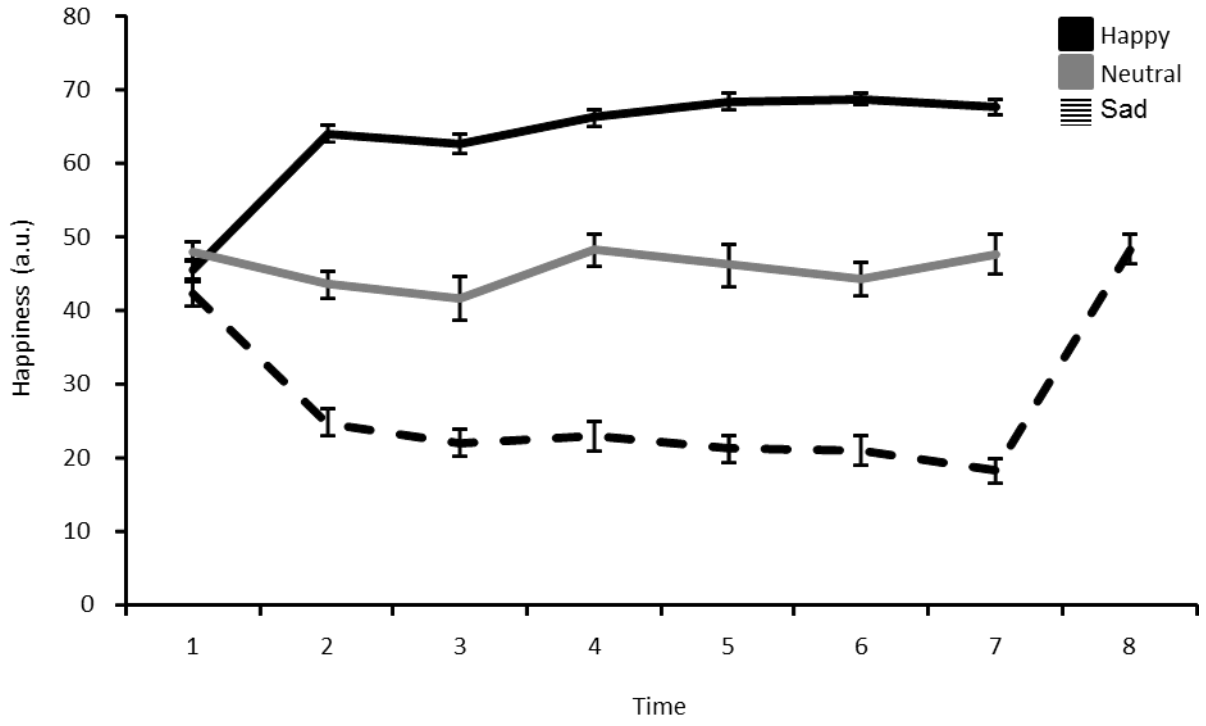

Figure 1. Happiness ratings shown separately for the happy, neutral and sad mood group as a function of time. Each point represents the mean and the error bar 1 standard error of the mean.

Independent samples t-tests for direct comparison between the happy and the neutral, and the neutral and the sad group showed that there were no significant differences at baseline, whereas in each subsequent measure (hence following a MIP each time) the happy group showed an increase compared to the neutral group, while the sad group showed a marked decrease in levels of pleasantness (Table 1). The same was true for happiness ratings (Figure 1, Table 2). Independent t-test comparisons for the sadness ratings (Table 3) showed that the happy and the neutral group had comparable, low and unchanged levels of sadness, whereas in the sad group sadness increased after the first MIP and stayed significantly higher than in the neutral group throughout the duration of the experiment (except for the last measurement following a positive MIP meant to restore a neutral mood state in this group; see methods). A paired sample t-test for the sad group showed that after the final happy MIP ($M=14.5$, $SD=12.2$), this group had significantly lower sadness scores than after the last sad MIP ($M=31.36$, $SD=17.64$), ($t(16)=4.51$, $p<.01$). At the same time, happiness scores increased

significantly from last sad MIP (M=18.43, SD= 14.14) to the happy MIP (M=48.43, SD=17.09), (t(16)=-7.04, p<.01). The happy and the sad group did not differ significantly in arousal levels (Table 4), but the sad group showed somewhat lower arousal scores than the neutral group.

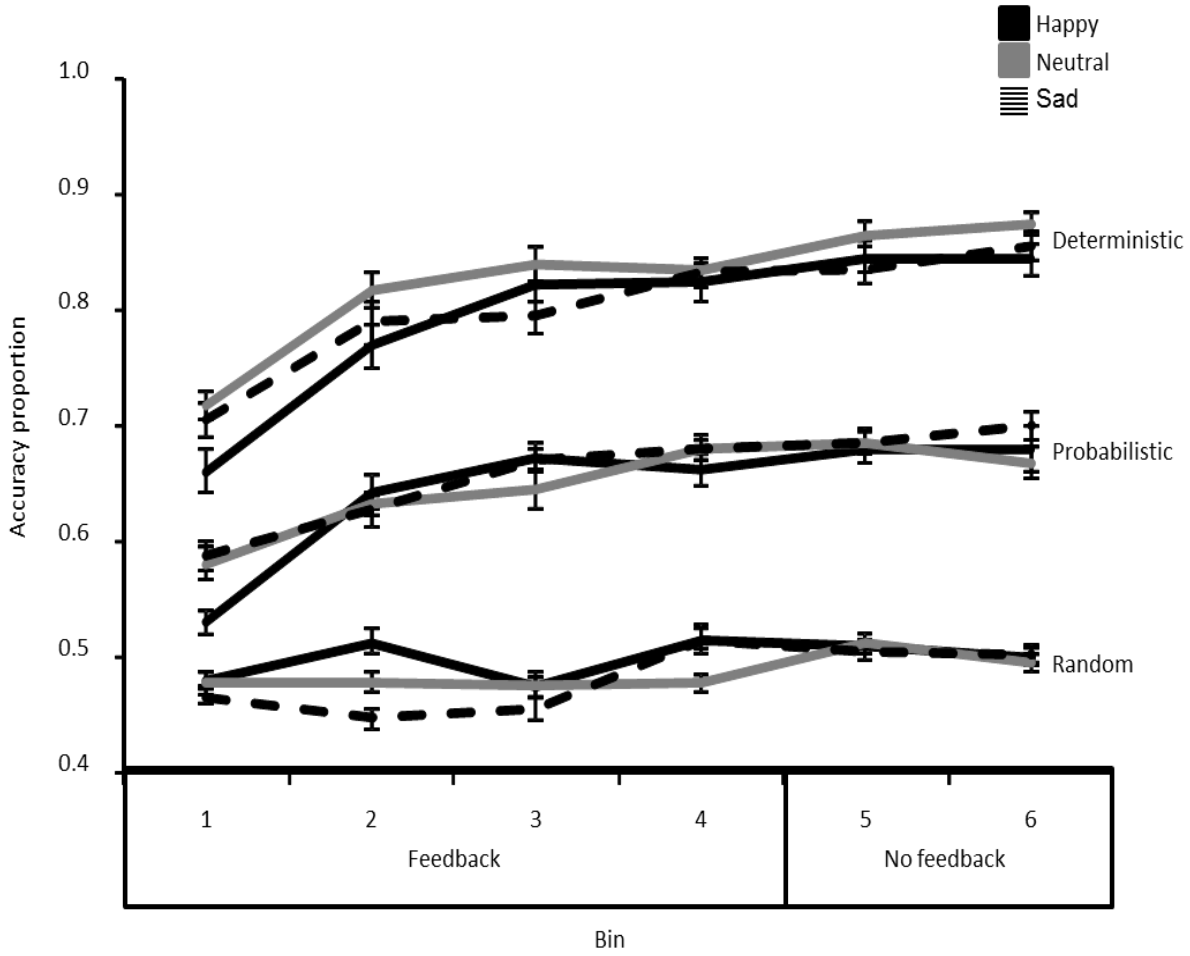

Figure 2. Accuracy data (i.e., proportion of correct responses) decomposed as a function of bin, condition and group. The error bar corresponds to 1 standard error of the mean.

Too late responses. The number of too late responses was modest (M=1.64, SD= 0.93) and not different between the three mood groups (p's>.05). There was a significant Group*Bin interaction (F(10, 210) =3.02, p<.01, η^2=0.13), showing that neutral group had increased number of too late responses compared to the other two groups for the final two bins without

feedback. Additionally, there was a significant main effect of condition ($F(2, 84) = 9.64$, $p<.01$, $\eta^2=0.19$), and bin ($F(3.71, 155.95) = 2.94$, $p<.05$, $\eta^2=0.10$).

Accuracy. Results showed a significant Condition*Bin interaction ($F(10,420)=6.13$, $p<.01$, $\eta^2=.13$), as well as significant main effects of Condition ($F(1.70, 71.89)=334.96$, $p<.01$, $\eta^2=.89$), and Bin ($F(3.92, 164.70)=34.27$, $p<.01$, $\eta^2=.45$). This interaction indicated, as can be seen from Figure 2, that accuracy was higher and that learning was steeper in the deterministic than in the probabilistic condition, while there was no learning (across time) whatsoever in the random condition.

Next, we averaged the two last Bins in the learning phase with feedback and compared performance to the two bins of the "no feedback" phase in order to assess whether learning still increased once feedback information on task performance had been removed, and exploitation only was required (Figure 2). This analysis showed a significant main effect of Phase ($F(1,42)=10.02$, $p<.01$, $\eta^2=.19$), suggesting that learning still increased reliably after removing the feedback. There was also a significant main effect of Condition ($F(2,84)=365.28$, $p<.01$, $\eta^2=.90$). However, there were no significant group-related effects (all p's>.05).

Reaction times (RTs) for correct responses. Results showed a significant effect of condition ($F(1.95, 81.95)=3.78$, $p<.05$, $\eta^2=0.08$). Paired samples t-tests showed that the random condition (M=412.61, SD=39.32) had marginally significantly larger RTs than the probabilistic (M=408.62, SD=35.66), ($t(44)=-1.85$, $p=.07$) and significantly larger than the deterministic condition (M=406.78, SD=35.22), ($t(44)=-3.05$, $p<0.01$). These two latter conditions did not differ significantly from each other (p.> .05). Unexpectedly, a significant main effect of group was evidenced too, ($F(2,42)=10.82$, $p<.01$, $\eta^2=0.34$). Follow-up independent t-tests showed that the sad group (M=382.81, SD=34.20) had overall

significantly faster RTs (see Figure 3) than the happy (M=424.22, SD=19.81), (t(29)=4.00, p<.01) and the neutral mood group (M=426.68, SD=32.35), (t(29)=3.64, p<.01).

Learning rate. The analysis showed a significant main effect of feedback valence (F(1,42)=172.78, p<.01, η^2=0.80), showing that this parameter was overall larger for positive than negative feedback, as already found in our previous study (Bakic et al., 2014). Other effects were all non-significant (all p's >.05; see Figure 4).

Exploration parameter. The one way ANOVA showed no significant group differences in exploration (Figure 5).

Table 1. Results of the Pleasantness scores. Means (+1 Standard Deviation) and results of the group comparison (based on independent-samples t-tests) between the Happy and Neutral (df=26) or the Neutral and Sad (df=29) mood Group.

Measure point	Pleasantness				
				t-test	
	Happy	Neutral	Sad	Happy vs. Neutral	Neutral vs. Sad
Baseline	47.32 (15.91)	46.38 (15.93)	43.90 (13.28)	0.16	0.47
1	65.01 (7.14)	47.86 (12.49)	32.90 (12.23)	4.46**	3.36**
2	64.57 (10.18)	45.25 (21.56)	24.67 (11.96)	3.03**	3.36**
3	66.36 (8.38)	51.14 (13.97)	27.68 (14.72)	3.49**	4.52**
4	67.12 (10.78)	49.06 (19.19)	22.47 (15.30)	3.07**	4.30**
5	68.35 (6.80)	45.74 (20.24)	24.39 (15.08)	4.96**	3.36**
6	67.46 (7.71)	46.62 (21.47)	21.60 (13.15)	3.42**	4.00**

*p<.05, **p<.01

Table 2. Results of the Happiness scores. Means (+1 Standard Deviation) and results of the group comparison (based on independent-samples t-tests) between the Happy and Neutral

Measure point	Happiness				
				t-test	
	Happy	Neutral	Sad	Happy vs. Neutral	Neutral vs. Sad
Baseline	45.59 (10.84)	48.05 (10.04)	42.55 (15.55)	-0.62	1.14
1	64.16 (9.19)	43.64 (13.33)	24.75 (15.03)	4.74**	3.66**
2	62.87 (9.79)	41.73 (22.46)	22.12 (15.69)	3.23**	2.87**
3	66.28 (8.26)	48.29 (16.76)	23.05 (15.38)	3.60**	4.37**
4	68.56 (7.88)	46.27 (20.40)	21.30 (16.61)	3.81**	3.86**
5	68.81 (5.97)	44.28 (17.64)	21.07 (16.91)	4.93**	3.73**
6	67.88 (7.52)	47.71 (20.22)	18.43 (14.14)	3.50**	4.74**

*p<.05, **p<.01

Table 3. Results of the Sadness scores. Means (+1 Standard Deviation) and results of the group comparison (based on independent-samples t-tests) between the Happy and Neutral (df=26) or the Neutral and Sad (df=29) mood Group.

Measure point	Sadness				
				t-test	
	Happy	Neutral	Sad	Happy vs. Neutral	Neutral vs. Sad
Baseline	13.59 (10.48)	8.82 (8.38)	11.64 (9.47)	1.33	-0.87
1	6.39 (7.54)	7.04 (5.80)	31.67 (14.68)	-0.25	-5.90**
2	5.41 (7.32)	9.32 (12.17)	33.39 (15.73)	-1.03	-4.68**
3	7.21 (7.95)	8.06 (6.70)	33.51 (17.83)	-0.31	-5.04**
4	6.28 (7.25)	5.66 (6.98)	28.06 (18.06)	0.23	-4.37**
5	5.31 (6.61)	9.35 (11.25)	27.60 (19.17)	-1.16	-3.14**
6	8.82 (10.02)	7.94 (6.32)	31.36 (17.64)	0.28	-4.72**

*p<.05, **p<.01

Table 4. Results of the Arousal scores. Means (+1 Standard Deviation) and results of the group comparison (based on independent-samples t-tests) between the Happy and Neutral (df=26) or the Neutral and Sad (df=29) mood Group.

Measure point	Arousal				
				t-test	
	Happy	Neutral	Sad	Happy vs. Neutral	Neutral vs. Sad
Baseline	4.79 (1.12)	5.07 (1.82)	3.65 (1.46)	-0.50	2.43*
1	5.93 (1.77)	5.21 (1.37)	4.29 (1.86)	1.19	1.54
2	5.93 (2.09)	5.64 (1.69)	3.47 (1.59)	0.40	3.68**
3	5.64 (2.41)	5.21 (1.37)	3.76 (1.48)	0.58	2.81*
4	5.71 (2.64)	4.64 (1.60)	4.12 (1.65)	1.30	0.89
5	5.93 (2.20)	4.93 (1.69)	4.18 (1.74)	1.35	1.21
6	5.86 (2.31)	3.86 (1.46)	3.94 (1.64)	2.73*	-0.15

*p<.05, **p<.01

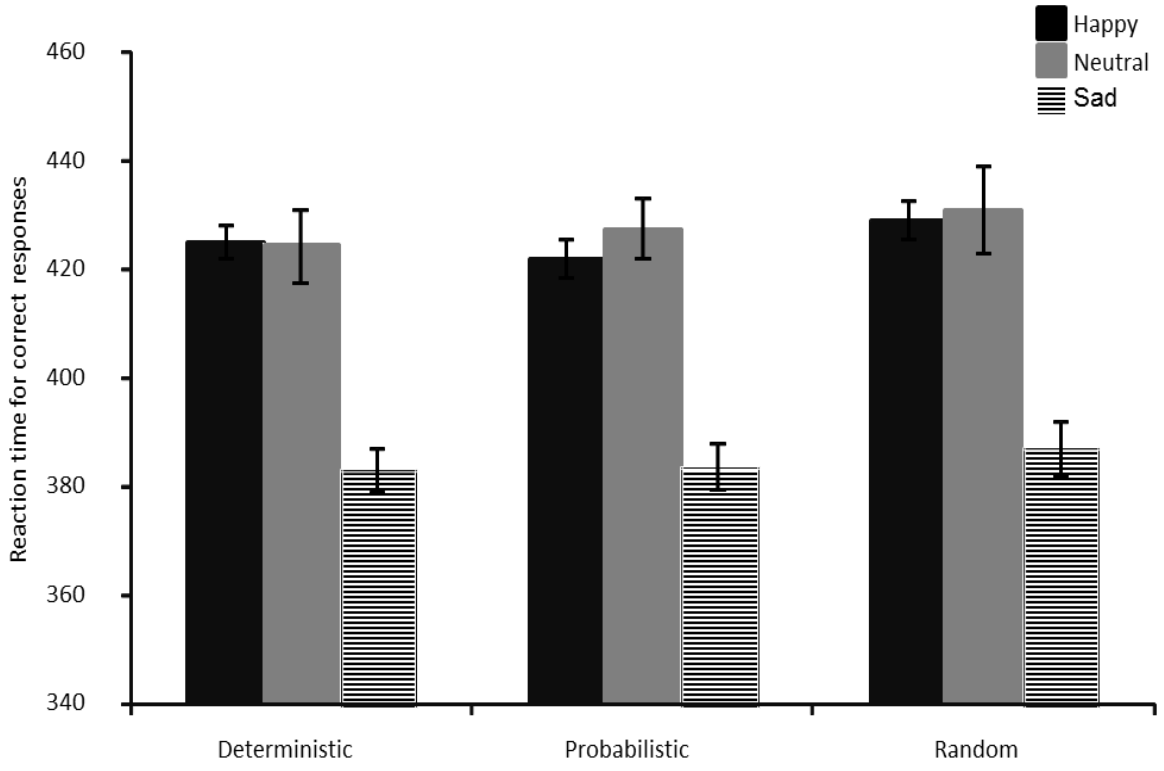

Figure 3. Reaction times for correct responses decomposed as a function of bin, condition and group. The error bar corresponds to 1 standard error of the mean.

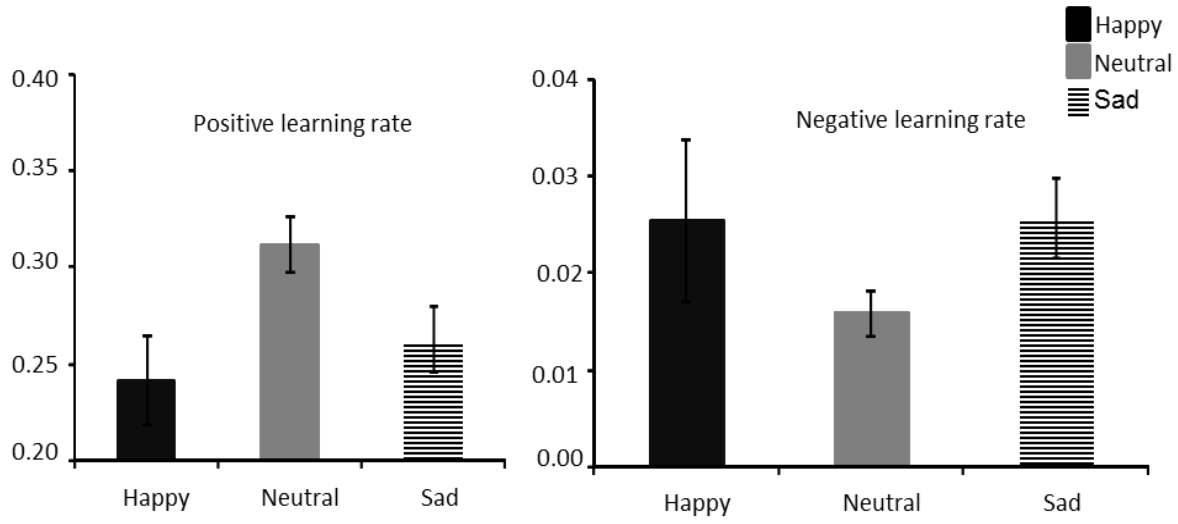

Figure 4. Learning rate for positive (left panel) and negative (right panel) feedback, separately for the Happy, Neutral and Sad mood group. The error bar corresponds to 1 standard error of the mean.

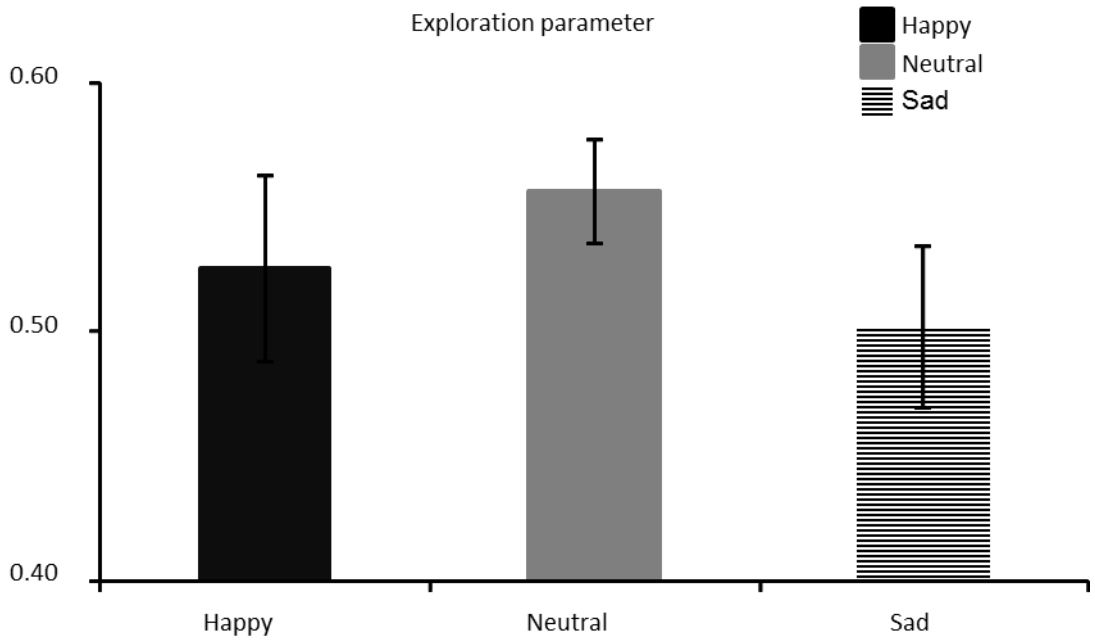

Figure 5. Exploration parameter shown separately for the Happy, Neutral and Sad mood group. The error bar corresponds to 1 standard error of the mean.

Post-experiment ratings. The mixed-model ANOVA carried out on the clarity ratings showed a significant main effect of Condition ($F(1.70, 71.61)= 708.28$, $p< .01$ $\eta 2=0.94$), showing that clarity varied monotonically as a function of probability. The analysis performed on the certainty ratings revealed significant Phase*Condition interaction ($F(2,84)=76.97$, $p< .01$, $\eta^2= 0.65$), main effect of phase ($F(1,42) =53.12$, $p<.01$, $\eta^2=0.56$) and condition ($F(2,84)=228.89$, $p< .01$, $\eta^2= 0.85$). This significant interaction was followed up by a paired t-tests to compare certainty across the two phases for each condition separately. The only significant difference was found in the deterministic condition, where certainty in the no feedback phase ($M= 82.93$, $SD=5.66$) was significantly higher than in the phase with feedback ($M=61.49$, $SD= 11.48$), ($t(44)= -15.32$, $p<.01$).

The analyses pertaining to subjective reports about the amount of positive vs. negative feedback received during the whole experimental session, as well as the like/dislike reactions to them revealed no significant group differences (all p's>.05).

*Questionnaires.*There were no significant group differences on BDI or RS-nl.

4. DISCUSSION

In this study, we sought to assess whether happy or sad mood could change RL, when compared to an active control condition or group with a neutral mood content. Even though no general consensus has emerged yet in the literature regarding effects of mood valence on learning (Gray, 2001; Huntsinger, 2012; Nadler et al., 2010; van Steenbergen et al., 2010), it is usually agreed that being in a state of increased emotionality (either positive or negative) alters motivational processes activated by cues signaling reward or punishment (Lang & Bradley, 2010), and hence learning by extension when this process is based on the direct exploitation of these incentives, like in the present case. More specifically, our primary goal was to assess if inducing happy mood could eventually lead to a gain in performance during RL, especially when externally-provided feedback information on task performance (hence cues signaling reward or punishment) were omitted and exploitation was fostered. For this purpose, we adapted a previously validated probabilistic learning task (Bakic et al., 2014; Eppinger et al., 2008; Unger et al., 2012) and introduced a second phase during the experiment where feedback on task performance was removed and hence learning could no longer be based on these (external) incentives, i.e., negative or positive feedback regarding task performance. During the initial phase of the experiment where this feedback information was still available, we expected to replicate the results of our previous study (Bakic et al., 2014), where we found that inducing positive mood led to an increase in the learning rate. Based on this previous study as well as the evidence currently available in the literature, we formulated a specific prediction: if the valence of the mood plays an important role in RL (Bolte & Goschke, 2010; Chiew & Braver, 2014), then happy and sad participants should

behave in opposite ways. More specifically, we expected that the learning rate would be larger in the happy compared to the neutral mood group, while it would be lower in the sad mood group compared to the control mood group. Moreover, by removing feedback as soon as learning was established, we hoped to exacerbate possible group differences in RL, bearing in mind that only standard accuracy and RT data could be extracted during this specific phase of the experiment (while computational modeling parameters could be estimated during the first phase of the experiment, as in our previous study; see Bakic et al., 2014) .

The results of this study confirm that guided imagery provides a valid method to induce specific mood states, either happy or sad. The happy mood group had a substantial increase in self-reported levels of happiness, comparable in size to the increase of sadness in the sad mood group.

Learning was clearly evidenced during the first phase of the experiment, equally strongly in the three mood groups however, challenging our assumption that mood valence (either positive or negative) could influence this process. Moreover, when considering two standard learning parameters extracted from a computational model (Bakic et al., 2014; Jepma & Nieuwenhuis, 2011), we still failed to disclose significant group differences, unlike what we found in our previous study where happy mood was associated with a larger learning rate (without concurrent change in exploration) compared to neutral mood. Strikingly, our results for the second phase of the experiment showed that participants (in all three groups) continued to learn in the absence of direct feedback information regarding task performance, suggesting that they unambiguously used or exploited abstract mental representations to comply with the task demands, as opposed to using or exploring externally provided cues signaling punishment or reward solely or primarily. However, neither happy nor sad mood did influence this learning process, even when feedback exploration was no longer feasible (second phase of the experiment).

Nieuwenhuis and colleagues (2005) already discussed the importance of feedback delivery: when it is delivered on a trial level, subjects tend to rely on external information more than on the internal monitoring system, and this may lead to reduced uncertainty levels as there is always an external check of the prediction. When the feedback is removed, it is expected that internal monitoring processes and the knowledge of the associations will be even more activated. Our current results add indirect support to this claim as we saw that removing feedback did not lead to a cost, but learning still progressed as if a boost of exploitation was triggered by this manipulation.

The failure to replicate our previous findings (Bakic et al., 2014) for the learning rate during the first phase of the experiment is puzzling at first sight, given that aside from the inclusion of a sad mood group in the current study, the experimental procedure was kept identical between these two studies. However, a closer look at the subjective ratings in these two studies might give us some hints on some of the reasons underlying this apparent discrepancy. When comparing mood changes directly between the two studies (i.e., the present one and Bakic et al., 2014), we found that the MIP in our previous study led to a large and significant difference between the two mood groups not only in valence, but also in arousal ($t(30)=3.10$, $p<.01$), while this was not the case in the current study ($p=.10$) (see Figure 6). Hence, in our previous study (Bakic et al., 2014), participants in the happy mood group were not only more happy than in the neutral mood group, but also more aroused by the MIP; an effect that was not found in the current study. Moreover, in our previous study, we found that the increase in happiness following the MIP (relative to the pre-MIP baseline measurement) correlated strongly with the increase in arousal ($r=.63$, $p<.01$), and importantly with both the positive ($r=.44$, $p<.01$) and the negative learning rate ($r=.51$, $p<.01$) as well. It seems that in our previous study (Bakic et al., 2014), the higher learning rate in the positive than the neutral mood group was likely explained by changes occurring both in the valence

and arousal dimensions in the former group as a function of the MIP. By comparison, no similar correlation was evidenced in the current study.

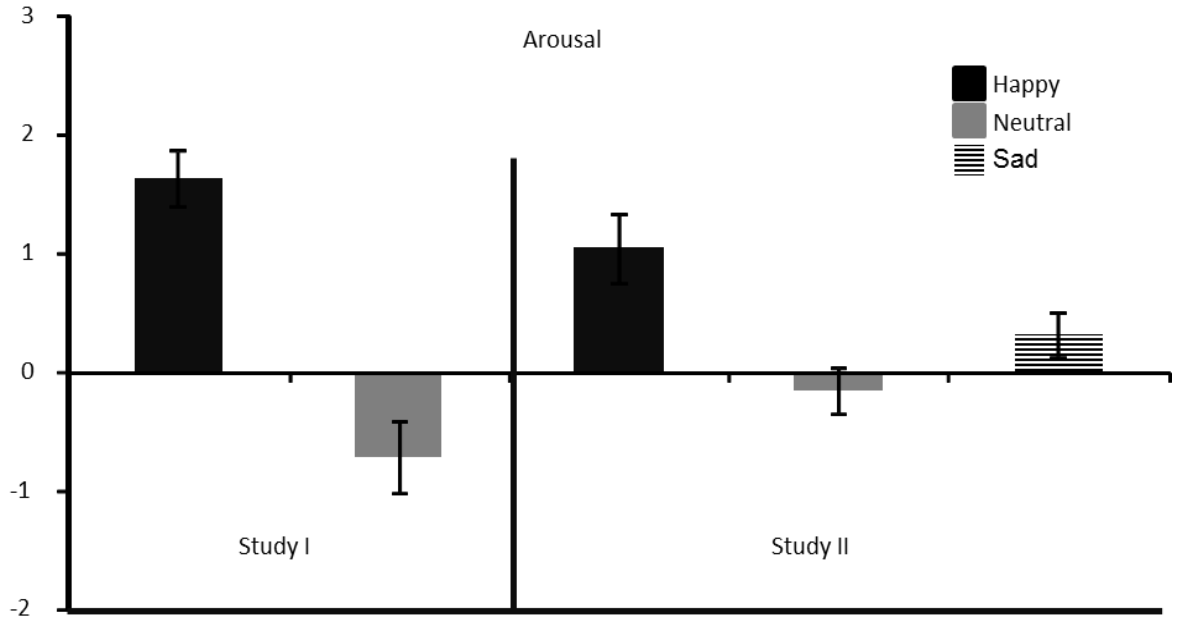

Figure 6. Arousal ratings shown separately for the happy and neutral mood group of Bakic et al. (2014 (left panel), and the three different mood groups of the current study (right panel).

Likewise, we also failed to find the evidence in this study for significant group differences regarding the perceived amount and like/dislike reactions to the feedback given during the RL task, while we did so in our previous study (see Bakic et al., 2014), suggesting that the elected MIP had probably a stronger impact (in terms of emotional changes brought about) in the positive mood group in our previous compared to the current study. Accordingly, it is tempting to conclude that our failure to replicate our previous findings for the learning rate (which was increased in the happy compared to the neutral mood group; see Bakic et al., 2014) in the present study might be imputed to the failure to elicit a reliable increase in levels of arousal in the happy compared to the neutral or sad mood group with our MIP. More generally, we believe this systematic comparison between our two studies is valuable because it confirms that arousal is probably an important dimension to consider (besides valence per se) in order to better understand modulatory effects on RL as a function

of positive mood, as we previously observed (Bakic et al., 2014) but failed to replicate here. As a matter of fact, arousal has often been conceived as an important determinant of learning (and more specifically the exploration-exploitation tradeoff) in the past, providing salience information to the organism. For example, according to the adaptive gain theory of Aston-Jones and Cohen (2005), exploration and exploitation are two decision strategies that depend directly on tonic and phasic changes in arousal, which is modulated by salience estimation generated in prefrontal areas. Jepma and Nieuwenhuis (2011) previously used this specific framework using a "four-armed bandit" task in healthy adult participants (without any mood induction) and showed that changes in the pupil diameter (a putative index of locus coeruleus activity) correlated with transitions from exploration to exploitation. Recent study corroborated the finding that arousal-related process indeed contributed to shape learning in a volatile setting (Browning, Behrens, Jocham, Reilly, & Bishop, 2015). Speculatively, it may therefore be the case that our MIP in the presents study failed to increase arousal substantially in the happy mood group (unlike what we found in Bakic et al., 2014), which in turn did not change the exploration-exploitation tradeoff and/or the learning rate in this group. Alternatively, arousal (resulting from the MIP we used here and in our previous study) could also foster probabilistic learning, when elicited to a sufficient degree, by influencing specific (short-term) memory processes needed to resolve the task, given that arousal usually heightens memory (Mather &Carstensen, 2005; Clewett & Mather, 2014). Accordingly, future studies are needed to assess whether (positive) mood valence could create changes in RL if and only if this specific mood state is accompanied by variations along the arousal dimension too.

Several limitations of our study warrant comment. First, if arousal plays an important role in mediating effects of mood on RL, then it is likely that a MIP tailored to increase arousal selectively (rather than valence) might provide a more promising avenue to evidence

effects of mood on the exploration-exploitation tradeoff during RL (Knutson, Katovich, & Suri, 2014). Here by contrast, we created three groups differing primarily regarding the valence of the mood induced (happy, sad or neutral), which may eventually have blurred rather than cleared out some of the group differences during RL. The choice of a (low-arousing) sad mood as comparison for the happy mood group was motivated by many earlier studies and models in the literature arguing that sadness is the opposite of happiness, as well as a good proxy of the anhedonic component in depression, for which there is already good evidence for modulatory effects on learning behavior, especially when it is based on either reward or punishment incentives/cues (Chase et al., 2010; Liu et al., 2014; Padrão, Mallorquí, Cucurell, Marco-Pallares, & Rodriguez-Fornells, 2013; Pizzagalli, 2014; Taylor Tavares et al., 2008)

Second, even though happy or sad mood did not alter learning at the behavioral (or computational modeling) level, we cannot rule out the possibility that these mood states could influence specific electrophysiological markers of RL, including the ERN and FRN components. Noteworthy, in our previous study (Bakic et al., 2014), we found that happy compared to neutral mood increased the ERN in the deterministic condition selectively. Accordingly, it would be important in the future to add EEG measurements to this probabilistic learning task in order to assess for example whether mood can dynamically alter early stages of error monitoring (in the absence of obvious changes at the behavioral level) or not, especially when feedback information on task performance has been removed and learning has to operate primarily based on exploitation of pre-existing knowledge or mental representations.

Finally, the probabilistic learning task used here (relying on a simple speeded two-alternatives forced choice task; see Eppinger et al., 2008) may not be sensitive enough to capture subtle changes in learning related to labile mood states, such as elicited after the MIP

(based on guided imagery) used in this study. Perhaps mood does change choice behavior, but not decision making per se, a hypothesis that would require the use of other experimental paradigms than the one used here, and where not only the amount but also the type of learning strategy at stake could be probed (see Frank et al., 2015). Presumably, the specific structure of the task used here, as well as the specific task requirements, may have weakened the expression of mood-related changes during RL. Likewise, reversal learning paradigms (see Chase, Swainson, Durham, Benham, & Cools, 2011) or more complex and volatile learning environments based on the use of more than two-alternatives forced choice task (see Browning et al., 2015; Jepma & Nieuwenhuis, 2011) could perhaps help to reveal clearer and stronger modulatory effects of either positive or negative mood on RL.

5. REFERENCES

Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529–550. Retrieved from <Go to ISI>://000082133000004

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Bakic, J., Jepma, M., De Raedt, R., & Pourtois, G. (2014). Effects of positive mood on probabilistic learning: Behavioral and electrophysiological correlates. *Biological Psychology*, *103*, 223–232. doi:10.1016/j.biopsycho.2014.09.012

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588–597. doi:10.1207/s15327752jpa6703_13

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–21. doi:10.1038/nn1954

Bolte, A., & Goschke, T. (2010). Thinking and Emotion: Affective Modulatoin of Cognitive Processing Modes. In B. Glatzeder, V. Goel, & A. Müller (Eds.), *Towards a Theory of Thinking* (pp. 261–277). Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/index/10.1007/978-3-642-03129-8

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. doi:10.1038/nn.3961

Chase HW, Frank MJ, Michael A, Bullmore ET, S. B. R. T. (2010). Approach and avoidance learning in patients with major depression and healthy controls : Relations to anhedonia. *Psychol Med*, *40*, 433–440.

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience*, 23(4), 936–46. doi:10.1162/jocn.2010.21456

Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, 14(2), 509–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24733296

Clewett, D. V., & Mather, M. (2014). Not all that glittered is gold: neural mechanisms that determine when reward will enhance or impair memory. *Frontiers in Neuroscience* . http://www.frontiersin.org/Journal/FullText.aspx?f=55&name=neuroscience&ART_DOI=10.3389/fnins.2014.00194

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 933–42. doi:10.1098/rstb.2007.2098

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

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. *Journal of Neuroscience*, 35(2), 485–494. doi:10.1523/JNEUROSCI.2036-14.2015

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci*, *359*(1449), 1367–1378. doi:10.1098/rstb.2004.1512

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, 530–547. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24659000

Gray, J. R. (2001). Emotional modulation of cognitive control: approach-withdrawal states double-dissociate spatial from verbal two-back task performance. *Journal of experimental psychology*. General (Vol. 130, pp. 436–452).

Holmes, E. A., Coughtrey, A. E., & Connor, A. (2008). Looking at or through rose-tinted glasses? Imagery perspective and positive mood. *Emotion* (Washington, D.C.), 8(6), 875–879. doi:10.1037/a0013617

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). Positive interpretation training: effects of mental imagery versus verbal training on positive mood. *Behavior therapy* (Vol. 37, pp. 237–247). doi:10.1016/j.beth.2006.02.002

Holmes, E. a, & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, 30(3), 349–62. doi:10.1016/j.cpr.2010.01.001

Huntsinger, J. R. (2012). Does positive affect broaden and negative affect narrow attentional scope? A new answer to an old question. *Journal of Experimental Psychology: General*, *141*(4), 595–600.

Isen, A. M. (1984). Toward understnding the role of affect in cognition. In R. Wyer & T. Srull (Eds.), *Handbook of social cognition* (pp. 174–236). Hillsdale, NJ: Erlbaum.

Isen, A. M. (1993). Positive affect and decision making. In M. Lewis & J. Haviland (Eds.), *Handbook of emotion* (pp. 261–277). New York: Guilford Press.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*(6), 1122–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3598858

Isen, A. M., Johnson, M. M., Mertz, E., & Robinson, G. F. (1985). The influence of positive affect on the unusualness of word associations. *Journal of Personality and Social Psychology*, *48*(6), 1413–1426.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–96. doi:10.1162/jocn.2010.21548

Knutson, B., Katovich, K., & Suri, G. (2014). Inferring affect from fMRI data. *Trends in Cognitive Sciences*, *18*(8), 422–428. doi:10.1016/j.tics.2014.04.006

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology*, *84*(3), 437–450.

Liu, W.-H., Wang, L.-Z., Shang, H.-R., Shen, Y., Li, Z., Cheung, E. F. C., & Chan, R. C. K. (2014). The influence of anhedonia on feedback negativity in major depressive disorder. *Neuropsychologia*, *53*, 213–20. doi:10.1016/j.neuropsychologia.2013.11.023

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review*, 10(6), 669–697.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: the positivity effect in attention and memory. *Trends in Cognitive Sciences*, 9(10), 496–502. doi:10.1016/j.tics.2005.08.005

Nadler, R. T., Rabi, R., & Minda, J. P. (2010). Better mood and better performance. Learning rule-described categories is enhanced by positive mood. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(12), 1770–1776. doi:10.1177/0956797610387441

Nieuwenhuis, S., Nielen, M. M., Mol, N., Hajcak, G., & Veltman, D. J. (2005). Performance monitoring in obsessive-compulsive disorder. *Psychiatry Research*, 134(2), 111–122. doi:10.1016/j.psychres.2005.02.005

Padrão, G., Mallorquí, A., Cucurell, D., Marco-Pallares, J., & Rodriguez-Fornells, A. (2013). Neurophysiological differences in reward processing in anhedonics. *Cognitive, Affective & Behavioral Neuroscience*, *13*(1), 102–15. doi:10.3758/s13415-012-0119-5

Pizzagalli, D. a. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, (January), 1–31. doi:10.1146/annurev-clinpsy-050212-185606

Portzky, M., Wagnild, G., De Bacquer, D., & Audenaert, K. (2010). Psychometric evaluation of the Dutch Resilience Scale RS-nl on 3265 healthy participants: a confirmation of the association between age and resilience found with the Swedish version. Scandinavian *Journal of Caring Sciences*, 24 Suppl 1, 86–92. doi:10.1111/j.1471-6712.2010.00841.x

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception*, 33(2), 217–236. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15109163

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: an introduction.* Cambridge, MA: MIT Press.

Taylor Tavares, J. V, Clark, L., Furey, M. L., Williams, G. B., Sahakian, B. J., & Drevets, W. C. (2008). Neural basis of abnormal response to negative feedback in unmedicated mood disorders. *NeuroImage*, *42*(3), 1118–26. doi:10.1016/j.neuroimage.2008.05.049

Unger, K., Kray, J., & Mecklinger, A. (2012). Worse than feared? Failure induction modulates the electrophysiological signature of error monitoring during subsequent learning. *Cognitive, Affective, & Behavioral Neuroscience*.

Van Steenbergen, H., Band, G. P., & Hommel, B. (2010). In the mood for adaptation: how affect regulates conflict-driven control. *Psychological Science*, *21*(11), 1629–1634. doi:10.1177/0956797610385951

Vanlessen, N., Rossi, V., De Raedt, R., & Pourtois, G. (2014). Feeling happy enhances early spatial encoding of peripheral information automatically: electrophysiological time-course and neural sources. *Cognitive, Affective & Behavioral Neuroscience*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24570275

Watkins, E. R., & Moberly, N. J. (2009). Concreteness training reduces dysphoria: a pilot proof-of-principle study. *Behaviour Research and Therapy*, 47(1), 48–53.

Westermann, R., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity o f mood induction procedures : a meta- analysis. *European Journal of Social Psychology*, *26*(January 1995), 557–580. doi:10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

CHAPTER 5

Preserved reinforcement learning mechanisms in major depressive disorder: evidence from behavioral, computational modeling and electrophysiological data[1]

Major depressive disorder (MDD) is characterized by abnormally high levels of negative affect and blunted positive affect concurrently. Accordingly, this mood disorder creates debilitating effects on cognition and learning. In this study, we examined possible changes or impairments in reinforcement learning (RL) as a function of MDD. Thirty-five treatment resistant MDD patients and 44 matched healthy controls (HCs) performed a standard probabilistic learning task, while 64-channels EEG was recorded concurrently. Both the amount of RL achieved (quantitative effects), as well as its underlying exploration-exploitation tradeoff (qualitative effects) were similar between these two groups. Moreover, the error related negativity (ERN), an electrophysiological marker of reward prediction error, was similar between the two groups alike, and showed normal amplitude variations as a function of reward probability in this task. These results suggest therefore relatively preserved cognitive abilities during RL in a group of treatment resistant MDD patients (characterized by high levels of Anhedonia), compared to a group of matched HCs. We interpret these new findings in terms of spared motivational processes in these patients that may in turn compensate for the otherwise abnormal processing of (positive and negative) reinforcers during learning.

[1] Bakic, J., Pourtois, G., Jepma, M., Duprat, R., De Raedt, R., & Baeken, C. (2014) Preserved reinforcement learning mechanisms in major depressive disorder: evidence from behavioral, computational modeling and electrophysiological data. Manuscript in preparation

1. INTRODUCTION

In a world of ever changing contingencies, each action is potentially costly. Therefore, there is a high demand for adaptive (i.e. reducing action-outcome uncertainty) and efficient (i.e. minimizing resource consumption) updating. All of this does not happen in a vacuum, but is highly dependent on the current affective state of the individual. Undoubtedly, there is a strong affective component associated with higher cognitive functions like reinforcement learning (RL). Emotional factors such as fear, sadness, worry, or pleasure, enjoyment, and happiness shape these processes, presumably with differential effects depending on the specific mood content or affective state induced or experienced. One could even argue that these affective components could potentially determine to what extent an outcome is eventually evaluated as rewarding or punishing. Ideally, affect, both of positive and negative valence, given its intrinsic motivational properties, could timely inform learning and lead to performance optimization (e.g. favoring maybe exploration when the affective state is colored by positive emotion, while leading to a more cautious and exploitation-driven behavior when the state or environment is perceived as negative or potentially threatening). But, what happens when this system somehow breaks down, because negative mood and sensitivity to punishment prevail, and they become chronic? When punishment sensitivity dominates over reward seeking, will it lead to a profound change in the way RL is accomplished? This question can be addressed by exploring RL in patients suffering from major depressive disorder (MDD).

A central hypothesis nowadays in the psychopathology literature is that disruption in reward processing (and Anhedonia) is one of the key deficits in MDD (Downar et al., 2014; Elliott, Sahakian, Michael, Paykel, & Dolan, 1998; Mayberg et al., 2005). Anhedonia is

defined as the "loss of pleasure or lack of reactivity to pleasurable stimuli", probably originating from dysfunctional interactions or neural communication between stress and brain reward systems (i.e., in specific fronto-striatal loops), and it seems to be predictive of recovery/remission (or the lack thereof) from MDD (Pizzagalli, 2014). This loss of pleasure or reactivity against rewarding or positive events has direct (negative) repercussions on behavior and learning, when it is based on the direct use or exploitation of rewarding cues or incentives available in the proximal environment. Consistent with this assumption, Pizzagalli, Iosifescu, Hallett, Ratner, & Fava (2009) previously found that MDD patients had a lower reward-based (probabilistic) learning compared to healthy controls (HC). These authors suggested that this deficit was due to reduced hedonic capacity in MDD group, arising from attenuated left prefrontal activations. Compared to HCs, MDD patients are also typically less motivated to increase effort for gaining rewards (Sherdell, Waugh, & Gotlib, 2012; Treadway, Bossaller, Shelton, & Zald, 2012). But MDD does not simply blunt the processing of positive reinforcers. This mood disorder is also characterized by increased negative affect and worry (Disner, Beevers, Haigh, & Beck, 2011). Depressed individuals display difficulty disengaging from negative thoughts (Conway, Howell, & Giannopoulos, 1991), and have for example increased chance of repeated error commission (i.e., inadvertently making two consecutive response errors; see Elliott et al., 1996). This worry-based, or apprehension-oriented mode of cognitive control (as opposed to arousal for example) has previously been related to altered performance monitoring, with effects visible at the electrophysiological level (Moser, Moran, Schroder, Donnellan, & Yeung, 2013). Even for the remitted MDD patients, there is evidence of increased electrophysiological response to negative feedback (in a learning task with intermittent negative or punishing outcomes), despite the fact that their (learning) performance was balanced with HCs (Santesso et al., 2008). Additionally, MDD patients were found to not be able to ignore "misleading" negative feedback in a reversal learning task,

which eventually led them to switch "unnecessarily" to a less favorable response option and decreased performance/accuracy (Murphy, Michael, Robbins, & Sahakian, 2003). In a different study, MDD patients made much more unnecessary adjustments after invalid negative feedback compared to HCs (Mies et al., 2011). Thus, timely tracking of reinforcement history (during learning) in order to make sure that overall the best outcome is eventually achieved appears to be compromised by MDD, because of an enhanced and abnormal sensitivity/hyper-reactivity to negative outcome that is apparently hard to overcome in this specific emotional disorder.

These findings suggest that MDD patients exhibit altered usage of reinforcers (positive and negative) during learning, which influences the adaptiveness and flexibility of their behavior. Noteworthy, this deficit is sometimes associated with improvements or benefits during learning: MDD patients can sometimes outperform HCs, for example, in a sequential choice task, where their extreme cautiousness (and reluctance) actually pays off (von Helversen, Wilke, Johnson, Schmid, & Klapp, 2011). Likewise, in an avoidance learning task, Cavanagh et al. (2011) found increased accuracy for avoiding less favorable options (accompanied by heightened fronto-central theta activation related to negative feedback) in the MDD group, even though overall they did not show better performance compared to HCs. These findings indicated that learning in MDD was mostly influenced by avoiding negative outcome, an element or bias which can be beneficial to the learning performance in specific situations or cases. However, it is worth mentioning that this potentiated processing of negative information (with the aim to avoid it presumably) is inversely related to the strength or severity of Anhedonia: the higher the anhedonic symptoms, the lower the feedback processing (Mueller, Pechtel, Cohen, Douglas, & Pizzagalli, 2015). Alternatively, this enhanced avoidance behavior towards negative feedback or outcome might stem from excessive worry (Disner et al., 2011; Moser et al., 2013). Interestingly, this conclusion has

also been corroborated recently for anxiety-related worry (at the subclinical level) in volatile learning environments (Browning, Behrens, Jocham, Reilly, & Bishop, 2015).

Changes in learning as a function of MDD are usually explored not only by standard behavioral measures (either accuracy or speed), but also by using electroencephalography (EEG), and more specifically, frontal midline event related brain potentials (ERPs) arising from midcingulate cortex (MCC). Usually, the focus is put on response-locked or feedback-locked ERPs because they relate to learning, and more specifically reward prediction (error) learning (see Holroyd & Coles, 2002). Recently Cavanagh and Shackman (2014) provided a comprehensive review of this psychophysiology literature and they argued that, as they show sensitivity to both changes in reinforcement contingencies and affective states, these fronto-central ERPs likely reflect the involvement of the MCC in both cognitive control and emotion (regulation) processes. The error related negativity (ERN,or negativity error/Ne) is probably the most studied ERP component in this context. It is a negative deflection peaking ~0-100 ms at fronto-central electrodes after the onset of incorrect (usually response errors) compared to correct actions/responses (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN/Ne is thought to originate from the MCC, and more specifically the rostral cingulate zone (Dehaene, Posner, & Tucker, 1994; Ullsperger, Fischer, Nigbur, & Endrass, 2014). Moreover, the ERN, although showing a (very) rapid electrophysiological time-course following (incorrect) response onset, is thought to reflect the backdoor of a large brain system composed of the MCC and interconnected limbic (basal ganglia) and midbrain dopaminergic structures involved in RL (Holroyd & Coles, 2002). In this framework, the ERN is therefore not simply reflecting/detecting error making, but instead is tracking the value of a specific action and assigning a given valence to it (i.e., an unwanted response error can be seen as a worse than expected event or an event which is compatible with a reward prediction error, thereby “automatically” acquiring a

negative meaning or connotation). According to the RL theory (see also Frank, Woroch, & Curran, 2005), the ERN corresponds to a phasic dopaminergic burst that signifies the occurrence of an unexpected and negative outcome, as well as the need to change or adapt behavior accordingly (Holroyd & Coles, 2002). Interestingly, a large body of evidence suggests that the ERN shows amplitude variation depending on specific emotional or motivational factors, and more specifically negative affect and punishment sensitivity (Olvet & Hajcak, 2008; Vaidyanathan, Nelson, & Patrick, 2012). In high anxious (and to a lesser extent depressed; Weinberg, Riesel, & Hajcak, 2011) individuals, the ERN is usually larger or "overactive", compared to HCs or low anxious participants, especially using standard laboratory interference tasks (such as the Flanker, Stroop or Go/noGo tasks) where learning is usually kept minimal or even neutralized. Recently, we used a standard probabilistic learning task in healthy adults (Eppinger et al., 2008) and found that the ERN component was actually increased after the induction of happy mood, relative to a control condition with a neutral mood content (Bakic, Jepma, De Raedt, & Pourtois, 2014). These ERP results challenged the notion that negative affect (necessarily) increases the ERN component. Hence, while the ERN is usually considered to reflect (response-locked) RL brain mechanisms (Frank et al., 2005), the actual functional meaning of its amplitude variation as a function of (negative) affect (trait-related) remains debated in the literature (Moser et al., 2013).

The goal of this study was to characterize changes in RL as a function of MDD, when this process was explored at multiple levels concurrently. To this end, we capitalized on a well-validated probabilistic learning task (Eppinger et al., 2008) that we previously used and adapted to explore changes in RL following the induction of positive mood in healthy adult participants (Bakic et al., 2014). Here, we administered it to a large group of MDD patients and matched HCs. RL was titrated at the behavioral and electrophysiological levels concurrently. At the behavioral level, learning accuracy was calculated following standard

practice (Eppinger et al., 2008). Moreover, we also used computational modeling to extract two parameters from the single trial (behavioral) data (Jepma & Nieuwenhuis, 2011; Sutton & Barto, 1998). (i) Learning rate is considered to be a measure of exploitation of previous response options and can be modeled separately for rewards (positive feedback) and punishments (negative feedback). (ii) The second parameter measures "exploration", or more precisely the amount of exploration of the different response options before choosing one. At the ERP level, we focused on the ERN component since it is related to RL (Frank et al., 2005; Holroyd & Coles, 2002), in particular to reward probability in this task (Eppinger et al., 2008), as we previously evidenced in healthy adult participants (Bakic et al., 2014). Based on the evidence reviewed here above, we formulated the following predictions. (i) Even though the overall learning performance could be balanced between MDD patients and HCs (see Cavanagh et al., 2011), we reckoned that MDD patients might show a larger learning rate following negative feedback (or alternatively a lower learning rate following positive feedback) than HCs, in keeping with their heightened sensitivity to cues signaling punishment and/or blunted reactions to reward (Pizzagalli, 2014). (ii) At the ERP level, we predicted that MDD patients would show a larger/overactive ERN compared to the HCs, especially when the clarity and visibility of the stimulus-response mapping/associations (and hence reward probability) was high (i.e., in the deterministic condition of the experimental design previously devised by Eppinger et al., 2008; see also Bakic et al., 2014 and Unger, Heintz, & Kray, 2012).

2. METHODS

2.1. Participants

Sixty non-depressed HCs (35 females, 25 males, mean age: 37.90, SD=12.82) and forty two individuals meeting the *Diagnostic and Statistical manual of Mental Disor*ders 5 criteria

(American Psychiatric Association, 2013) for unipolar MDD (30 females, 12 males, mean age: 41.40, SD=12.04) participated in the current study. All participants had normal or corrected to normal vision.

HCs were recruited via advertisements for comparison purposes with the MDD patients and received a 60 euro compensation for their participation. They were matched on group level with MDD sample for age, sex and education. They gave informed consent prior to the beginning of the study. The MDD ambulatory or hospitalized patients were recruited from several Belgian psychiatric clinics and were all diagnosed with MDD by using the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Depression severity was assessed with the 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967), and the 21-item Beck Depression Inventory (Beck, Steer, & Brown, 1996). Additionally, they also filled in the Snaith-Hamilton Pleasure Scale (SHAPS, Snaith et al., 1995), a self-report questionnaire that measures hedonic responses in four different domains: interests and pastimes, social interaction, sensory experience, food and drink. They also filled in the Temporal Experience of Pleasure Scale (TEPS, Gard, Gard, Kring, & John, 2006) that measures consummatory and inhibitory Anhedonia. All the patients underwent the structured clinical interview (MINI) and HDRS carried out by a certified psychiatrist. These patients were deemed treatment resistant (Fava, 2003), and were classified as at least Stage I treatment resistant (Rush et al.,2003). Patients had had at least one unsuccessful treatment trial with a serotonin reuptake inhibitor/ noradrenaline or serotonin reuptake inhibitor (SSRI/NSRI), in adequate dose and over a sufficient amount of time (> 4 weeks). They actually took part in a larger clinical study at the Ghent University Hospital that examined (beneficial) effects of accelerated intermittent theta burst stimulation (iTBS) applied over the left dorsolateral prefrontal cortex (dlPFC) during four consecutive days (5 sessions/day). However, the results reported in this article pertain exclusively to the baseline testing moment, prior to the start of

the iTBS protocol. After a washout period, at the time of testing, patients were free of medication for at least 2 weeks. All patients were free from any antidepressant (AD), neuroleptic and mood stabilizer for at least two weeks before entering the iTBS treatment protocol. Only habitual benzodiazepine agents were allowed. The maximum allowed dose of benzodiazepines was the equivalent of 40 mg diazepam. These benzodiazepines equivalent doses are described by the British National Formulary (No. 66, London: British Medical Association and Royal Pharmaceutical Society of Great Britain; September 2013, pp. 218-22). Exclusion criteria were (i) bipolarity, (ii) the use of antipsychotics, tricyclic antidepressant, and (iii) a history of neurological disorders, including epilepsy, head injury, and a loss of consciousness, (iv) a history of electroconvulsive therapy, (v) a past or present substance abuse, (vi) past or present experience of psychotic episodes, (vii) learning disorders. Finally, some of those admitted to the study were excluded a posteriori for the following reasons: (viii) matching for age (n=4 HC), (ix) insufficient or no learning during the main task, (i.e. learning curves below chance level). The data of 16 participants (11 in the HC and 5 in the MDD group) had to be excluded due to below chance learning achievement, and (x) additional 3 (1 in HC and 2 in MDD group) due to excessively noisy EEG signal. Based on these criteria men were excluded significantly more than women ($\chi^2(3)=9.44$, $p<.05$). The two groups did not differ significantly for the number of participants excluded ($p=.17$). Importantly, inclusion of these participants did not change the results of the analyses reported below, however it was decided not to include them in the analysis to reduce the noise in the data. The final sample consisted of 44 HCs and 35 MDD patients. Demographic and clinical data are presented in Table 1. The study was approved by the ethics committee of the Ghent University Hospital.

Table 1. Demographic and clinical data for HCs and MDD patients (means, and in parenthesis standard deviations). Individual samples T-test differences are provided for HAM_D (df=77), BDI II (df=72), Anhedonia subscale of BDI II (df=77), TEPS with the corresponding subscales (dfs=74) and SHAPS (df=77).

	HC	MDD	t
N	44	35	
Age	37.89 (12.23)	43.00 (11.67)	-1.88
Sex	28F/16M	27F/8M	
Age at onset		24.6 (11.03)	
Lenght of episode (months)		20.81 (32.05)	
Number of episodes		3.14 (2.61)	
HAM_D	1.42 (2.37)	21.83 (5.63)	-21.79**
BDI_II	5.98 (6.75)	30.21 (10.27)	-12.16**
Anhedonia	0.98 (1.37)	4.66 (2.25)	-8.97**
TEPS	75.02 (19.22)	58.97 (17.04)	3.81**
Consumatory	36.05 (9.57)	28.76 (9.02)	3.39**
Inhibitory	38.89 (10.94)	30.21 (8.95)	3.76**
SHAPS	0.55 (2.16)	7.31 (4.09)	-9.45**

*p<.05, **p<.01

Based on iTBS treatment outcome (assessed for each patient four weeks after the start of the study), it was also possible to differentiate MDD patients as treatment responders or not (i.e., treatment outcome or response rate). It is important to mention that the iTBS protocol followed a randomized sham-controlled cross-over pattern and all patients only responded to the real iTBS and not sham. The criterion to decide whether the treatment was successful or not, was standard (Downar et al., 2014; Mayberg et al., 2005) and corresponded to a 50% reduction of symptoms' severity (as captured by the HSRD) between the baseline moment and the post treatment moment (i.e., occurring four weeks later). Based on this standard clinical criterion, 21 patients were considered as non-responders (baseline Hamilton:

M=22.00, SD=6.21, post-treatment Hamilton: M= 17.10, SD=7.24) and 14 patients as responders (baseline Hamilton: M=21.57, SD=4.85, post-treatment Hamilton: M= 6.85, SD=1.91). The HRSD difference between responders and non-responders was not significant at baseline ($p>.05$) but was significant post-treatment ($t(31) =4.97$, $p<.01$). Additionally there were no significant group differences at baseline for BDI, SHAPS, TEPS (all p's >.05).

2.2. Probabilistic learning task

We used a probabilistic learning task previously validated by Eppinger et al. (2008) and used in Bakic et al. (2014), as well as (Unger et al., 2012). In this task, participants were asked to decipher and learn, by trial and error, several hidden stimulus-response (S-R) mappings. For each trial, participants were asked to decide, with a strict time limit, whether the stimulus shown on the screen was arbitrarily associated with response 1 or 2. Feedback on the choice made by the participant was given following every response made.

Participants were presented with 6 different visual stimuli (A-B-C-D-E-F), belonging to three conditions (unknown to the participants) that differed regarding the actual probability of the S-R mapping (100%, 80% or 50%). In the condition 100%, each stimulus of the pair was always associated with one of the two response keys, corresponding to a "deterministic" S-R mapping (i.e., response 1 was always correct for stimulus A, and response 2 for the stimulus B). In the condition 80%, the S-R mapping was "probabilistic", given that stimulus C was associated 80% of the time with response 1 (and thus 20% of the time with the concurrent response 2), while stimulus D had a symmetric probability for the S-R mapping. Finally, in the condition 50% ("random" S-R mapping), each stimulus of the pair was associated equally often to each of the two response keys (i.e., stimuli E and F were associated 50% of the time with response 1 and 50% of the time with response 2).

Colorful line drawings (Rossion & Pourtois, 2004) were used as visual stimuli, presented against a white homogenous background on a 17-inch computer screen. These stimuli were visual objects belonging to different semantic categories (artifacts, buildings, musical instruments, clothes, vehicles, furniture). Their mean size was 7 cm width x 5 cm height, corresponding to 5 x 3,6 degrees of visual angle at 80 cm viewing distance.

The trial structure was as follows: it began with a fixation cross of 250 ms duration, followed by a 250 ms blank screen. Then, the stimulus was presented for 500 ms, followed by a blank screen of 300 ms. Response deadline was set to 800 ms following the onset of the visual stimulus on the screen. After 500 ms, performance feedback was presented for 500 ms. The feedback was provided in the form of a written word (in Dutch) shown in black against a white homogenous background. This word was "goed" (correct), "fout" (incorrect), or "te traag" (too late). The inter-trial interval was constant (500 ms) and it corresponded to a blank screen, after which a new trial would ensue. Manual responses (i.e., key presses) were recorded using the Cedrus response box.

Each participant completed two blocks of 240 trials. Each block had six different stimuli, each repeated forty times. Accordingly, participants had to learn new S-R mappings in each block. Trial order within a block, as well as the order of the two blocks was alternated across participants.

2.3. Procedure

Prior to the actual testing session, HCs and MDD patients were asked not to consume any caffeine or nicotine for a period of two hours at least. First, they were prepared for the EEG recording. In order to get acquainted with the task, they completed a short practice session of 20 trials, after which the task would ensue. After each block, participants were asked to indicate, for each of the 6 stimuli, the clarity and certainty of each of the six

stimulus-response (S-R) associations, by means of a horizontal 10-cm VAS. Furthermore, they were asked to rate the amount of positive vs. negative feedback they thought they received during this last block (using a 10 cm VAS going from "exclusively negative" to "exclusively positive"), as well as how much they liked or disliked this positive vs. negative feedback when receiving them (using a Likert scale spanning from 0 to 100). The whole experiment lasted about 2 hours.

2.3.1. EEG recording

EEG was recorded continuously using 64-channels by means of a Biosemi Active Two system (www. Biosemi.com). The EEG was sampled at 512 Hz, with CMS-DRL serving as the reference-ground. The EEG signal was filtered off line, using a 0.016 to 70 Hz filter (12db/oct), with a 50 Hz notch and re-referenced using the linked (average) mastoids. Individual epochs were segmented using a -/+ 500 ms interval around the response (see Aarts & Pourtois, 2010; Aarts, Vanderhasselt, Otte, Baeken, & Pourtois, 2013; Pourtois, 2011). Eye blinks were removed automatically via vertical ocular correction (Gratton, Coles, & Donchin, 1983), using two electrodes, placed above and below the right eye. Individual epochs were baseline corrected using the first 200 ms of the pre-response time-interval (i.e., from -500 to -300 ms prior to response onset; see also Bakic et al., 2014).

Artifact rejection was based on a ± 100 μV amplitude cutoff which led to 84.64% of the individual segments being kept and eventually included in the individual averages. No significant group difference [HCs: M=85.00, SEM=0.87; MDD patients: M=84.22, SEM=1.13; $t(77)= 0.55$, $p=.59$] was found for this metric. Finally, individual epochs were averaged separately for the different conditions and subjects, and an additional low pass filter set to 30 Hz was applied on the individual averages before grand-averaging.

2.4.Data analysis

Accuracy analyses. Accuracy data were expressed in proportions of correct responses (relative to the total number of trials), separately for each condition (n=3). Moreover, for each condition (2 stimuli x 40 repetitions), changes of learning performance as a function of time were captured by grouping the data into four bins of equal size (i.e., 20 trials/condition) (see Eppinger et al., 2008; Bakic et al., 2014 for a similar approach). These data were then submitted to a mixed model ANOVA with group as between subjects factor, and condition and bin as within subjects factors. An additional analysis was performed for the MDD patients only, with the responders vs. non-responders factor. We also used the number of switches after receiving negative feedback as additional dependent variable, given that this metric was related to exploration previously (i.e., the larger the number of switches after negative feedback, the more sensitive the participant/patient is presumably to this external feedback, because it translates somehow a change in learning/behavior as a function of exploration of the feedback; see Gonzalez & Dutt, 2011; Hills & Hertwig, 2012). Where necessary, Greenhouse-Geisser correction for sphericity was performed, and corrected p-values were reported, together with uncorrected effect size.

Reinforcement-learning model. We used two complementary measures based on the modeling procedure described in more detail in Bakic et al. (2014) and Jepma & Nieuwenshuis (2011). We computed first the learning rate parameter (α), which is a prediction error measure, so that the higher this value, the bigger the updating value of the new feedback compared to the previous response history. This was done separately for positive and negative feedback. We also calculated a second parameter, β, which is a "noise" parameter that reflects how much randomness there is in participant's choices (and hence exploration indirectly).

EEG analyses. We focused on the ERN/Ne component, a well-documented error-related ERP activity, which is related to reward prediction error at the response level (Holroyd & Coles,

2002; Eppinger et al., 2008; Frank, et al., 2005; Mies et al., 2011) given that we previously reported an amplitude modulation of this early error-related ERP component by (positive) mood in this task (Bakic et al., 2014). The mean amplitude for the ERN was calculated in an interval spanning 100 ms after response onset at electrode FCz (see Bakic et al., 2014 for a similar procedure). A mixed-model ANOVA was performed on the mean ERN amplitudes with group (first comparing HC vs. MDD; and subsequently responders vs. non-responders) as between subjects and condition and response accuracy as within subjects factor. In a second step, we computed difference waveforms by subtracting the ERP activity of incorrect from correct trials, following standard practice (Eppinger et al., 2008; Bakic et al., 2014).

Topography. While the previous ERP analysis informed about (local) amplitude changes of the ERN component occurring as a function of condition and group, we also wanted to use a data-driven ERP mapping analysis that took into account all 64 channels concurrently as well as all (512) time-frames in order to assess whether the experimental factors might also alter the topography (and not only the amplitude measured at FCZ) of the ERN electric field. To this aim, we used a spatial cluster analysis, which can summarize the ERP data set into a limited number of dominant field configurations (topographies), or functional microstates. The rationale and underlying principles of this electrical neuroimaging method have already been explained extensively elsewhere (Lehmann & Skrandies, 1980; Michel & Murray, 2012; Michel, Seeck, & Landis, 1999; Pourtois, Delplanque, Michel, & Vuilleumier, 2008). Following standard practice (see Pourtois et al., 2008), we first performed a topographic pattern analysis on the group-averaged data from -500 to 500 ms after response onset (512 consecutive time frames at a 512-Hz sampling rate), using a standard cluster (or spatiotemporal segmentation) method (i.e., K-means algorithm; see Pascual-Marqui, Michel, & Lehmann, 1995). Then, the segmentation results were fitted back to individual data for subsequent statistical testing. The spatiotemporal segmentation algorithm is derived from

spatial cluster analysis (Pascual-Marqui et al., 1995) and allows for the identification of the most dominant scalp topographies appearing in the group-averaged ERPs of each condition and over time, while minimizing the biases for the selection of time frames or electrodes of interest. Noteworthy, this procedure allows the identification of these dominant scalp topographies, irrespective of (local or global) changes in amplitude (Michel et al., 1999; Murray et al., 2008). The optimal number of topographic maps explaining the whole data set is determined objectively using both cross-validation (Pascual-Marqui et al., 1995) and Krzanowski-Lai (Tibshirani, Walther, & Hastie, 2004) criteria. The dominant scalp topographies are then fitted to the ERPs of each individual participant using spatial-fitting procedure to quantitatively determine their representation across accuracy and condition. This procedure thus provides fine-grained quantitative values, such as the duration of a specific topographic map or its global explained variance (GEV, or goodness of fit). GEV is a sum of the explained variance weighted by the global field power at each moment in time. These topographic analyses were carried out using CARTOOL software (Version 3.34; developed by D. Brunet, Functional Brain Mapping Laboratory, Geneva, Switzerland). The GEV of the topography corresponding to the ERN (or the correct-related negativity in the case of correct responses, CRN) was entered as dependent variable in a mixed model ANOVA with group as between subjects and condition, accuracy and map configuration as within subjects factors.

3. RESULTS

3.1. Behavioral results

Accuracy. The analysis carried out on the proportions of correct responses (Figure 1a) showed a significant Condition*Bin interaction ($F(4.72, 363.30)=31.92$, $p<.01$, $\eta^2=0.29$), as well as significant main effect of condition ($F(2, 154) =295.14$, $p<.01$, $\eta^2=0.79$), and bin ($F(2.74, 210.86) =73.86$, $p<.01$, $\eta^2=0.49$). As expected, these effects translated a steep learning across

time in the deterministic condition, intermediate in the probabilistic condition and no learning in the random condition. Groups did not differ significantly with respect to accuracy.

The analysis taking into account treatment outcome (Figure 1b) revealed a significant three-way Condition*Bin*Group interaction ($F(6, 198)=3.53$, $p<.02$, $\eta^2=.10$). However, followed up t-tests did not show any significant group difference (all p's>.05).

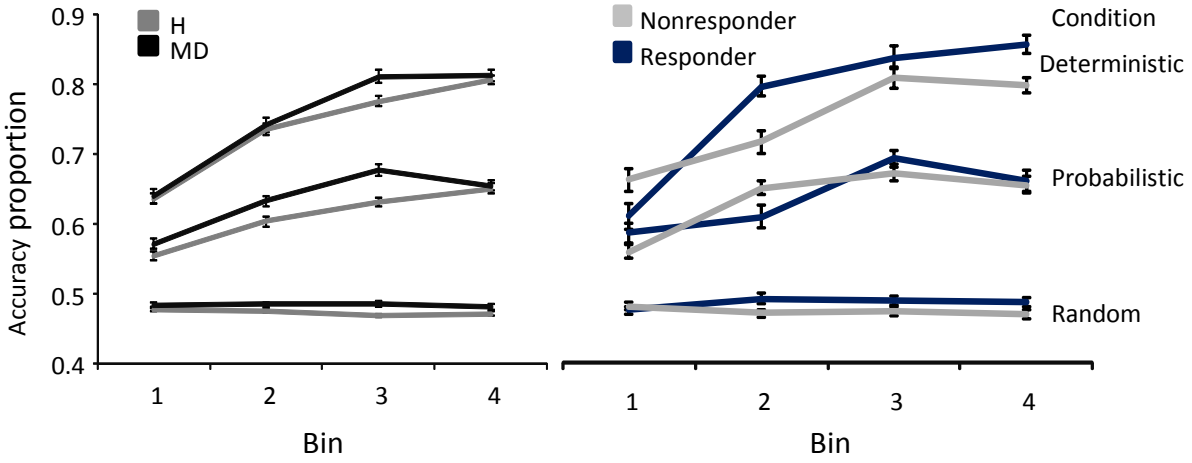

Figure 1. Accuracy data (i.e., proportion of correct responses) decomposed as a function of bin, condition and group: (a) HCs vs. MDD patients and (b) Responders vs. Non-responders. The error bar corresponds to 1 standard error of the mean.

Reaction times (RTs) for correct responses. The MDD group (M=444.85, SD=47.69) was significantly slower than the HC group (M=412.31, SD= 51.80), ($F(1, 77)= 8.25$, $p<.01$, $\eta^2=0.10$). Significant main effects of Condition ($F(2, 154)= 9.23$, $p<.01$, $\eta^2=0.11$), and Bin ($F(2.15, 165.39)= 4.47$, $p<.01$, $\eta^2=0.10$), were followed up by t-tests that showed that RTs for the deterministic condition (M=422.80, SD=51.67) was significantly faster than for the probabilistic condition (M=426.44, SD=51.88), ($t (78) =-2.00$, $p<.05$), which was, in turn, significantly faster than for the random condition (M=430.94, SD=55.96), ($t(78)= -2.40$, $p<.05$). Additionally, during bin 1 (M=431.86, SD=60.87), participants were significantly

slower than during bin 2 (M=420.92, SD=56.43), (t(78)=3.27, p<.01), but at bin 3 (M=430.22, SD=54.42) they slowed down again compared to bin2 (t(78)= -2.99, p<.01). There was increase in speed again at the end at bin4 (M=423.90, SD=50.11) compared to bin 3 (t(78)= -2.75, p<.01).

When entering treatment outcome in the statistical analysis, a significant Bin*Group interaction was evidenced ($F(3,99)=3.04, p<.03, \eta^2=0.08$). Independent samples t-tests showed that responders were somewhat slower than non-responders at bins 3 and 4, but these differences were only marginally significant.

Too late responses. The number of too late responses was modest (M=3.45, SD=1.83). It was significantly higher for the MDD group than for the HC group ($F(1, 77) =9.51$, $p<.01$, $\eta^2=0.11$). The number of too late responses also varied significantly with condition ($F(2, 154) =9.01$, $p<.01$, $\eta^2=0.11$), and bin ($F(3, 231) =3.12$, $p<.05$, $\eta^2=0.04$). Follow-up paired t-tests showed that there was significantly less too late responses in the deterministic (M=3.10, SD=2.0) compared to the probabilistic condition (M=3.59, SD= 2.00), t(78)=-3.16, p<.01), which did not differ significantly from the random condition (M=3.67, SD= 1.9), (p>.05). The auxiliary analysis taking into account treatment outcome did not show any significant difference between responders and non-responders for this variable (p>.05).

Switches after negative feedback. This analysis showed a significant Group*Bin interaction ($F(3, 231)= 3.47$, $p<.05$, $\eta^2=.04$; see Figure 2a). Individual t-tests showed that in the first half of the task the difference between the two groups was not significant (p>.05), while during the second half of the experimental session the MDD group (M= 0.24, SD= 0.10) had a lower number of switches after negative feedback compared to the HCs (M=0.30, SD= 10), (t(77)=2.88, p<.01). There was a significant main effect of Condition ($F(2, 154)= 8.13$, $p<.01$,

η^2=.10), and Bin (F(3, 231)= 2.89, p<.05, η^2=.04). Main effect of group was not significant (p>.05). When considering treatment outcome, no significant group difference was observed (Figure 2b).

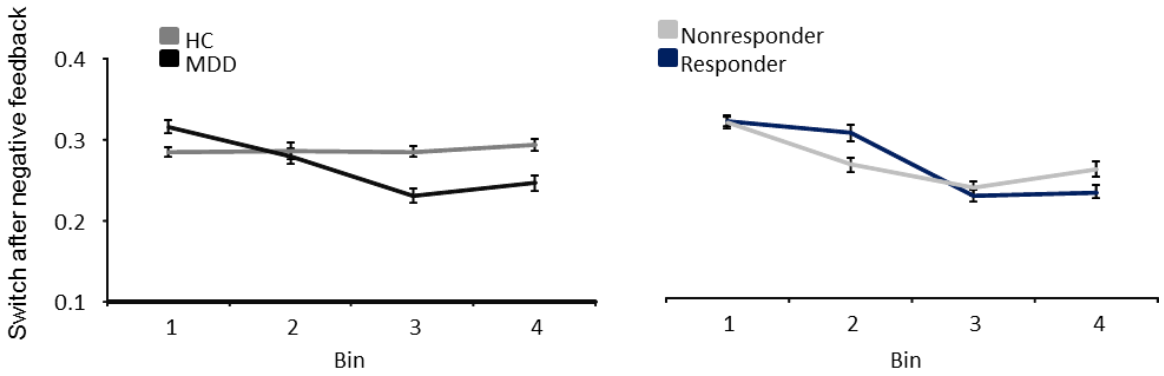

Figure 2. Mean number of switches after negative feedback (expressed here in proportion) decomposed as a function of bin and group: (a) HCs vs. MDD patients and (b) Responders vs. Non-responders. The error bar corresponds to 1 standard error of the mean.

Learning rate and exploration parameter. There was a main effect of Feedback Valence (F(1, 77)= 145.93, p<.01, η^2=.66) showing a higher learning rate following positive feedback (M=0.32, SD= 0.23) than negative feedback (M=0.04, SD=0.08), replicating our previous results (Bakic et al., 2014). All the other effects remained non-significant (p>.05). The group comparison performed on the inverse-gain parameter (β) revealed no significant effect (p>.05).

Table 2. Correlations of β (randomness) parameter with accuracy, general switch rate and number of switches after negative feedback.

Learning measure	Correlation with β parameter	
	MDD	HC
Accuracy		
Deterministic	-.57**	-.49**
Probabilistic	-.72**	-.41**
Random	-.03	.02
Overall number of switches		
Deterministic	.65**	.44**
Probabilistic	.66**	.42**
Random	.40*	.30*
Switches after negative feedback		
Deterministic	-.01	-.06
Probabilistic	-.03	-.01
Random	-.20	-.12

*p<.05, **p<.01

Interestingly, β correlated negatively with accuracy for the deterministic and probabilistic conditions, and positively with the overall number of switches (for all three conditions together), but not with the number of switches after negative feedback (Table 2), corroborating the assumption that β reflects exploration indirectly. The responders were not significantly different from the non-responders on any of these two measures, learning rate and inverse-gain parameter (all p's>.05).

Ratings. Clarity ratings (Figure 3a) showed a significant Group*Condition interaction ($F(2, 154)= 3.04$, $p<.05$, $\eta^2=.04$) and a main effect of Condition ($F(2, 154)= 311.70$, $p<.01$, $\eta^2=.80$). Independent T-tests showed that in the deterministic condition, the HC group (M=77.09, SD=11.33) rated the S-R associations to be clearer than the MDD group (M=70.78, SD=13.93), ($t(77)=2.22$, $p<.05$). There was no significant group difference for the two other conditions. Overall, the deterministic condition (M=74.29, SD=12.86) had significantly higher clarity ratings than the probabilistic condition (M=65.97, SD=10.51), ($t(78)=5.61$, $p<.01$), which was itself higher than the random condition (M=38.79, SD=8.13), ($t(78)=18.76$, $p<.01$). Certainty ratings (Figure 3b) revealed a significant main effect of Group ($F(1,77)= 5.23$, $p<.05$, $\eta^2=.06$) showing that the HC group was overall more certain regarding the accuracy of the responses given than the MDD group. As expected, there was also significant main effect of Condition ($F(1.67, 128.52)= 60.80$, $p<.01$, $\eta^2=.44$), showing that, irrespective of group, certainty was higher in the deterministic (M=55.78, SD=15.70) than in the probabilistic condition (M=44.32, SD=13.64), ($t(78)=6.08$, $p<.01$), which was itself also significantly higher than in the random condition (M=33.87, SD=15.03), ($t(78)=6.45$, $p<.01$). Additionally, the HC group (M=40.73, SD=10.67) rated that they had received overall significantly more positive feedback than the MDD group (M=25.74, SD=9.84), ($t(77)=4.68$, $p<.01$). The HC group (M=52.74, SD=9.84) also reported liking the positive feedback significantly more than the MDD group (M=44.39, SD=23.73), ($t(77)=2.12$, $p<.05$). The two groups did not differ significantly with respect to how much they disliked receiving negative feedback ($p>.05$). Additionally, responders (M=17.84, SD=14.80) also rated receiving less often positive feedback than non-responders (M=31.01, SD=17.57), ($t(33)=2.31$, $p<.05$).

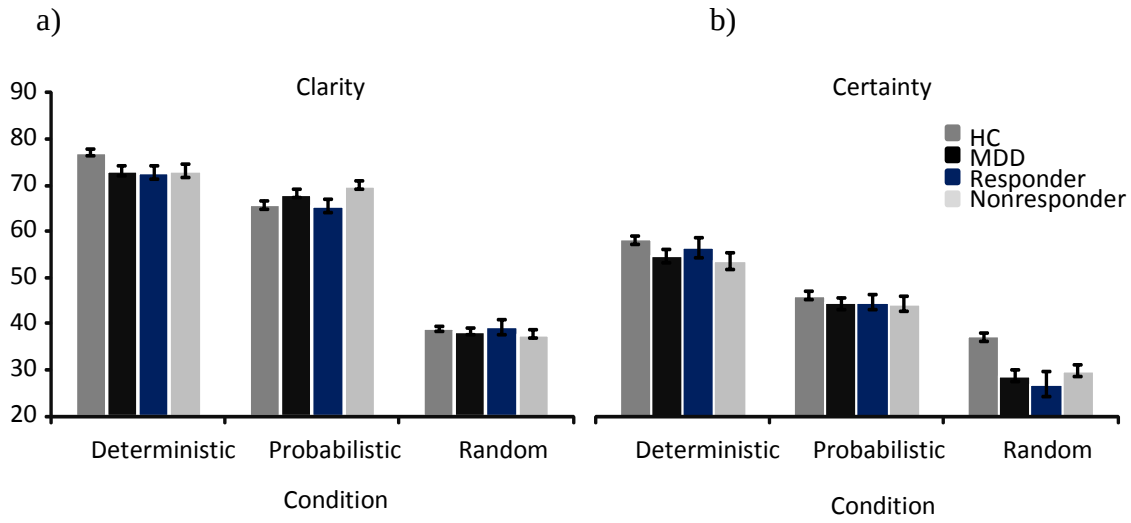

Figure 3. Clarity (a) and certainty (b) ratings decomposed as a function of condition and group. The error bar corresponds to 1 standard error of the mean.

3.2. ERP results

The analysis carried out on the ERN mean amplitudes at channel FCz showed a significant Condition*Accuracy interaction ($F(1.84, 139.98)= 34.59$, $p<.01$, $\eta^2=.31$), and main effects of Condition ($F(2,152)= 9.32$, $p<.01$, $\eta^2=.11$) and Accuracy ($F(1,76)= 49.25$, $p<.01$, $\eta^2=.39$). The main effect of Group was not significant ($p>.05$) (see Figure 4, two upper panels). In order to break-down the significant interaction effect, an analysis of the difference scores (incorrect-correct) was carried out and it showed a significant main effect of Condition ($F(1.59, 122.67)= 27.08$, $p<.01$, $\eta^2=.26$), while all the other effects were not significant (all p's $>.05$). As expected, the deterministic condition ($M=-2.19$, $SD=2.35$) had significantly higher ERN than the probabilistic condition ($M=-0.31$, $SD= 1.44$), ($t(78)= -6.45$, $p<.01$), which was significantly higher than the merely absent ERN in the random condition ($M=0.33$, $SD=2.40$), ($t(78)=-2.00$, $p<.05$). There were no differences in ERN amplitude at FCz related to treatment outcome either (Figure 4, two lower panels).

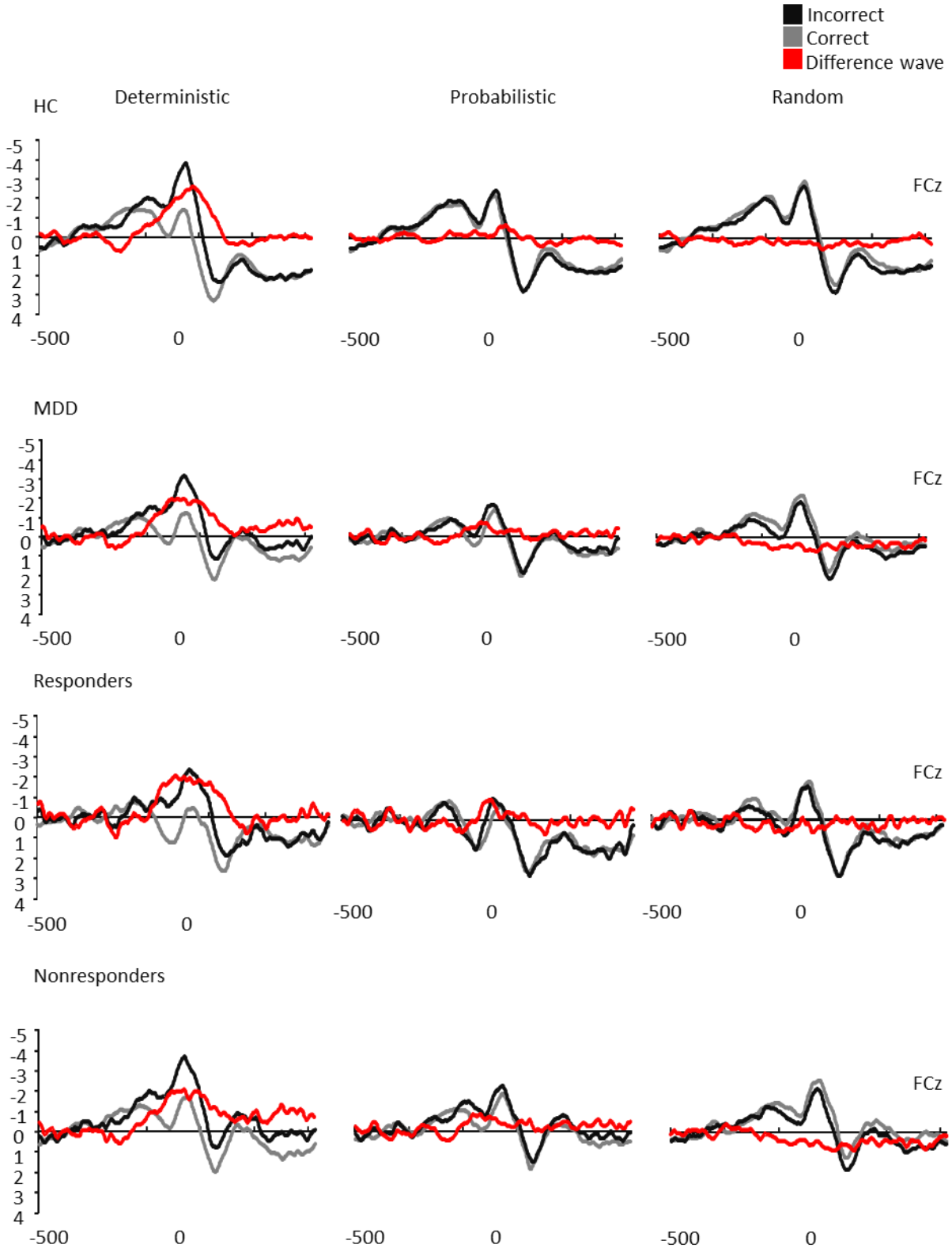

Figure 4. Grand average ERP waveforms for the response-locked ERP data (electrode FCZ), separately for condition and accuracy level for HCs, MDD patients, Responders, Nonresponders

In order to further explore depression-related differences in ERN size, we divided our MDD sample into a group of severe (n=17) vs. extremely severe levels of depression (n=18), a border marked by a Hamilton score of 22 (Hamilton, 1960). We performed a comparison between these two groups (relative to the HCs), by means of a repeated measures ANOVA. This analysis revealed significant main effects of Group ($F(1,2)= 3.61$, $p<.05$, $\eta^2=.10$), Condition ($F(2,150)= 12.10$, $p<.01$, $\eta^2=.10$) and Accuracy ($F(1,75)= 42.80$, $p<.01$, $\eta^2=.36$), as well as a significant Condition*Accuracy interaction ($F(1.84,137.94)= 27.90$, $p<.01$, $\eta^2=.27$). Noteworthy, the ERN was larger for group with moderate severity (M= -4,27, SD= 3.97) than in the group with extreme severity (M=-0.33, SD=2.87), ($t(33)=-3.38$, $p<.01$).

Topographical components. A solution with 5 dominant maps explained 94% of the variance. During the time interval corresponding to the ERN and CRN (from -20 ms to 100 ms following the response onset), two different main topographies were revealed by the cluster analysis, one expressed for the ERN (incorrect responses) and another one for the CRN (correct responses). While the CRN map was characterized by a broad and diffuse negative activity spreading over prefrontal sites (bilaterally), the ERN map had a slightly different configuration/geometry showing a more narrow/circumscribed distribution of the negative counterpart of the dipolar field over fronto-central electrode positions along the midline. This outcome is compatible with earlier ERP results reported in the literature for the topography of the ERN and CRN components (see Vocat et al., 2008; Aarts & Pourtois, 2010; Pourtois, 2011). This dissociation confirmed that partially non-overlapping neural networks were involved in the processing of correct vs. incorrect responses at the ERN/CRN level. Importantly, no group difference was evidenced for this dissociation. These two dominant maps were then fitted back to the individual ERP data to quantify their estimate across conditions and groups.

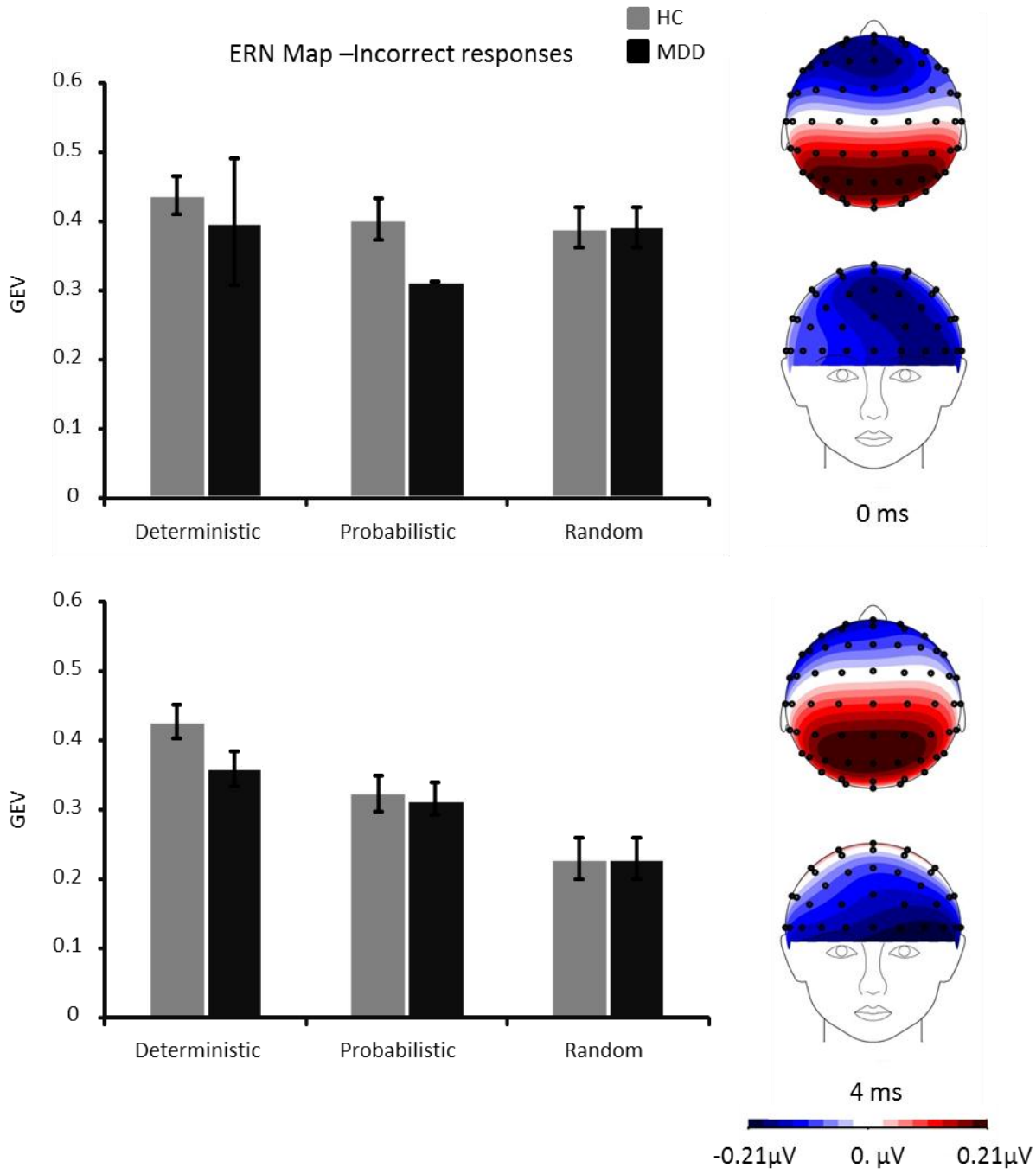

Figure 5. Main topographical components extracted after ERP data clustering (see methods for details) during the time interval (~100 ms post-response onset) corresponding to the ERN (response errors) or the CRN component (correct responses). The (spatial) correlation between the ERN and the CRN map is .90.

The results of the mixed ANOVA carried out on the GEV (Figure 5) values extracted after fitting showed a significant three way Accuracy*Condition*Map interaction ($F(1.79, 135.93) = 20.28$, $p < .01$, $\eta 2 = .21$). Additionally, there was a significant Accuracy*Map

interaction (F(1, 76)= 21.20, p<.01, η2=.22). Results of the pairwise t-tests used as a follow-up analyses for the three way interaction are presented in Table 3. Results showed that for the deterministic condition, the ERN map clearly explained more variance for incorrect responses than for correct responses, while the CRN map showed the exact opposite pattern. This accuracy-specific topographical effect was not found in the probabilistic and the random conditions. Noteworthy, these topographical effects did not differ between the two groups (HCs vs. MDD patients) (p>.05).

Table 3. Means (standard deviations in parenthesis) and paired samples t-test (df=77) for the GEV values obtained after fitting for the ERN and the CRN topographical map, separately for each accuracy and condition. P-values are Bonferroni corrected for multiple testing (p=.008)

Map	Accuracy	Condition		
		Deterministic	Probabilistic	Random
ERN	Incorrect	3.36 (2.83)	2.70 (2.88)	2.62 (2.71)
	Correct	1.88 (2.84)	2.33 (2.90)	2.64 (2.78)
	t-test	-5.27**	-2.12	0.18
CRN	Incorrect	1.14 (1.81)	1.93 (2.13)	1.87 (2.24)
	Correct	2.63 (2.28)	2.18 (2.13)	1.72 (2.14)
	t-test	5.73**	1.22	-1.10

The additional analysis taking into account treatment outcome for the MDD patients showed significant Accuracy*Condition*Map (F(1.88, 61.95)= 7.54, p<.01, η^2=.219), Accuracy*Map*Group (F(1, 33)= 4.51, p<.05, η^2=.12) and Accuracy*Map (F(1, 33)= 18.74, p<.01, η^2=.36) interactions. As shown in Table 4, while the accuracy specific effect for the ERN/CRN map was clearly evidenced in the group of responders (in the deterministic

condition only), this distinction was much weaker and it did not survive a stringent Bonferroni correction in the group of non-responders.

Table 4. Means (standard deviations in parenthesis) and paired samples t-test (df=35) for the GEV values obtained after fitting for the ERN and the CRN topographical map, decomposed for accuracy and condition in the MDD group. P-values are Bonferroni corrected for multiple testing (p=.008).

Group	Map	Accuracy	Condition		
			Deterministic	Probabilistic	Random
Responders	ERN	Incorrect	3.02 (1.94)	1.90 (2.28)	2.05 (2.40)
		Correct	1.24 (2.17)	1.12 (2.02)	1.77 (2.32)
		t-test	-3.47*	-1.49	-0.94
	CRN	Incorrect	0.77 (1.90)	2.14 (2.14)	1.80 (2.17)
		Correct	2.83 (1.90)	2.69 (2.0)	2.08 (2.10)
		t-test	3.33*	1.10	0.91
Non-responders	ERN	Incorrect	3.39 (3.24)	2.53 (2.82)	2.49 (2.60)
		Correct	2.05 (2.65)	2.54 (2.53)	2.41 (2.72)
		t-test	-2.30	0.02	-0.50
	CRN	Incorrect	1.20 (1.64)	2.25 (2.38)	1.84 (2.51)
		Correct	2.23 (2.39)	1.82 (2.14)	1.78 (2.44)
		t-test	2.21	-0.89	-0.25

4. DISCUSSION

RL assumes action-value to be learned in a habitual system, through (phasic) prediction-error-like signals (such as the ERN component) which are timely generated by dopaminergic-dependent regions in the human brain (Rangel, Camerer, & Montague, 2008). Because this mental process is tightly linked to the processing of reward and punishment incentives, MDD, which is characterized by a break-down in positive affect/reward processing (and an increase in negative affect), stands out as a unique mood disorder for the study of changes in RL

occurring as a function of emotional or motivational factors. In this study, we therefore sought to characterize changes in RL during a standard probabilistic learning task (Eppinger et al., 2008) in a group of treatment resistant unipolar MDD patients, when compared to age matched HCs. RL was studied at three levels concurrently: behavioral (including computational modeling), subjective (ratings and self-reports) and electrophysiological (ERN component). We hypothesized that learning rate could be different between MDD patients and HCs. More specifically, we tested the prediction that because negative affect prevails in MDD, these patients could show a larger learning rate following negative feedback than HCs, or alternatively a lower learning rate following positive feedback than HCs (or maybe both effects). Moreover, we predicted a larger ERN component in MDD patients compared to HCs, in keeping with earlier studies and models available in the literature linking overactive ERN to negative affect, harm avoidance, neuroticism or internalizing disorders broadly speaking, such as evidenced in MDD (Olvet & Hajcak, 2008; Vaidyanathan et al., 2012; Olvet, Klein, & Hajcak, 2010; Pizzagalli, 2014). In auxiliary analyses, we also took into account treatment outcome for the MDD patients (defined as MDD symptoms‘ severity reduction four weeks later) to analyze these (behavioral and EEG) data collected at baseline (prior to the start of the iTBS treatment), with the aim to assess whether the response rate could be predicted, at least in part, based on these baseline RL measurements (for a similar approach with fMRI data; see Downar et al., 2014).

Our first prediction, namely that MDD patients might show increase in the learning rate following negative feedback compared to the HCs, was not confirmed. Although MDD patients had overall slower RTs and more too late responses than HCs, which are often reported in the literature (Pizzagalli et al., 2009; Pizzagalli, 2014; Vanderhasselt et al., 2014), learning slope and accuracy were actually matched between the two groups. Moreover, neither learning rate, nor exploration (as captured by beta) differed between the two groups

either, when using a specific RL model able to extract these two parameters based on the single trial time series (see Bakic et al., 2014). The similar learning rate found between MDD patients and HCs casts doubt on the assumption that these MDD patients have trouble integrating the reinforcement history (especially when it is based on negative outcome or feedback) to guide learning (Padrão, Mallorquí, Cucurell, Marco-Pallares, & Rodriguez-Fornells, 2013). Previous studies (Pizzagalli et al., 2009; Treadway et al., 2012) have found that MDD patients have difficulty developing a positivity bias towards more favorable or rewarded responses. In the current experiment, we failed to observe this negative mood-related effect regarding approach motivation or reward processing. Noteworthy, an important difference between our study and previous ones is that monetary reward was often used in them (Padrão et al., 2013; Pizzagalli et al., 2009), while we did not do so in the present case. Our reward vs. punishment incentives were primarily related to the perceived task-success/failure, as opposed to primary or secondary rewards or punishments, the former which presumably activate more abstract motivational processes, and possibly more dorsal prefrontal cortical areas than the latter (see Sescousse, Caldú, Segura, & Dreher, 2013). Speculatively, being depressed could perhaps dampen reward processing and/or boost punishment sensitivity during RL if and only if incentives used to guide learning are disconnected somehow from core motivational processes, such as competence or self-determinism, and they therefore entail primarily monetary losses or gains (Deci & Ryan, 2002). Hence, paradoxically at first sight, the use of incentives bound to competence or self-efficacy might actually provide MDD with cues that boost (or challenge in the case of negative feedback) their motivation, and in turn yield a learning slope and accuracy being eventually comparable to that of HCs. At any rate, our results do not support a global or obvious deficit in the way RL is achieved in a group of severe treatment resistant MDD patients, including the processing of negative feedback (Mies et al., 2011). While the learning

rate (or accuracy) failed to evidence any group difference, we found however that MDD patients had a lower number of switches after negative feedback than HCs, during the second phase of the experimental session selectively (hence when learning of the arbitrary S-R associations was reliably established). Given the link between this metric and exploration (Gonzalez & Dutt, 2011; Hills & Hertwig, 2012), this result may point to a weaker or smaller exploration during RL in MDD patients than HCs, especially when this process is triggered by negative feedback on task/learning performance. Alternatively, this decrease might arise due to a drop in motivation in these patients, even though this explanation can then hardly account for the fact that MDD patients and HCs actually achieved the same level of learning at the end of the experimental session. Further, it is important to note that the overall number of switches (hence when taking into account both positive and negative feedback) was actually balanced between the two groups, ruling out the possibility of a drop in motivation or increase in fatigue accounting for this result. In a way, towards the end of the experimental session, MDD patients appeared to become less flexible in their decision making process, when it was driven by the processing of negative feedback. Remarkably, despite a learning performance that was matched with the HCs, and as has often been reported in the literature on MDD (Dunn, Dalgleish, Lawrence, & Ogilvie, 2007), these MDD patients judged that they had received less often positive feedback (and they then liked them less) throughout the experimental session than HCs (which was not the case obviously), unambiguously translating blunted positive affect at the subjective level in these patients. They also evaluated the clarity of the S-R associations in the deterministic condition to be lower than the HCs, and they felt overall less certain about the accuracy of their responses than the HCs, even though they actually performed equally high.

Concerning our second hypothesis, we did not find a larger (or smaller) ERN component in MDD patients compared to the HCs. More specially, in the deterministic

condition, where the clarity of the S-R associations was the highest (and therefore error commission the least ambiguous or uncertain), the size of the ERN was not different between MDD patients and HCs, challenging the assumption that early error-related brain activities (or reward prediction error signals) are "always" overactive in negative affect or internalizing disorders, and they can therefore be used as reliable endophenotypes for these disorders (Olvet & Hajcak, 2008; Vaidyanathan et al., 2012). As a matter of fact, abnormal or overactive ERNs in high anxious individuals or depressed patients are often reported in tasks devoid of learning, such as Flanker or Stroop, and for which response errors have probably a different "affective" or cognitive meaning than errors committed during a probabilistic learning task (see also Endrass & Ullsperger, 2014). In the latter case, unwanted response errors are not necessarily negatively connoted actions, but they actually serve as potent learning signals, unlike the former case. Accordingly, it is possible that negative affect influences (i.e., increases) the ERN when error making is happening in an experimental setting where response errors are somehow perceived (even implicitly) as negative events or acquiring a negative value (i.e., they reflect lapses of attention or concentration, or are perhaps associated with losses; see also Inzlicht & Al-Khindi, 2012), but not when these worse than expected events are instrumental to learning, as in the present case. Of note, in both groups, the ERN shows the expected amplitude dependency based on reward probability, suggesting that this ERP deflection was indeed related to reward prediction error in this task (see Eppinger et al., 2008; Bakic et al., 2014; Unger et al., 2012). Moreover, when considering the topography, as opposed to the peak of this ERP component, we found a clear interaction effect between the map configuration, accuracy and reward probability, equally strongly in both groups. Thus, in both groups alike, the ERN topography explained most of the variance for response errors in the deterministic condition, while the CRN topography did so for correct responses, with no such dissociation found for the probabilistic and random

conditions. This topographic ERP mapping analysis therefore confirmed the sensitivity of the ERN component to accuracy and reward probability concurrently, without any difference between the two groups though. Interestingly, when we split these ERP data according to depression severity at baseline (using the HRSD)[2], patients with severe MDD had lower ERN than patients with moderate MDD, suggesting that the relationship between depression/negative affect and the ERN is probably more complex than initially thought (Olvet & Hajcak, 2008; Olvet, Klein, & Hajcak, 2010), and not merely linear (i.e., the higher the symptoms' severity, the larger the ERN component). Similar results were already reported by Schrijvers and colleagues (2009), who found no increase in the amplitude of the ERN related to depression severity (in patients with comparable, high levels of MDD), nor to treatment.

When considering treatment outcome and analyzing therefore these baseline data separately for the responders vs. non-responders, we found interesting group differences which are worth discussing further, as they might shed light on the complex interplay between state effects (induced by the iTBS treatment in the present case) and trait or disposition effects pertaining to the patients themselves (see also Downar et al., 2014). Responders showed slightly better performance and shorter RTs during the probabilistic learning task at baseline compared to non-responders, even though this behavioral effect was marginally significant only. Importantly, error monitoring was clearly different between the two groups however. More specifically, in the deterministic condition, the responders had a sharper dissociation between the ERN (errors) and the CRN (correct responses) topographical map than the non-responders. This effect was not explained or conflated by MDD severity at baseline (see footnote 1). This result is potentially interesting in the light of the recent evidence showing

[2] Depression severity at baseline was not predictive of treatment outcome: a chi square statistic showed that the number of patients with moderate vs. high MDD severity at baseline did not explain the rate of responders vs. non-responders four weeks later ($\chi^2(1)=0.89$, $p>.05$).

that increased ERN (or better ERN-CRN difference) might be related to positive affect in healthy adult participants (see Bakic et al., 2014). Hence, a larger or clear ERN-CRN difference (already present at baseline) might translate "relatively spared" phasic activation within a dedicated reward prediction (error) system (Holroyd & Coles, 2002; Frank et al., 2005), which might be a necessary condition or pre-requisite in order to enable later beneficial iTBS-induced plasticity in this specific and remote brain networks (treatment effect), and in turn improve depressive symptoms (see Downar et al., 2014).

To conclude, this study shows relatively preserved cognitive abilities during RL in a group of treatment resistant MDD patients (characterized by high levels of Anhedonia), compared to a group of matched HCs. Both the amount of RL achieved (quantitative effects), as well as its underlying exploration-exploitation tradeoff (qualitative effects) are similar between these two groups. Therefore, we speculate that the lack of obvious impairments during RL as a function of MDD might be explained by "spared" motivational effects (Baeken et al., 2010) or processes in these psychiatric patients. Because reward and punishment are bound to task performance here (and not to losses or gains in money or points for example), it may be the case that "intrinsic" motivation is timely and dynamically engaged, and it compensates in turn for the otherwise abnormal reward (hyposensitivity) and punishment (hypersensitivity) usually observed in these patients. This conclusion awaits however direct empirical validation, and the use of experimental designs where RL could be compared directly between two situations or contexts: one where it is driven primarily by "intrinsic" motivation, as in the present case, versus another one where "extrinsic" motivation prevails (Pizzagalli et al., 2009). Presumably, MDD could impair the latter process to a larger degree, while leaving perhaps the former relatively spared, suggesting then context-dependent impairments in MDD during RL. Further, our study is also informative as it clearly shows that baseline differences (i.e., prior to the non-invasive neurostimulation treatment) in a standard

ERP correlate of reward prediction error (namely the ERN component) could partly predict treatment outcome four weeks later in these patients, regardless of MDD severity (see also Downar et al., 2014). We found that a blunted ERN-CRN differentiation in the deterministic condition at baseline in this probabilistic learning task (when considering the topography/configuration of these ERP components) was actually associated with a worse/unfavorable treatment outcome. This result is potentially clinically important/relevant because it suggests the existence of a specific predisposition (perhaps endophenotype) concerning the integrity or (still normal) functioning of a dedicated fronto-striatal brain network/loop (linking the rostral cingulate zone to deep subcortical structures in the basal ganglia), from which the ERN/CRN is thought to arise (Ullsperger et al., 2014). In other words, one of the electrophysiological correlates of error monitoring during RL (i.e., the ERN component) could very well be used (as a diagnostic tool) to assess whether treatment resistant MDD patients might eventually respond "positively" (i.e., to show symptoms' reduction in the short run, and possibly even remission in the long run) to accelerated theta-burst neurostimulation of their left DLPFC or not.

5. REFERENCES

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective & Behavioral Neuroscience*, *10*(4), 479–492. doi:10.3758/CABN.10.4.479

Aarts, K., Vanderhasselt, M.-A., Otte, G., Baeken, C., & Pourtois, G. (2013). Electrical brain imaging reveals the expression and timing of altered error monitoring functions in major depression. *Journal of Abnormal Psychology*, *122*(4), 939–50. doi:10.1037/a0034616

Baeken, C., De Raedt, R., Santermans, L., Zeeuws, D., Vanderhasselt, M. A., Meers, M., & Vanderbruggen, N. (2010). HF-rTMS treatment decreases psychomotor retardation in medication-resistant melancholic depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *34*(4), 684–687.

Bakic, J., Jepma, M., De Raedt, R., & Pourtois, G. (2014). Effects of positive mood on probabilistic learning: Behavioral and electrophysiological correlates. *Biological Psychology*, *103*, 223–232. doi:10.1016/j.biopsycho.2014.09.012

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the Beck depression inventory-II. *San Antonio, TX: Psychological Corporation*, 1–82.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. doi:10.1038/nn.3961

Cavanagh, Bismark, Frank, & Allen. (2011). Larger Error Signals in Major Depression are Associated with Better Avoidance Learning. *Front Psychol*, *2*, 331. doi:10.3389/fpsyg.2011.00331

Cavanagh, J. F., & Shackman, A. J. (2014). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology Paris*. doi:10.1016/j.jphysparis.2014.04.003

Conway, M., Howell, A., & Giannopoulos, C. (1991). Dysphoria and thought suppression. *Cognitive Therapy and Research*, *15*(2), 153–166.

Deci, E. L., & Ryan, R. M. (2002). Self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 3–34). The University of Rochester Press.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*(5), 303–305. doi:10.1111/j.1467-9280.1994.tb00630.x

Disner, S. G., Beevers, C. G., Haigh, E. a P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews. Neuroscience*, *12*(8), 467–77. doi:10.1038/nrn3027

Downar, J., Geraci, J., Salomons, T. V, Dunlop, K., Wheeler, S., McAndrews, M. P., … Giacobbe, P. (2014). Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biological Psychiatry*, *76*(3), 176–85. doi:10.1016/j.biopsych.2013.10.026

Elliott, R., Sahakian, B. J., McKay, A. P., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1996). Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. *Psychological Medicine*, *26*(5), 975–989.

Elliott, R., Sahakian, B. J., Michael, A., Paykel, E. S., & Dolan, R. J. (1998). Abnormal neural response to feedback on planning and guessing tasks in patients with unipolar depression. *Psychological Medicine*, *28*(3), 559–571. doi:http://dx.doi.org/10.1017/s0033291798006709

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

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

Fava, M. (2003). Diagnosis and definition of treatment-resistant depression. *Biological Psychiatry*, *53*(8), 649–659.

Frank, Woroch, & Curran. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. doi:10.1016/j.neuron.2005.06.020

Gard, D. E., Gard, M. G., Kring, A. M., & John, O. P. (2006). Anticipatory and consummatory components of the experience of pleasure: A scale development study. *Journal of Research in Personality*, *40*(6), 1086–1102.

Gehring, W. J., Goss, B., & Coles, M. G. H. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390. doi:10.1111/j.1467-9280.1993.tb00586.x

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–51. doi:10.1037/a0024558

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

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62.

Hills, T. T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: comment on Gonzalez and Dutt (2011). *Psychological Review*, *119*(4), 888–92; discussion 893–8. doi:10.1037/a0028004

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–96. doi:10.1162/jocn.2010.21548

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*(6), 609–621.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., … Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*(5), 651–660. doi:10.1016/j.neuron.2005.02.014

Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal Dynamics of Human Cognition. *News in Physiological Sciences : An International Journal of Physiology Produced Jointly by the International Union of Physiological Sciences and the American Physiological Society*, *14*, 206–214. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11390852

Mies, G. W., van der Veen, F. M., Tulen, J. H. M., Birkenhäger, T. K., Hengeveld, M. W., & van der Molen, M. W. (2011). Drug-free patients with major depression show an increased electrophysiological response to valid and invalid feedback. *Psychological Medicine*, *41*(12), 2515–25. doi:10.1017/S0033291711000778

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*(August), 466. doi:10.3389/fnhum.2013.00466

Mueller, E. M., Pechtel, P., Cohen, A. L., Douglas, S. R., & Pizzagalli, D. A. (2015). Potentiated processing of negative feedback in depression is attentuated by anhedonia. *Depression and Anxiety*. doi:10.1002/da.22338

Murphy, F. C., Michael, a., Robbins, T. W., & Sahakian, B. J. (2003). Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychological Medicine*, *33*(3), 455–467. doi:10.1017/S0033291702007018

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*.

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clinical Psychology Review*, *28*(8), 1343–54. doi:10.1016/j.cpr.2008.07.003

Olvet, D. M., Klein, D. N., & Hajcak, G. (2010). Depression symptom severity and error-related brain activity. *Psychiatry Research*, *179*(1), 30–37.

Padrão, G., Mallorquí, A., Cucurell, D., Marco-Pallares, J., & Rodriguez-Fornells, A. (2013). Neurophysiological differences in reward processing in anhedonics. *Cognitive, Affective & Behavioral Neuroscience*, *13*(1), 102–15. doi:10.3758/s13415-012-0119-5

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, *42*(7), 658–665.

Pizzagalli. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, (January), 1–31. doi:10.1146/annurev-clinpsy-050212-185606

Pizzagalli, Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2009). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87.

Pourtois, G. (2011). Early error detection predicted by reduced pre-response control process: An ERP topographic mapping study. *Brain Topography*, *23*(4), 403–422.

Pourtois, G., Delplanque, S., Michel, C., & Vuilleumier, P. (2008). Beyond conventional event-related brain potential (ERP): exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topography*, *20*(4), 265–277. doi:10.1007/s10548-008-0053-6

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556. doi:10.1038/nrn2357

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15109163

Santesso, D. L., Steele, K. T., Bogdan, R., Holmes, A. J., Deveney, C. M., Meites, T. M., & Pizzagalli, D. A. (2008). Enhanced negative feedback responses in remitted depression. *Neuroreport*, *19*(10), 1045–1048.

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

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The

development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. In *Journal of Clinical Psychiatry* (Vol. 59, pp. 22–33).

Sherdell, L., Waugh, C. E., & Gotlib, I. H. (2012). Anticipatory pleasure predicts motivation for reward in major depression. *Journal of Abnormal Psychology*.

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *The British Journal of Psychiatry*, *167*(1), 99–103. Retrieved from http://bjp.rcpsych.org/content/167/1/99.abstract

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: an introduction.* Cambridge, MA: MIT Press.

Tibshirani, R., Walther, G., & Hastie, T. (2004). Estimating the Number of Data Clusters via the Gap Statistic. In *BIOSTAT*.

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*.

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–67. doi:10.1016/j.tics.2014.02.009

Unger, K., Heintz, S., & Kray, J. (2012). Punishment sensitivity modulates the processing of negative feedback but not error-induced learning. *Frontiers in Human Neuroscience*, *6*(June), 1–16. doi:10.3389/fnhum.2012.00186

Vaidyanathan, U., Nelson, L. D., & Patrick, C. J. (2012). Clarifying domains of internalizing psychopathology using neurophysiology. *Psychological Medicine*. doi:10.1017/S0033291711001528

Vanderhasselt, M.-A., De Raedt, R., De Paepe, A., Aarts, K., Otte, G., Van Dorpe, J., & Pourtois, G. (2014). Abnormal proactive and reactive cognitive control during conflict processing in major depression. *Journal of Abnormal Psychology*, *123*(1), 68–80. doi:10.1037/a0035816

Von Helversen, B., Wilke, A., Johnson, T., Schmid, G., & Klapp, B. (2011). Performance benefits of depression: sequential decision making in a healthy sample and a clinically depressed sample. *Journal of Abnormal Psychology*, *120*(4), 962–8. doi:10.1037/a0023238

Weinberg, A., Riesel, A., & Hajcak, G. (2011). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. *Motivation and Emotion*, *36*(1), 84–100. doi:10.1007/s11031-011-9269-y

CHAPTER 6

GENERAL DISCUSSION

"Toto, I don't think we're in Kansas anymore."

"Wizard of Oz", the movie

1. General summary

The above quote represents one of the probably most famous vocalizations of prediction error (PE) ever. Reading it, one can almost feel the fear and confusion amounting with the creeping, tiding realization that things are not what one expected them to be, in a way that is crucial for one's existence. Poor Dorothy's feedback related negativity (FRN) must have been through the roof! But, had she decided not to leave the mundane, familiar Kansas, she would have deprived herself of one of the most adored adventures of the world's cinematography. It is moments like these that create new insights, and those immensely valuable learning experiences that we depend on for our very survival. It is time for us to leave Kansas as well. Armed with positive mood, it is time to step out of the safety of the knowledge gathered in the previous four experimental chapters, summon the results, and ready ourselves for some well-prepared, goal-directed exploration. The foundations of the yellow brick road have been laid, and now it is time to take critical look at our first steps, and plan the ones that lay ahead.

The goal binding together the four experimental chapters of this thesis was to examine the effects of positively (and negatively) valenced mood on several components of reinforcement learning (RL) using a standard probabilistic learning task. The sources of motivation for this idea were three-fold. Firstly, there are many high-level cognitive processes for which affect-related modulation is evidenced in empirical studies and theoretical models in psychology. However, dominant learning theories usually did not take into account these "hot cognition" effects (especially when considering mood), and as a result, it was therefore unclear what their specific role or contribution to learning mechanisms might be. Secondly, not much is known about the effects of positive mood in comparison to negative mood or affect, which has prevailed for a long time in the literature. The importance of happiness for wellbeing and health, as well as many aspects of human cognitive functioning, and therefore adaptation to different challenges, has only recently become a focus of intense research in psychology. In fact, it is hard to shake off the impression that positive affect has been of interest especially in the field of work psychology, and of course clinical, therapy related work. By comparison, systematic research on this topic in cognitive or experimental psychology is scant. As a result, a clear understanding of the exact neuroanatomical and functional underpinnings of these effects is still lacking. The initial ideas need to be subjected to experimental scrutiny to discern the actual mechanisms and processes involved. And finally, juxtaposing the consequences of healthy affective experiences of either valence to pathological affective states can have important implications to our understanding of affect-related pathology, and differentiation between different disorders. In this respect, studying mood, probably more than any other aspect of affective experience, can be particularly valuable.

As the literature on this specific topic, namely positive mood- RL integration, is scarce, our initial hypotheses were inevitably broadly formulated: will positive mood change

the so-called exploration-exploitation trade-off in a way that mimics dopamine boosts in the anterior cingulate cortex (ACC) and interconnected brain regions at the cortical and subcortical levels? Does positive mood lead to the creation of a safe, exploration-favoring cognitive style or mental set, or does it lead to worse performance due to distraction away from the main task goal? Do participants in a positive mood learn better than the ones in a neutral affective state? Do they learn better exclusively from actions associated to rewarding, rather than punishing feedback, in a valence-specific way? Or do they learn equally well as participants in a neutral mood, but negative mood worsens RL due to worry related disruptions in active goal-maintenance? Do participants in a positive mood experience reward as more rewarding, and the ones in a negative mood punishment more punishing? Do they consciously appraise these learning signals differently? These were the type of questions we wanted to address in this thesis in a systematic way, using a stringent experimental approach in psychophysiology, effects of (positive) mood on mechanisms of RL in humans.

1.1. Guided imagery (GI) as potent means to induce and sustain positive mood

Unlike emotions that can be induced in a relatively quick and easy way, moods, as slow moving background states (Kohn et al., 2014), are not so easy to induce, and yet, induction techniques are crucial for examining their effects on brain and behavior, as well as their relation and differentiation to mood disorders. Most of the knowledge of the effectiveness of the mood induction procedures (MIPs) currently available in the field comes from review studies, with few direct experimental comparisons however (Westermann, Stahl, & Hesse, 1996). Therefore, our first challenge was to find and validate a way to induce positive mood in a potent and reliable way in (unselected) healthy adult participants, that would enable to

study changes in RL as a function of different mood states. The results of Chapter 1 showed that GI is as efficient in inducing positive mood as watching happy/funny movie clips, a more broadly used MIP, but that this former method is more personalized and tailored to the background and specificities of each individual participant, making it in turn less automatic and more engaging for higher cognitive functions like voluntary effort, social cognition/empathy, and episodic memory. Moreover, unlike movie clips, GI does not require any external medium or support; the mood is actually elicited based on the active "internal" reliving of personal memory episodes. Additionally, in Chapter 1, we replicated the results previously reported in Eppinger, et al. (2008) and confirming the dominant RL theory put forward by Holroyd and Coles (2002), where the amount or rate of learning was found to depend strongly on reward probability [linking specific stimulus (S)-to specific response (R)], which is a hallmark of RL (Holroyd & Coles, 2002; Sutton & Barto, 1998). In this study (Chapter 1), using standard learning parameters, such as accuracy, reaction times (RT), and the number of switches after negative feedback (a metric that reflects exploration; see Gonzalez & Dutt, 2011; Hills & Hertwig, 2012), we nevertheless found no strong modulatory effect of (positive) mood on RL, although the MIP used (based on GI) turned out to be efficient.

1.2. Positive mood increases learning rate and the ERN component

The lack of systematic group difference in RL depending on mood in Chapter 1 led us to search for alternative measures of RL that would perhaps be more sensitive to mood-related modulations. Note that a priori, we did not necessarily expect that the short-term induction of positive mood would lead to a higher learning accuracy per se, compared to a control condition with a neutral mood state. Rather, we were interested in showing that the positive,

compared to the neutral mood, might be accompanied by the usage of different routes or strategies during RL to reach the same goal; hence we surmised that mood could be associated with qualitative changes during RL (as opposed to quantitative ones). In order to prove that, in Chapter 2, we therefore focused on finer-grained measures of RL and eventually assessed if positive mood could influence either exploration or exploitation (or perhaps both), besides the amount of learning. To this aim, we compared the two groups (positive and neutral mood) with respect to learning rate and exploration (or randomness of choice), which provide two complementing parameters that derive from computational modeling (Jepma & Nieuwenhuis, 2011; Sutton & Barto, 1998) and allow to gain insight into the strategy likely at stake during RL (as opposed to the amount of learning only). Using these mathematical parameters, we found that positive mood, compared to a neutral mood state, led to a significant increase in the learning rate for both positive and negative feedback, showing that in this specific context (namely in a probabilistic learning task), there was no proof of a mood congruency effect that would imply a valence-specific change of this (or the other) parameter. To correctly interpret this change in learning rate depending on (positive) mood, it was important to consider the demands imposed by the task we have chosen (Eppinger et al., 2008). Even though the S-R associations are of probabilistic nature, and not easy to discern- they are the same from the beginning to the end of the task, hence they remain constant throughout it. Thus, once the participant internalizes the rule, or learns the different (hidden) S-R associations, the learning demands are drastically decreased, unlike, for example, in reversal paradigms or more volatile learning environments (Behrens, Woolrich, Walton, & Rushworth, 2007; Cools, Lewis, Clark, Barker, & Robbins, 2007). In that sense, this paradigm offers a stable learning setting, so it is expected that learning rate will decrease with time: new encounters of the same S-R associations bring no further knowledge to the learning history. In a way, α (corresponding to the learning rate) is a measure of reinforcement efficacy; it shows the degree of reinforcement

induced by the reward and punishment (Beeler, 2012), here conveyed by specific evaluative feedback on task performance. The fact that positive mood led to increases in both positive and negative learning rates therefore meant that in the positive mood group, even though the (total) learning level was comparable to the neutral mood group, the more recent information was probably used more intensively for choosing the next response as compared to the history of previous responding. In other words, participants in the positive mood group seemed to be more sensitive to the alternations in the reinforcement value compared to the neutral mood group, even when the internalization of the task rules had been achieved. Behrens and colleagues (2007) previously argued that higher response to volatility in the outcome phase, as measured by functional magnetic resonance imaging (fMRI) increased ACC activity when subjects are awaiting the outcome of a trial, is matched with higher average learning rates. The individuals with a greater effect of volatility in the ACC during the monitor/exploration period had a higher mean learning rate, and therefore gave more weight to the most recent piece of information (see also Quilodran, Rothe, & Procyk, 2008). By comparison, we found no effect of positive mood on randomness/exploration, or β parameter, which we would expect if we would assume that the encounter and experience of positive mood simply mimic somehow the neural effects produced by tonic changes in the endogenous level of dopamine in the prefrontal cortex (Ashby, Isen, & Turken, 1999). This null finding was surprising to some extent, as we were expecting that positive mood could act as a signal for the presence of a safe environment (or environment where reward or opportunities are available), which should in turn boost its exploration.

In Chapter 2, RL was not only characterized at the behavioral (and computational modeling) level, but also at the electrophysiological one concurrently, capitalizing directly on amplitude changes of the ERN and/or FRN component as a function of reward probability (see Frank, 2005; Holroyd & Coles, 2002). Results showed that participants in the positive

mood group had a larger error-related negativity (ERN) than in the neutral mood group, however in the deterministic condition exclusively. In other words, positive mood seemed to make participants more responsive/sensitive to internal markers of error commission in the case where the response error was presumably clearly recognized or perceived as such based on a motor or mental representation of the action. This neurophysiological result was important because it showed a larger ERN amplitude associated with positive mood. So far, such boost in reaction to internal indicators of errors was mostly associated with internalizing disorders or even traits/dispositions (at the subclinical level), such as depression and anxiety (Aarts & Pourtois, 2010; Moser, Moran, Schroder, Donnellan, & Yeung, 2013; Olvet & Hajcak, 2008; Vaidyanathan, Nelson, & Patrick, 2012), which obviously provide (psychopathological) conditions that stand in sharp contrast to the experience of happy mood by healthy adult participants, as we have used throughout this thesis (see Chapters 1-3). Accordingly, our new ERP results (Chapter 2) suggested that it is not merely the valence of the current affective state or condition that determines this neurophysiological effect (with a putative link between negative affect and an enhanced or overactive ERN), but that "affect", as an extension of motivation circuitry, likely shapes the rapid and automatic reaction to less than favorable (motor/internal) outcome. It has already been proposed in the motivational theory of ERN (and FRN) that the (perceived) salience of error events is also an important factor, besides negative affect, influencing the size of this ERP component (Olvet & Hajcak, 2008). Hence, in our study (Chapter 2), response errors could have been perceived as more salient events after the induction of positive mood, when these (unwanted) errors occurred in a context when the probability of getting a reward was high (i.e., in the deterministic condition). Although the FRN amplitude was numerically larger in the positive compared to the neutral mood group during RL, this difference did not reach significance.

1.3. A specific contribution of arousal to changes in RL as a function of mood?

Encouraged by the observation of multiple differences (i.e., learning rate and ERN component) during RL between the happy and the neutral mood group in Chapter 2, we set out to explore these differences further (in Chapter 3) with the aim to understand their functional significance (as well as possible implications) better. As primary goal, we also wanted to replicate these findings (regarding the learning rate) in a second (behavioral) study, using independent samples. Moreover, we added a third group of participants that received a sad mood induction to provide a clear(er) contrast effect with the happy mood group. Crucially, in Chapter 3, we also adapted the probabilistic learning task used so far (Eppinger et al., 2008) in combination with the MIP, in a very specific way. We surmised that removing the evaluative feedback once learning was established might yield or exacerbate opposite (RL) effects in the happy and sad mood group. We conjectured that this manipulation (meant to foster exploitation and block or neutralize exploration) might be especially detrimental for the latter, but not so much for the former group. In other words, while the (affective) need for exploration and external confirmation of the decision taken could prevail in the sad mood group (given the negative affect, rumination and uncertainty characterizing the current mood state), the happy mood group could perhaps more flexibly cope with this (sudden) change in the environment and eventually not suffer so much from the withdrawal of external rewards and/or punishments provided to guide learning. Nevertheless, the results obtained (Chapter 3) did not confirm these predictions, even though the MIP was found to be especially successful (with clear-cut mirror symmetric effects regarding the mood change produced by it in the happy and the sad mood group). First, we failed to replicate a differential (larger) learning rate in the happy compared to the neutral mood group. However, a systematic comparison between the results of Chapter 2 and Chapter 3 enabled us to gain insight into this apparent

discrepancy. As a matter of fact, we found out that arousal (for the happy mood group) was reliably different between these two studies, quite unexpectedly. More specifically, the arousal level of the happy mood group (following the MIP) was significantly lower in Chapter 3 than Chapter 2. Moreover, while arousal was larger for happy than neutral participants in Chapter 2, this difference was no longer evidenced in Chapter 3: there was no significant differences in terms of arousal between the happy, the neutral, and the sad mood group following the MIP, for reasons that were not entirely clear to us at first sight. More importantly, the importance of arousal, besides valence, was confirmed a posteriori by correlation analyses (Chapter 2) showing that increases in arousal did predict boosts in the learning rate (following both positive and negative feedback), a relationship that was absent in Chapter 3. Interestingly, the unexpected difference in arousal levels between the two studies also likely created imbalances in the conscious appraisal of positive and negative reinforcers. Only in Chapter 2 did participants in the positive mood group overestimate (when asked using specific post-experiment subjective ratings) the amount of positive feedback they received during the experimental session, as well as did report stronger positive and negative emotional reactions regarding the encounter of rewarding and punishing feedback, respectively. Hence, while the MIP led to the expected mood change in terms of valence, both the level of arousal reached and the conscious appraisal of the different reinforces used were found to be milder in the positive mood group in Chapter 3 than in Chapter 2. The sad mood group (Chapter 3) showed no different learning rate either compared to the happy or the neutral mood group. What sad mood did do however, was to decrease reaction times (RT) uniformly (across all conditions)- without altering accuracy however. This latter result might be taken as evidence that negative mood perhaps led to a more narrow focus of attention, even though additional studies are needed to confirm this interpretation. Secondly, we found that removing the evaluative feedback once learning was established did not impair the sad mood group, nor did

it influence learning performance in the happy mood group. Remarkably, in all three groups alike (hence regardless of the specific mood state induced), learning still progressed/increased after the withdrawal of the evaluative feedback information, unambiguously translating the use of exploitation (or some internal knowledge) to guide RL during this phase of the experiment.

1.4. The interplay of (abnormal positive) mood with (intrinsic) motivation during RL

In Chapter 4, we adopted a radically different approach- rather than inducing (positive) mood and explore its effects on RL in healthy adult participants (see Chapters 1-3), we studied effects of mood on this process by capitalizing on a patient population characterized by profound alterations in positive mood/affect (and Anhedonia), namely treatment resistant unipolar major depressive disorder (MDD) patients. This way, we could circumvent the problems related to mood induction, and "directly" test the prediction that Anhedonia, or the pathological and chronic lack of pleasure or reward experienced for stimuli or activities that were previously associated with positive emotion, causally changed the strength and expression of RL. Again, in analogy to Chapter 2, we used a multi-level characterization of RL in Chapter 4, using behavioral, computational modeling and ERP markers/correlates. Regarding the latter, it is important to note that although negative affect (or internalizing disorders or dimension of psychopathology) was repeatedly found to augment the ERN component (Weinberg, Riesel, & Hajcak, 2011), for MDD, the picture is usually much less clear and previous ERP studies have reported rather mixed results. This statement is especially true for studies focused on RL, and effects of MDD on this process when it is titrated at the ERP level. The results of Chapter 4 showed that compared to a group of age-matched healthy controls, the ERN and FRN components of a large group of very severe

(treatment resistant) MDD patients were not deemed abnormal during RL. The same conclusion was drawn regarding learning rate and exploration. Accordingly, these new results were important because they challenged the assumption of a generic deficit or impairment during reward processing in MDD (Pizzagalli, 2014). To account for this spared ability in MDD, we reckoned that the specific experimental context and task settings used (Eppinger et al., 2008) may paradoxically have fostered the use of preserved (intrinsic) motivational processes in these patients to solve the challenges posed by the probabilistic learning task, eventually enabling them to compensate somehow for an otherwise abnormal reward processing, and this way eventually reach the same level of performance compared to the healthy controls. At any rate, we acknowledged that this interpretation awaits further empirical validation before the contribution of specific motivational processes in MDD during RL could be confirmed. We will come back to this issue later in this closing section.

2. Labile effects of (positive) mood on reinforcement learning: the aftermath

From the experimental results reviewed here above and gathered in the four empirical chapters forming the bulk of this thesis, new insights regarding effects of (positive) mood on RL have been gained, and specific conclusions can therefore be drawn. First, the MIP elected in this work and based on GI was without doubt successful in inducing either happy or sad mood, but also in providing an active control condition characterized by a neutral mood state or content against which effects of happy (or sad) mood on RL could be compared properly. This latter (control) condition was an important pre-requisite at the methodological level to ascertain that effects of positive mood on RL were not merely resulting from uncontrolled changes (between the different mood groups) in motivation or attention processes. Even though its effects were titrated at the subjective level (using dedicated visual analog scales)

only, we could replicate them across different studies and independent samples, each time with a potent change of the mood state resulting from the MIP being reported, further emphasizing its valuable properties to alter the current mood state of adult healthy participants. Secondly, in each and every study performed, we were able to replicate the (behavioral and event related brain potential -ERP) results previously reported by Holroyd & Coles (2002) and Eppinger et al. (2008), and showing that (the amount of) learning was clearly dependent upon reward probability linking specific S to specific (arbitrary) R in this task; this linkage being reinforced by means of evaluative feedback (on task performance). Hence, the experimental paradigm chosen was particularly suited to study mechanisms of RL. Moreover, RL was established at multiple levels concurrently, including (standard) behavioral measures of accuracy, the number of switches after negative feedback, as well as subjective ratings pertaining to the clarity and certainty of the S-R associations learned and reinforced differentially throughout the task. Similarly, the electrophysiological correlates of RL were evidenced each time (see Chapters 2 and 4). The amplitude of the event related negativity (ERN) was the highest in the deterministic condition, lower/intermediate in the probabilistic one, and strongly attenuated in the random condition, as expected (see Eppinger et al., 2008; Frank et al., 2005). By comparison, the amplitude of the FRN showed the exact opposite variation (see Chapter 2), unambiguously translating their dependence on RL. However, although the MIP was successful each time (or positive mood was strongly altered from the outset in Chapter 4 without the need to induce it) and the chosen probabilistic learning task yielded the desired RL effects (at the behavioral and ERP levels), we have to acknowledge that when we consider the four empirical chapters together, evidence for a strong effect of mood on this process is actually lacking. As they turn out, the changes brought about by the MIP have only limited transfer to RL. Likewise, in a pathological condition where negative affect prevails and positive affect is basically missing or strongly reduced (namely MDD, see

Chapter 4), RL mechanisms do not appear to be compromised by this debilitating state or condition. Hence, we can conclude that mood (or a specific disorder of mood) does not influence RL mechanisms in a transparent, direct or straightforward manner. Even though the reasons underlying this observation are probably multiple and complex, hereafter we outline possible explanations for it, starting first with methodological considerations and ending up with more theoretical issues pertaining to the rather loose connections between positive affect/mood, enhanced cognition and dopamine.

2.1. Methodological reasons

2.1.1 Primary or secondary reward versus self-efficacy or self-determinism as reward signal

Very often, modulatory effects of reward on cognition and behavior (including learning) are actually explored by means of providing participants with money or points (either on a trial by trial basis or using block designs), hence by using what is usually called secondary reward in the literature (Botvinick, 2007; Christakou, Brammer, Giampietro, & Rubia, 2009; Frank, Doll, Oas-Terpstra, & Moreno, 2009; Fröber & Dreisbach, 2014; Knutson & Greer, 2008; Sescousse, Caldú, Segura, & Dreher, 2013; Small et al., 2005; Stürmer, Nigbur, Schacht, & Sommer, 2011; Van Steenbergen, Band, & Hommel, 2009). This conclusion also holds for the study of psychopathology (Padrão, Mallorquí, Cucurell, Marco-Pallares, & Rodriguez-Fornells, 2013; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2009; Vrieze et al., 2013). However, not all rewards are the same or serve the same function or goal (see Sescousse et al., 2013), and it is therefore very well possible that effects of (positive) mood on RL (or the lack thereof) could tentatively be explained by the use of rather abstract reward or punishment

signals/reinforcers in our studies. More specifically, in our studies (see Chapters 1-4), we used a very different approach and rewarding type or content. Reward versus punishment incentives were not related to (winning or losing) money or points during learning, but instead to a primary motivational cue, namely self-efficacy, hence to self-determinism and (intrinsic) motivation (Deci & Ryan, 2002). We used evaluative feedback that "directly" informed about a potential (mis)match between the expected or desired and actual decision, without consequences or relation to a secondary goal or function, such as winning/losing money or points for example. In this context, this evaluative feedback is a real "primary" learning signal that conveys critical information about the goal-conduciveness (or goal-obstructiveness) of the action/decision just performed. Accordingly, the (external) reinforcers used throughout this thesis were deemed "abstract" in the sense of being related to self-determinism or efficacy, a situation that might mitigate a direct effect of (either positive or negative) mood on RL. As already pointed out in Chapter 4, we have good reasons to believe that this factor might have explained the lack of clear or obvious deficits during RL in MDD patients. Hence, it might be valuable in future studies to compare directly effects of this kind of (abstract) reward to more concrete ones (using for example secondary reward) on mechanisms of RL when a specific mood state is induced and sustained throughout the task.

2.1.2. Choice behavior versus decision making, and the need to move beyond the use of a simple two-alternative forced choice task

Besides the specific type of reinforcer used in our studies, it is also important to interpret our new empirical findings in light of the specifics of the probabilistic learning task used here (see also Eppinger et al., 2008). In this standard experimental paradigm (see Holroyd & Coles, 2002), "simple" decision making is requested. Participants have each time (for each trial) to decide between two response alternatives based on a specific learning history colored by

specific reward probabilities. In other words, the task demands remain rather simple: a two-alternative forced choice task (with time pressure) has to be carried out. While we found only limited evidence for a strong effect of positive mood on this process (decision making per se), it might be the case that choice behavior would be influenced by it to a larger extent in this same task. To this aim, an extension or adaption of this paradigm, such as previously performed by Frank et al. (2005), would be needed or desired. This extension entails the inclusion of choice trials (intermixed with the standard decision making trials) where participants are asked to choose/select the stimulus (usually out of two available) they actually prefer. By manipulating which two stimuli (and differential reward contingencies) are provided in the pair, one can assess (using both behavioral and EEG measurements) whether for example participants are deemed positive learners (they look mostly for rewards and opportunities), or instead negative learners (they mostly avoid punishments or threats) (Frank et al., 2005). Although speculative, positive mood could perhaps influence behavior during these choice trials (hence revealing an effect of positive mood on one facet of RL), to a larger degree than the simpler decision making trials we have focused on in this work. Alternatively, it might be valuable to explore in the future effects of (positive) mood on RL using more complex (and volatile) experimental paradigms, such as an n-armed bandit task for example (Jepma & Nieuwenhuis, 2011).

Although we supplemented each time the standard behavioral measures of RL with specific computational modeling parameters (learning rate and exploration), because a two-alternative forced choice task was always used (see Chapters 1-4), the extracted β parameter (estimating "exploration") necessarily corresponded more to the randomness of responding than to exploration per se. Accordingly, we would not like to conclude too hasty that positive mood did not influence exploration during RL for sure. Instead, our estimate of exploration (β) probably did not reflect exploration as such, given the specifics of the probabilistic

learning task used here (see also Eppinger et al., 2008). The notion that β reflected the randomness of choices/decisions more than exploration was confirmed indirectly in Chapter 4 where we carried out auxiliary correlation analyses. Interestingly, these analyses showed that β correlated negatively with learning success, but positively with the total number of switches. However, β did not correlate (positively) with the more direct measure of exploration that could be extracted from the task, namely the number of switches made after negative feedback (Gonzalez & Dutt, 2011; Hills & Hertwig, 2012).

2.1.3. Phasic RL effects versus tonic mood changes

The lack of systematic modulatory effects of positive mood on RL could also be imputed to the apparent discrepancy (or mismatch) in terms of time-course or temporal dynamics between the putative mood change (presumably lasting several minutes and being mild in intensity) and the (behavioral or ERP) markers of RL (unfolding on a millisecond/second time scale). The MIP used in our studies and based on GI (see Chapters 1-3) is supposed to create tonic brain changes in a set of interconnected brain structures (Foland-Ross, Cooney, Joormann, Henry, & Gotlib, 2014; Habel, Klein, Kellermann, Shah, & Schneider, 2005; Kohn et al., 2014; Phan, Wager, Taylor, & Liberzon, 2002), perhaps all innervated by dopamine (Fiorillo, Tobler, & Schultz, 2003). Likewise, the lower (positive) mood in MDD patients was deemed stable or chronic (Chapter 4). By comparison, our measures of RL (α, β, ERN, FRN) were deemed phasic, and they primarily depend on the integrity of a circumscribed (cortical) brain area, namely the dorsal ACC (Holroyd & Coles, 2002). Accordingly, it may very well be the case that our MIP did produce a tonic change in dopamine in a rather diffuse manner in the prefrontal cortex and interconnected cortical and subcortical structures, whose nature,

strength or origin may therefore have been suboptimal to influence reliably phasic changes of this neurotransmitter, as captured by the ERN or FRN component (or α and β). Perhaps the use of alternative EEG data analyses (based on time-frequency decompositions and the study of induced or endogenous neural effects; see Tallon-Baudry & Bertrand, 1999) that take into account longer epochs or time periods might provide a way to overcome this inherent limitation. We note however that it appears especially challenging to marry (better) RL (which is phasic in nature) and (positive) mood (which is typically outlasting RL effects, is labile and has no specific object).

2.2. The "magic triangle" positive affect - enhanced cognition – dopamine revised

Given that we found little (or not consistent) evidence for systematic transfer effects of (positive) mood on RL across these four empirical chapters, this begs then the question whether, after all, the MIP triggered the (brain) changes we hypothesized to take place from the beginning or not. Our initial idea was that putting someone in a certain mood, or examining (psychiatric) patients with lowered (positive) mood level as "default mode", could in turn influence the way incentives or reinforces are eventually used to guide reward-based learning. Crucial to this reasoning was the assumption that inducing positive mood leads to a tonic increase in dopamine levels in the human brain, and more specifically the prefrontal cortex (see Ashby et al., 1999). Moreover, independent evidence clearly links this specific neurotransmitter to RL (Frank, 2005; Krugel, Biele, Mohr, Li, & Heekeren, 2009; Rangel, Camerer, & Montague, 2008). As a matter of fact, neither for the first, nor for the last statement we have actually obtained direct empirical evidence in this thesis. However, based on the results of related studies using similar MIP (see Foland-Ross et al., 2014; Habel et al.,

2005; Phan et al., 2002; Ridderinkhof et al., 2012), we can presume that our MIP was indeed accompanied by a tonic increase of dopamine in the human brain.

Moreover, there are good reasons to believe that this tonic dopaminergic change as a function of the MIP was mostly influencing the ACC and interconnected brain regions (Amiez et al., 2013; Quilodran et al., 2008), which play a critical role in the generation of the ERN (and FRN) component at the ERP level, as well as the learning rate and exploration parameter at the computational modeling level. Doya (2008) previously argued that the learning rate can be influenced by two changes; either by manipulating the environment or the subject's experience. Translated to our specific experimental set-up, the environment had presumably a more positive outlook in the positive compared to the neutral mood group, while previous experience was presumably balanced between the two groups (they both encounter the same task structure, with three hidden S-R probabilities). Doya (2008) further elaborated on the assumption that the tonic dopamine level provides a marker of opportunity cost, so that the more dopamine, the less the person/agent values potential damages, and the easier he/she takes risks (see also Mohr, Biele, & Heekeren, 2010). Moreover, Behrens and colleagues (2007) stated that the function of the ACC (and the ventral striatum, as well as the VTA) is to estimate volatility of the environment, whereas the action value signals are estimated in the putamen. Accordingly, we have good (and independent) reasons to expect that changes in volatility of the environment should be visible on the one hand in the amplitude of the ERP components originating from the ACC (such as the ERN or FRN; see Chapters 2 & 4), and on the other in the learning rate. Additionally, the dorsal ACC is also considered to be involved in self-regulation processes and modulation of emotional arousal (Luu, Collins, & Tucker, 2000; Shackman et al., 2011), thus making the prediction of mood-related changes arising in this specific medial frontal brain structure and influencing RL a valid one.

2.2.1. Multiple dopamine pathways and brain sites

However, the "easy" or straightforward link between positive affect and (enhanced) cognition (including RL), via changes in dopamine (within the prefrontal cortex), as originally put forward in a dominant theory in the field (see Ashby et al., 1999), might actually be too simple to account for mood-related changes during RL, or the lack thereof, as reported in this thesis. In their dominant model, Ashby et al. (1999) already referred to the existence of multiple dopaminergic pathways in the human brain, likely serving different functions. They contrasted the nigrostriatal pathway to the mesocorticolimbic one, providing projections from the VTA to several cortical and limbic brain areas. This latter pathway is the one considered to be primarily involved in affective modulations of cognitive processes and therefore, this is also the pathway via which we expected (positive) mood to influence mechanisms of RL. Interestingly, Ridderinkhof et al. (2012) recently showed that positive mood induction in Parkinson patients (whose nigrostriatal dopaminergic pathway is impaired) led to improvements in reward-based decision making, as if this mood manipulation compensated somehow for the abnormal dopaminergic neurotransmission, and in turn restored a cognitive process that strongly depends on it. On the other hand, the model of RL proposed by Holroyd and Coles (2002) stipulates that the basal ganglia (and more specifically phasic dopaminergic changes occurring therein) send critical inputs to the ACC, which is the locus of the ERN and FRN component. A central assumption of the RL theory is that upon error commission (or after the presentation of a negative evaluative feedback), a phasic depression of dopamine neurons takes place in the midbrain. As a result, there is a reduction of extracellular DA levels in the anterior midcingulate cortex, which can be registered as change in the amplitude of the ERN or FRN component. This system is associated with adjustments of the behavioral

outputs through motor controllers that take this dopamine drop as a signal that adaptation is timely needed. But this dopaminergic reward prediction error (RPE) signal is also conveyed to dorsal and ventral striatum, where it is used as a teaching signal, to improve outcome prediction (Ullsperger, Fischer, Nigbur, & Endrass, 2014). Of note, there are still some unresolved issues with this model, including the fact that the dopamine (neuro)transmission from the VTA to the anterior midcingulate cortex (aMCC) is probably too slow to underlie the generation of the ERN and the FRN, which therefore occur too fast to reflect the integrity of this specific neural system (Ullsperger, Danielmeier, & Jocham, 2014).

While several previous studies (see Foland-Ross et al., 2014; Habel et al., 2005; Phan et al., 2002) already pinpointed the ACC (and some other areas, including the amygdala and hippocampus) as the locus of MIP-related changes (using memory-based GI), other ones argued that the striatum actually lies at the origin of these changes (Floresco, Yang, Phillips, & Blaha, 1998). Positive mood induction was expected to increase tonically striatal dopamine, augmenting the availability of the DA molecules in the synaptic cleft, and as a consequence, creating phasic dips triggered by the response errors or negative evaluative feedback (Frank, Seeberger, & O'reilly, 2004; Frank, 2005; Ridderinkhof et al., 2012). As a result, it is expected that positive mood induction would actually lead to lower ERN and/or FRN amplitudes, which is the exact opposite of what we have found in Chapter 2 (i.e., ERN was increased by positive mood, for the deterministic condition only however).

Furthermore, even though many models previously emphasized the pivotal role of dopamine in RL (e.g., Holroyd & Coles, 2002; Schultz, 2010), other neurotransmitters probably also contribute to this process (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Ullsperger, Danielmeier, et al., 2014). Same is true for positive affect/mood that does not presumably depend upon a single neurotransmitter system. For example, recent studies

confirmed and emphasized the role of Serotonin as an important neurotransmitter in the experience of positive affect, including in healthy participants (see Haase et al., in press).

To sum up this part, even though there is independence evidence to believe that dorsal ACC activations (during RL) can be modulated by mood or affect (Shackman et al., 2011), there are many factors that might complicate this relationship, and by extension the existence of a direct link between positive mood on the one hand, and RL on the other. Among them, we have outlined the possibility that while positive mood might very well influence dopamine in the prefrontal cortex in a diffuse way, the actual locus in the human brain of phasic and dopaminergic effects related to RL is still a matter of intense debate and controversy in the literature (Ullsperger, Danielmeier, et al., 2014). Moreover, it is important to note that even if tonic mood changes would overlap somehow with phasic RL effects in the prefrontal cortex (as surmised by Ashby et al., 1999), including the ACC (a statement which awaits direct empirical validation), such an overlap would not necessarily entail that functional integration does indeed occur between these two process and thereby that (positive) mood alters RL. At any rate, clarifying these issues appears therefore necessary before one could assess in future studies if (positive) mood could eventually shape mechanisms of RL in a more systematic and reproducible way than what we have reported in this extensive work.

2.2.2. Arousal

Another reason likely explaining the limited impact of positive mood on RL may be found in the way mood was actually operationalized and measured in our studies (see Chapters 1-4). Following standard practice, we primarily emphasized (directional) changes in valence following the MIP, but merely overlooked arousal. Usually, a (right) combination of valence and arousal is however needed in order to produce the desired emotional effect (Knutson & Greer, 2008), which cannot be obtained by cranking up only one of these two components in

isolation. Presumably, mood makes no exception and also obeys this fundamental rule. Moreover, as has become apparent when comparing results of Chapter 2 and 3, arousal (titrated at the subjective level) might have played an important role in the strength of positive mood, and in turn the putative influences this specific mood state could have had on RL in our different studies. Noradrenaline-based arousal is usually considered to regulate exploration (Aston-Jones & Cohen, 2005; Jepma & Nieuwenhuis, 2011). A certain/intermediate/optimal level of arousal is needed in order to override the exploitative tendency (Daw et al., 2006). Aston-Jones and Cohen (2005) previously elegantly demonstrated that the Locus Coeruleus (LC) produces tonic increases in norepinephrine (NE), thereby mediating the level of arousal, given the widespread (reciprocal) anatomical connections between the LC and the cortex. Among them, there are prominent connections from the ACC and orbitofrontal cortex (OFC) to the LC, which are thought to optimize behavioral performance during decision making process. In this model, phasic LC bursts are associated with exploitation, while tonic changes are related to exploration. Hence, the effectiveness of the MIP could probably be increased further by taking into account levels of arousal. Possibly, "a right dose" of arousal (besides valence per se) could be needed in order to create a reliable mood change that would then be sufficient to influence reliably the exploration-exploitation tradeoff during RL. This issue too should be the focus of research in the future.

2.2.3. Intrinsic Motivation

Hints on the possible contribution of motivational processes to RL (besides mood) were obtained indirectly in Chapter 4. Very much like arousal, this effect was actually overlooked in our studies, which might eventually have weakened the demonstration of a strong link between positive mood and RL. As already alluded to in the General Introduction, we actually chose this specific experimental paradigm (Eppinger et al., 2008) to block or balance effects

of motivation across all three conditions (random, probabilistic and deterministic). Presumably, this manipulation was successful. However, because we used abstract reinforcers that tapped into self-efficacy (see point 2.2.1. here above), motivation (or engagement in the task) was actually not low, but rather high throughout the task. In other words, because the evaluative feedback timely informed about self-efficacy and learning success, it provided a very salient learning signal at the motivational level. By virtue of that, the (lingering) effects of (positive) mood in the background may have been weakened rather than increased, eventually accounting for a weak modulation of RL by positive mood found in our studies (see Chapters 1-3). Moreover, we believe that this framework could be used to account for the somewhat surprising results reported in Chapter 4 where treatment-resistant MDD patients had actually normal RL effects (at the behavioral and ERP levels) compared to healthy controls, again emphasizing no obvious link between mood (here a specific mood disorder) and RL. These MDD patients were actually enrolled in a treatment study based on neurostimulation (intermittent theta burst stimulation, iTBS), which was for most of them basically the last option left over to improve their (medical/psychological) condition and eventually restore their mood. Presumably, in such a (clinically-oriented) context, receiving feedback on task performance that inform about self-efficacy becomes "intrinsically" motivationally significant. For this specific reason, intrinsic motivation could then have overruled possible negative mood-related modulations of RL in these MDD patients. More generally, we have to acknowledge that additional research is needed to understand better the complex interplay of mood and motivation during RL. In this context, it might be valuable to measure specific (trait-related) motivational drives/tendencies (see Deci & Ryan, 2002) such as to assess whether they could perhaps contribute to better explain changes in RL as a function of (positive) mood.

3. General conclusions and recommendations

To conclude, while both the MIP and probabilistic learning task were found to be each successful, we have found little or limited evidence in this work for a strong and systematic modulation of mechanisms of RL by (positive) mood across the different empirical chapters, as if the mood effects induced did not influence RL directly. As emphasized here above at different places in this closing section, we believe however that the absence of (strong) evidence is not the evidence of absence. More specifically, as has become evident across our different chapters, different factors that were either overlooked or not considered carefully enough probably might eventually have mitigated modulatory influences of (positive) mood on RL. To end on a positive note, we therefore would like to formulate three specific recommendations, with the aim to increase the chance in future studies to reveal systematic effects of positive mood on RL. (i) As our new results suggest (see Chapters 2&3), arousal might be an important ingredient (very much like valence) of the emotional effect brought about by the MIP. Therefore, it appears essential to create potent MIP that deal with this factor, instead of assuming (like we did) that it did not play any role. In the same vein, measuring not only subjective but also physiological level of arousal might be beneficial in future work to evaluate the impact of arousal on RL, and its likely change by mood-related manipulations. Even though mood is usually characterized by a mild emotional intensity, the right level of arousal might actually be necessary to strengthen the mood's experience, and thereby reliably influence specific cognitive processes, including RL. (ii) Using a similar logic, we also would like to advocate for a better consideration (and perhaps theoretisation) of motivation effects influencing RL, or possibly weakening the link between mood and RL (see Chapter 4). In our task, participants (or psychiatric patients in Chapter 4) often showed a high level of engagement to solve it, presumably because the (external) incentives used to foster RL had a strong motivational value or drive for them. Maybe this (intrinsic) motivational or

"striving" effect outweighed the weaker mood effect present in the background. Accordingly, it appears important to better understand the complex interplay of motivation with RL in order to be able to assess then if mood provides yet another source of modulation on this complex and fundamental mental process. (iii) Finally, at the methodological level, it might be valuable in future studies to try to connect the mood changes better to the RL task itself (than what we did), for example by using mood contingency manipulations. In this context, an interesting avenue for future research is provided by EEG-based neurofeedback (for a recent example in psychopathology, see Konicar et al., 2015), where specific brain states can be measured online and then fed back to the participants (or patients) to teach them to alter a specific (mental) process. However, we still lack nowadays a good understanding of which brain changes (especially when considering scalp EEG measurements) are actually elicited during the encounter and experience of specific mood states. Hence, a better characterization of mood-related EEG states is probably required first, before it could later be used in combination with other tasks (such as the probabilistic learning task used in this work) by means of neurofeedback.

4. REFERENCES

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective & Behavioral Neuroscience*, *10*(4), 479–92. doi:10.3758/CABN.10.4.479

Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., & Procyk, E. (2013). The location of feedback-related activity in the midcingulate cortex is predicted by local morphology. *J Neurosci*, *33*(5), 2217–2228. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23365257

Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529–550. Retrieved from <Go to ISI>://000082133000004

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–50. doi:10.1146/annurev.neuro.28.061604.135709

Beeler, J. a. (2012). Thorndike's law 2.0: Dopamine and the regulation of thrift. *Frontiers in Neuroscience*, *6*(AUG), 1–12. doi:10.3389/fnins.2012.00116

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 356–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18189009

Christakou, A., Brammer, M., Giampietro, V., & Rubia, K. (2009). Right ventromedial and dorsolateral prefrontal cortices mediate adaptive decisions under ambiguity by integrating choice utility and outcome evaluation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(35), 11020–11028. doi:10.1523/JNEUROSCI.1279-09.2009

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. a, & Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 180–9. doi:10.1038/sj.npp.1301153

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), 876–9. doi:10.1038/nature04766

Deci, E. L., & Ryan, R. M. (2002). Self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 3–34). The University of Rochester Press.

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

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, N.Y.)*, *299*(5614), 1898–1902.

Floresco, S. B., Yang, C. R., Phillips, A. G., & Blaha, C. D. (1998). Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *The European Journal of Neuroscience*, *10*(4), 1241–1251.

Foland-Ross, L. C., Cooney, R. E., Joormann, J., Henry, M. L., & Gotlib, I. H. (2014). Recalling happy memories in remitted depression: a neuroimaging investigation of the repair of sad mood. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 818–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24146315

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

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*(8), 1062–8. doi:10.1038/nn.2342

Frank, M. J., Seeberger, L. C., & O'reilly, R. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science (New York, N.Y.)*, *306*(5703), 1940–3. doi:10.1126/science.1102941

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, 530–547. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24659000

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–51. doi:10.1037/a0024558

Haase, C. M., Beermann, U., Saslow, L. R., Shiota, M. N., Saturn, S. R., Lwi, S., … Levenson, R. W. (n.d.). *Short alleles, bigger smiles? Effects of 5-HTTLPR on smiling and laughing behavior. Emotion.*

Habel, U., Klein, M., Kellermann, T., Shah, N. J., & Schneider, F. (2005). Same or different? Neural correlates of happy and sad mood in healthy males. *NeuroImage*, *26*(1), 206–214.

Hills, T. T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: comment on Gonzalez and Dutt (2011). *Psychological Review*, *119*(4), 888–92; discussion 893–8. doi:10.1037/a0028004

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–96. doi:10.1162/jocn.2010.21548

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1511), 3771–3786. doi:10.1098/rstb.2008.0155

Kohn, N., Falkenberg, I., Kellermann, T., Eickhoff, S. B., Gur, R. C., & Habel, U. (2014). Neural correlates of effective and ineffective mood induction. *Social Cognitive and Affective Neuroscience*, *9*(6), 864–872. doi:10.1093/scan/nst055

Konicar, L., Veit, R., Eisenbarth, H., Barth, B., Tonin, P., Strehl, U., & Birbaumer, N. (2015). Brain self-regulation in criminal psychopaths. *Scientific Reports*, *5*, 9426. doi:10.1038/srep09426

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S.-C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17951–17956.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology. General*, *129*(1), 43–60.

Mohr, P. N. C., Biele, G., & Heekeren, H. R. (2010). Neural processing of risk. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(19), 6613–6619.

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*(August), 466. doi:10.3389/fnhum.2013.00466

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clin Psychol Rev*, *28*(8), 1343–1354. doi:10.1016/j.cpr.2008.07.003

Padrão, G., Mallorquí, A., Cucurell, D., Marco-Pallares, J., & Rodriguez-Fornells, A. (2013). Neurophysiological differences in reward processing in anhedonics. *Cognitive, Affective & Behavioral Neuroscience*, *13*(1), 102–15. doi:10.3758/s13415-012-0119-5

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*(2), 331–348.

Pizzagalli. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, (January), 1–31. doi:10.1146/annurev-clinpsy-050212-185606

Pizzagalli, Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2009). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87.

Quilodran, R., Rothe, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–325. doi:10.1016/j.neuron.2007.11.031

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*(7), 545–556. doi:10.1038/nrn2357

Ridderinkhof, K. R., van Wouwe, N. C., Band, G. P. H., Wylie, S. a., Van der Stigchel, S., Van Hees, P., … Van den Wildenberg, W. P. M. (2012). A tribute to Charlie Chaplin: induced positive affect improves reward-based decision-learning in Parkinson's disease. *Frontiers in Psychology*, *3*(JUN), 1–10. doi:10.3389/fpsyg.2012.00185

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions : BBF*, *6*, 24.

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

Shackman, A. J., Salomons, T. V, Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews. Neuroscience*, *12*(3), 154–167. doi:http://dx.doi.org/10.1038/nrn2994

Small, D. M., Gitelman, D., Simmons, K., Bloise, S. M., Parrish, T., & Mesulam, M.-M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. *Cerebral Cortex (New York, N.Y. : 1991)*, *15*(12), 1855–1865. doi:10.1093/cercor/bhi063

Stürmer, B., Nigbur, R., Schacht, A., & Sommer, W. (2011). Reward and punishment effects on error processing and conflict control. *Frontiers in Psychology*, *2*(NOV).

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning: an introduction.* Cambridge, MA: MIT Press.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*.

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

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–67. doi:10.1016/j.tics.2014.02.009

Vaidyanathan, U., Nelson, L. D., & Patrick, C. J. (2012). Clarifying domains of internalizing psychopathology using neurophysiology. *Psychological Medicine*. doi:10.1017/S0033291711001528

Van Steenbergen, H., Band, G. P. H., & Hommel, B. (2009). Reward counteracts conflict adaptation: Evidence for a role of affect in executive control. *Psychological Science*, *20*(12), 1473–1477.

Vrieze, E., Pizzagalli, D. a., Demyttenaere, K., Hompes, T., Sienaert, P., De Boer, P., … Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biological Psychiatry*, *73*(7), 639–645. doi:10.1016/j.biopsych.2012.10.014

Weinberg, A., Riesel, A., & Hajcak, G. (2011). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. *Motivation and Emotion*, *36*(1), 84–100. doi:10.1007/s11031-011-9269-y

Westermann, R., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity o f mood induction procedures : a meta- analysis. *European Journal of Social Psychology*, *26*(January 1995), 557–580. doi:10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

NEDERLANDSTALIGE SAMENVATTING

Invloed van positieve emotie op probabilistisch leren: een experimentele benadering

1. Inleiding

Een toenemende tendens in affectieve neurowetenschappen is om 'reinforcement learning' (RL) modellen te gebruiken om inzicht te verwerven in de effecten van affectieve componenten op cognitie. Dit is exact het doel van deze thesis. Op deze manier hopen we ook bij te dragen aan een groeiende literatuur die het belang van positieve emoties in verschillende domeinen erkent. Een beter begrip van hoe men leert in een bepaalde gemoedstoestand zal ons vermoedelijk helpen om nieuwe fundamentele kennis te verkrijgen over leren en positieve emoties, mogelijks met directe implicaties voor gezondheid en welzijn.

De invloedrijke "broaden and build" theorie van Frederickson (2004) stelt dat positieve emoties ons beschermen tegen een hoge mate van negatief affect (of angst) tijdens het verkrijgen van straf of wanneer een uitkomst slechter is dan verwacht. Omgekeerd kan de verwerking van beloning ook bevorderd worden in de context van positieve of gunstige factoren. Aan de andere kant van het spectrum, bij stemmingsstoornissen die gekenmerkt worden door een gebrek aan positief affect of Anhedonie, zoals depressie, wordt het verwerken van beloning negatief beïnvloed (Eshel & Roiser, 2010). In hun dominante theorie over de invloed van positief affect op cognitie (met een focus op creativiteit, of redeneren), stellen Ashby et al. (1999) dat mild positief affect kan leiden tot een toename in cognitieve flexibiliteit via de dopaminerge banen in de middenhersenen. Maar afgezien van hun kritische rol bij de mediatie van positief affect zijn diezelfde banen ook betrokken bij RL. Hieruit vloeit de hypothese dat een positieve gemoedstoestand leren kan beïnvloeden via gemeenschappelijke dopamine-afhankelijke banen (Ashby et al., 1999; Goschke & Bolte,

2013; Ridderinkhof et al., 2012). Experimentele evidentie die deze assumptie ondersteunt is echter gering.

De meeste theorieën van RL stellen dat optimalisatie van de prestatie gebaseerd is op de exploitatie van beloning, en de exploratie van minder gekende, maar potentieel betere alternatieven (Aston-Jones & Cohen, 2005; Behrens, Woolrich, Walton, & Rushworth, 2007; Cohen, McClure, & Yu, 2007; Jepma & Nieuwenhuis, 2011). Deze twee gelijktijdige processen, exploratie en exploitatie, hebben complementaire voordelen of functies: hoewel het belangrijk is om huidige doelen voorop te stellen en geen afleiding toe te laten (m.a.w. om exploitatie te bevorderen), is het ook belangrijk om de omgeving te controleren op mogelijke veranderingen die de prestatie mogelijks kunnen beïnvloeden (m.a.w. om exploratie te bevorderen). Onze hypothese is dat de balans tussen exploratie en exploitatie gevoelig is voor veranderingen in de gemoedstoestand van de deelnemer. Als een positieve gemoedstoestand zou leiden tot meer exploratie van minder gekende opties, en een negatieve gemoedstoestand tot een sterkere focus (Bolte & Goschke, 2010), gericht naar negatieve informatie, dan zou dit effect zichtbaar moeten zijn in de exploratie-exploitatie balans tijdens RL.

In alle studies voerden deelnemers een variant van de probabilistische leertaak uit, zoals opgesteld door Eppinger, Kray, Mock, and Mecklinger (2008), terwijl het electroencefalogram (EEG) werd geregistreerd, teneinde veranderingen in error en feedback gerelateerde negativiteit (ERN and FRN) in relatie tot leren en de gemoedstoestand te exploreren. In deze taak werden deelnemers gevraagd om verschillende verborgen stimulus (S) - respons (R) associaties te ontcijferen en leren, door ‘trial en error’. Voor elke beurt werden deelnemers, onder een constante tijdsdruk, gevraagd om te beslissen of de stimulus die op het scherm getoond werd geassocieerd was met respons 1 of 2 (overeenkomend met twee vooraf bepaalde toetsen op de responsbox). Feedback over de prestatie werd gegeven na elke respons. Deelnemers kregen 6 verschillende visuele stimuli te zien, die toegewezen

waren aan één van de drie condities (waarvan de deelnemers niet op de hoogte waren) die verschilden met betrekking tot de werkelijke (belonings)probabiliteit van de S-R associatie. In de 'deterministische' S-R associatie werd elke stimulus van het paar exclusief geassocieerd met een van de twee responstoetsen. M.a.w. respons 1 was steeds correct voor stimulus A, terwijl respons 2 steeds correct was voor stimulus B. In de probabilistische conditie werd stimulus C 80% van de tijd geassocieerd met respons 1 (en dus 20% van de tijd met respons 2), terwijl stimulus D de omgekeerde S-R associatie had. Tenslotte werd in de random conditie elke stimulus van het paar in gelijke mate geassocieerde met beide responstoetsen. M.a.w. stimuli E en F werden 50% van de tijd geassocieerd met respons 1 en 50% van de tijd met respons 2.

De ERN en FRN componenten zijn traditioneel gerelateerd aan foutverwerking gebaseerd op respectievelijk interne (motorische) en externe cues. De ERN/Ne is een negatieve deflectie over fronto-centrale elektroden die een piek vertonen ~0-100 ms na de incorrecte respons. Deze is dus gebaseerd op een interne (motorische) representatie (Carbonnell & Falkenstein, 2006; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Gehring, Goss, Coles, Meyer, & Donchin, 1993). De FRN daarentegen is zijn feedback-gerelateerde tegenhanger, die ongeveer 250-400 ms na de presentatie van negatieve feedback zichtbaar is (Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997).

2. Beknopt overzicht van de belangrijkste bevindingen

2.1. Verschillende manieren om blij te zijn: differentiële effecten van guided imagery en videoclips op de mate van blijheid en reinforcement learning

In hoofdstuk 1 voerden we een reeks studies uit om verschillende technieken gericht op de inductie van een gemoedstoestand te vergelijken, niet alleen wat betreft hun mogelijkheid om positieve veranderingen in de affectieve toestand van deelnemers te verwezenlijken, maar ook om na te gaan of RL systematisch beïnvloed werd door de gemoedstoestand. Hiervoor gebruikten we de probabilistische leertaak (zie Eppinger et al., 2008).

We baseerden ons hierbij op een eerder ontwikkelde theorie van guided imagery (GI) van Emily Holmes (Holmes, Coughtrey, & Connor, 2008; Holmes, Mathews, Dalgleish, & Mackintosh, 2006; Holmes & Mathews, 2010), die gekenmerkt wordt door het gebruik van veldperspectief, d.w.z. vanuit een zeer specifiek, persoonlijk perspectief, en vergeleken dit met de meer courante techniek gebruik makend van grappige videoclips. De resultaten van hoofdstuk 1 toonden aan dat GI even efficiënt is in het induceren van een positieve gemoedstoestand als het bekijken van grappige videoclips. GI is bovendien meer gepersonaliseerd en aangepast aan de achtergrond en specificiteiten van elke individuele deelnemer, waardoor het minder automatisch is en meer vergt van hogere cognitieve functies, zoals vrijwillige inspanning, sociale cognitie/empathie en episodisch geheugen. In dit hoofdstuk vonden we echter geen leerverschillen tussen de positieve en neutrale conditie.

2.2. Effecten van een positieve gemoedstoestand op probabilistisch leren: gedragsmatige en elektrofysiologische correlaten

In hoofdstuk 2 hebben we mechanismen van RL vergeleken op verschillende niveau's bij deelnemers met een positieve of neutrale gemoedstoestand. Hiervoor maakten we gebruik van het computationele model van Sutton en Barto om de leercurve en exploratieparameters af te leiden, bovenop de standaardmaten van leren (accuraatheid en snelheid). De leercurve parameter werd apart berekend voor de positieve en negatieve feedback. We wilden namelijk vooral nagaan of het potentiële effect van een positief gemoed op RL valentie-specifiek was, zodat een positief gemoed enkel de relatie met beloning zou veranderen/verhogen (zoals in Frank, 2005). Bovendien wilden we ook nagaan of een positief gemoed, in vergelijking met een controleconditie van deelnemers met neutrale gemoedstoestand (na een actieve inductiefase), de balans tussen exploratie en exploitatie kon beïnvloeden. Als een positief gemoed gezien kan worden als een tonische (dopamine-gerelateerde) verandering in het systeem (Aston-Jones & Cohen, 2005; Beeler, 2012; Berridge & Kringelbach, 2013), dan zouden we verwachten dat deze specifieke gemoedstoestand geassocieerd is met meer exploratie dan de neutrale gemoedstoestand. We hebben ook de ERN en FRN componenten gemeten in de twee groepen om te kijken of een positief gemoed zou leiden tot een amplitudeverandering van deze componenten. De resultaten toonden aan dat een positieve gemoedstoestand leidde tot een toename van zowel positieve als negatieve leercurves, wat aangeeft dat de deelnemers in de positieve gemoedsgroep gevoeliger waren voor veranderingen in de bekrachtigingswaarde in vergelijking met de neutrale gemoedsgroep, zelfs indien de taakregels geïnternaliseerd waren. Ten tweede hadden de deelnemers in de positieve gemoedsgroep een grotere error-gerelateerde negativiteit (ERN) dan deelnemers in

de neutrale gemoedsgroep, hoewel dit effect beperkt was tot de determinische conditie. Hieruit blijkt dat zij meer responsief/gevoeliger waren voor de interne foutmarkers in het geval dat de respons error duidelijk herkend werd. Dit neurofysiologische resultaat was belangrijk omdat het een grotere ERN amplitude vertoonde geassocieerd met een positief gemoed. Tot dusver waren zulke toegenomen reacties op interne indicators van fouten meestal geassocieerd met internaliserende stoornissen of zelfs karaktertrekken/disposities (subklinisch), zoals depressie en angst (Aarts & Pourtois, 2010; Moser, Moran, Schroder, Donnellan, & Yeung, 2013; Olvet & Hajcak, 2008; Vaidyanathan, Nelson, & Patrick, 2012).

2.3. What is in the feedback? Effecten van geïnduceerde blije vs. verdrietige gevoelens op probabilistisch leren met vs. zonder exploratie

In de volgende studie (hoofdstuk 3) trachtten we te testen of een positieve gemoedstoestand de exploitatie van verworven waarden of kennis tijdens RL zou kunnen bevorderen en zo het leren te doen toenemen/veranderen door de afhankelijkheid van externe informatie (feedback) te verminderen. Daarom verwijderden we de evaluatieve (visuele) feedback na elke beslissing van zodra deelnemers associaties hadden geleerd. Hierdoor werd exploratie geblokkeerd of geneutralizeerd en exploitatie bevorderd. Bovendien voegden we een derde groep toe waarbij een verdrietige gemoedstoestand werd geïnduceerd met behulp van dezelfde ‘guided imagery’ techniek. In de eerste plaats konden we in deze studie de verschillende (grotere) leercurve voor de blije versus de neutrale gemoedsgroep niet repliceren. Een systematische vergelijking van de resultaten van hoofdstuk 2 en 3 toonden echter aan dat het niveau van arousal in de blije groep (na de gemoedsinductieprocedure, GIP) significant lager was in hoofdstuk 3 dan in hoofdstuk 2. Bovendien, hoewel arousal groter was voor blije in vergelijking met neutral

deelnemers in hoofdstuk 2, was dit niet meer het geval in hoofdstuk 3. Er waren geen significante verschillen in arousal tussen de blije, neutrale en verdrietige groep na de GIP. Ten tweede vonden we dat het weglaten van de evaluatieve feedback van zodra leren had plaatsgevonden geen effect had op leerprestatie in de verdrietige en ook niet in de blije groep. Merkwaardig genoeg was nam leren nog steeds toe na het weglaten van de evaluatieve feedback informatie in alle drie de groepen (en dus onafhankelijk van de specifieke geïnduceerde gemoedstoestand). In deze fase van het experiment werd dus exploitatie gebruikt (of interne kennis) om RL te leiden.

2.4. Behoud van 'reinforcement learning' mechanismen bij depressive: evidentie van gedragsmatige, computationeel gemodelleerde en elektrofysiologische data

Tenslotte bestudeerden we in de laatste studie van deze thesis een grote cohort van behandelings-resistente unipolaire depressie (MDD), gekenmerkt door laag positief affect (en Anhedonie) en hoog/excessief negatief affect, die deelnamen aan een behandelingsstudie. We testten de voorspelling dat MDD gekenmerkt wordt door abnormale ERN componenten, in lijn met eerdere modellen en bevindingen in de literatuur (Olvet, Klein, & Hajcak, 2010). De resultaten van hoofdstuk 4 toonden aan dat de ERN en FRN componenten van een grote groep zeer zware (behandelingsresistente) MDD patiënten in vergelijking met een groep van gezonde controles gematched op leeftijd niet abnormaal waren tijdens RL. Dezelfde conclusie werd getrokken met betrekking tot de leercurve en exploratie. Deze resultaten zijn belangrijk omdat ze ingaan tegen de assumptie van een generisch tekort of abnormaliteit tijdens het verwerken van beloning bij MDD (Pizzagalli, 2014)). Om deze intacte vaardigheid bij MDD te verklaren, veronderstellen we dat de specifieke behandelings- en taakcontext (Eppinger et al., 2008) paradoxaal het gebruik van intacte (intrinsieke) motivationele processen heeft

bevorderd. Dit liet toe om te compenseren voor een normalerwijs abnormale verwerking van beloning, en op deze manier konden ze een gelijkaarde prestatie neerzetten als gezonde controles. We erkennen uiteraard dat deze interpretatie verdere empirische validatie vereist vooraleer de rol van specifieke motivationele processen in MDD tijdens RL bevestigd kunnen worden.

3. Discussie

Wanneer we de 4 experimentele hoofdstukken samennemen, vinden we weinig of beperkte evidentie voor een sterke en systematische modulatie van RL mechanismen door (een positief) gemoed, alsof de geïnduceerde gemoedseffecten RL niet direct beïnvloedden. Ook al was de GIP elke keer succesvol (of een positieve gemoedstoestand was sterk veranderd van bij het begin zonder dat dit geïnduceerde diende te worden, zoals in hoofdstuk 4) en ook al resulteerde de probabilistische leertaak telkens in de verwachtte RL effecten (op gedragsmatig en ERP niveau), toch moeten we erkennen dat op basis van de 4 hoofdstukken samen evidentie voor een sterk effect van de gemoedstoestand op dit proces niet kon aangetoond worden. Op basis van de experimentele resultaten zoals hierboven besproken, werden nieuwe inzichten inzake effecten van een positief gemoed op RL verkregen, en hieruit kunnen specifieke conclusies getrokken worden. Zo blijken de veranderingen veroorzaakt door de GIP slechts in beperkte mate een impact te hebben op RL. Hoewel de onderliggende verklaring voor deze observatie waarschijnlijk meervoudig en complex is, willen we toch enkele mogelijke verklaringen naar voren brengen. Zoals onze resultaten suggereren (zie hoofdstukken 2 en 3), zou arousal (evenals valentie) een belangrijke factor kunnen zijn in de emotionele effecten van de GIP. Daarom lijkt het ons essentieel om een sterke GIP te ontwikkelen die rekening houdt met deze factor, in plaats van ervan uit te gaan (zoals wij deden) dat dit geen rol speelde. Daarbovenop moeten toekomstige studies zich meer

rechtstreeks richten op het effect van motivationele factoren. Zoals aangetoond in hoofdstuk 4 kunnen deze invloeden namelijk het verband tussen gemoedstoestand en RL verzwakken. Tenslotte zou het methodologisch gezien waardevol zijn om in toekomstige studies veranderingen in gemoedstoestand beter te relateren aan de RL taak. Een eerste vereiste is waarschijnlijk een betere characterization van gemoedsgerelateerde EEG toestanden, vooraleer dit kan gebruikt worden in combinatie met andere taken (zoals de probabilistische leertaak die gebruikt werd in dit werk).

4. References

Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective & Behavioral Neuroscience*, *10*(4), 479–92. doi:10.3758/CABN.10.4.479

Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529–550. Retrieved from <Go to ISI>://000082133000004

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–50. doi:10.1146/annurev.neuro.28.061604.135709

Beeler, J. a. (2012). Thorndike's law 2.0: Dopamine and the regulation of thrift. *Frontiers in Neuroscience*, *6*(AUG), 1–12. doi:10.3389/fnins.2012.00116

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Berridge, K. C., & Kringelbach, M. L. (2013). Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology*, *23*(3), 294–303. doi:10.1016/j.conb.2013.01.017

Bolte, A., & Goschke, T. (2010). Towards a Theory of Thinking, 261–277. doi:10.1007/978-3-642-03129-8

Carbonnell, L., & Falkenstein. (2006). Does the error negativity reflect the degree of response conflict? *Brain Research*, *1095*(1), 124–130. doi:10.1016/j.brainres.2006.04.004

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical*

Transactions of the Royal Society of London. Series B, Biological Sciences, *362*(1481), 933–42. doi:10.1098/rstb.2007.2098

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

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*(2), 118–24. doi:10.1016/j.biopsych.2010.01.027

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

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, *51*(2-3), 87–107. doi:10.1016/S0301-0511(99)00031-9

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

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci*, *359*(1449), 1367–1378. doi:10.1098/rstb.2004.1512

Gehring, Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error-Detection and Compensation. *Psychological Science*, *4*(6), 385–390. Retrieved from <Go to ISI>://A1993MF61300008

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

Goschke, T., & Bolte, A. (2013). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*.

Holmes, E. A., Coughtrey, A. E., & Connor, A. (2008). Looking at or through rose-tinted glasses? Imagery perspective and positive mood. *Emotion (Washington, D.C.)*, *8*(6), 875–879. doi:10.1037/a0013617

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). *Positive interpretation training: effects of mental imagery versus verbal training on positive mood. Behavior therapy* (Vol. 37). doi:10.1016/j.beth.2006.02.002

Holmes, & Mathews. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349–362. doi:10.1016/j.cpr.2010.01.001

Holroyd, C., & Coles, M. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–96. doi:10.1162/jocn.2010.21548

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

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*(August), 466. doi:10.3389/fnhum.2013.00466

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: toward an endophenotype. *Clinical Psychology Review*, *28*(8), 1343–54. doi:10.1016/j.cpr.2008.07.003

Olvet, D. M., Klein, D. N., & Hajcak, G. (2010). Depression symptom severity and error-related brain activity. *Psychiatry Research*, *179*(1), 30–37.

Pizzagalli. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, (January), 1–31. doi:10.1146/annurev-clinpsy-050212-185606

Ridderinkhof, K. R., van Wouwe, N. C., Band, G. P. H., Wylie, S. a., Van der Stigchel, S., Van Hees, P., … Van den Wildenberg, W. P. M. (2012). A tribute to Charlie Chaplin: induced positive affect improves reward-based decision-learning in Parkinson's disease. *Frontiers in Psychology*, *3*(JUN), 1–10. doi:10.3389/fpsyg.2012.00185

Vaidyanathan, U., Nelson, L. D., & Patrick, C. J. (2012). Clarifying domains of internalizing psychopathology using neurophysiology. *Psychological Medicine*. doi:10.1017/S0033291711001528

Data Storage Fact Sheet (22 June 2015)
Name/identifier study: "Ways of being happy: differential effects of guided imagery and movie clips on happiness levels and reinforcement learning" (Chapter 2)
Author: Jasmina Bakic
Date: 22 June 2015

1. Contact details
==========
1a. Main researcher

- name: Jasmina Bakic
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Jasmina.Bakic@UGent.be
1b. Responsible Staff Member (ZAP)

- name: Gilles RC Pourtois, PhD.
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: gilles.pourtois@ugent.be

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:
/
* Which datasets in that publication does this sheet apply to?: the sheet applies to all the data used in the chapter

3. Information about the files that have been stored
==========
3a. Raw data
---
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:
* On which platform are the raw data stored?
- [X] researcher PC
- [ ] research group file server:
- [X] other (specify): external memory hard drive personal back-up
* Who has direct access to the raw data (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): co-researcher

3b. Other files
---
* Which other files have been stored?
- [X] file(s) describing the transition from raw data to reported results. Specify: Please refer to methodology section in the article
- [X] file(s) containing processed data. Specify: R script (.R) and exel (.xlsx) output files with in-between steps of the analysis for the behavioral data
- [X] file(s) containing analyses. Specify: Output files of spss analysis (.spv)
- [X] files(s) containing information about informed consent (.docx)
- [ ] a file specifying legal and ethical provisions.
- [ ] file(s) that describe the content of the stored files and how this content sho

uld be interpreted.
- [ ] other files. Specify: ...
* On which platform are these other files stored?
- [X] individual PC
- [ ] research group file server
- [X] other: external hard drive of responsible ZAP
* Who has direct access to these other files (i.e., without intervention of another pe
rson)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...
4. Reproduction
==========
* Have the results been reproduced independently?: [] YES / [X] NO
* If yes, by whom (add if multiple):
- name:
- address:
- affiliation:
- e-mail:

Data Storage Fact Sheet (24 September 2014)
Name/identifier study: "Effects of positive mood on probabilistic learning: behavioral and electrophysiological correlates", (Chapter 3)
Author: Jasmina Bakic
Date: 24 september 2014

1. Contact details
==========
1a. Main researcher

- name: Jasmina Bakic
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Jasmina.Bakic@UGent.be
1b. Responsible Staff Member (ZAP)

- name: Gilles RC Pourtois, PhD.
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: gilles.pourtois@ugent.be

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:
Bakic, J., Jepma, M., De Raedt, R., & Pourtois, G. (2014). Effects of positive mood on probabilistic learning: Behavioral and electrophysiological correlates.
Biological Psychology, 103, 223–232. doi:10.1016/j.biopsycho.2014.09.012
* Which datasets in that publication does this sheet apply to?: the sheet applies to a ll the data used in the publication

3. Information about the files that have been stored
==========
3a. Raw data
---
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:
* On which platform are the raw data stored?
- [X] researcher PC
- [ ] research group file server:
- [X] other (specify): external memory hard drive personal back-up
* Who has direct access to the raw data (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): co-researcher

3b. Other files
---
* Which other files have been stored?
- [X] file(s) describing the transition from raw data to reported results. Specify: Please refer to methodology section in the article
- [X] file(s) containing processed data. Specify: R (.R) script and exel output files with in-between steps of the analysis for the behavioral data,
as well as Brain Vision Analyzer files with analysis steps (.cnt, .ehst2, .hfinf) .
- [X] file(s) containing analyses. Specify: Output files of spss analysis (.spv), and the Brain Vision Analyser export

files (.txt) with electrophysiological data
Exceptionally, analysis script for the RL modelling of the behavioral data is stored externally, by a study colaborator, Marieke Jepma,
who holds a post-doctoral fellowship at the Colorado University at Boulder. She can be reached by email :Marieke.Jepma@Colorado.EDU
- [X] files(s) containing information about informed consent (.txt).
- [ ] a file specifying legal and ethical provisions.
- [ ] file(s) that describe the content of the stored files and how this content should be interpreted.
- [ ] other files. Specify: ...
* On which platform are these other files stored?
- [X] individual PC
- [ ] research group file server
- [X] other: external hard drive of responsible ZAP
* Who has direct access to these other files (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...
4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [x] NO
* If yes, by whom (add if multiple):
- name:
- address:
- affiliation:
- e-mail:

Data Storage Fact Sheet (22 June 2015)
Name/identifier study: "What is in the feedback? Effects of induced happiness vs. sadness on probabilistic learning with vs. without exploration", (Chapter 4)
Author: Jasmina Bakic
Date: 22 June 2015

1. Contact details
==========
1a. Main researcher

- name: Jasmina Bakic
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Jasmina.Bakic@UGent.be
1b. Responsible Staff Member (ZAP)

- name: Gilles RC Pourtois, PhD.
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: gilles.pourtois@ugent.be

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:
Bakic, J., De Raedt, R., Jepma, M., & Pourtois, G. (2014). What is in the feedback? Effect of induced happiness vs. sadness on probabilistic learning with vs. without exploration. Manuscript in preparation
* Which datasets in that publication does this sheet apply to?: the sheet applies to a ll the data used in the publication

3. Information about the files that have been stored
==========
3a. Raw data
---
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:
* On which platform are the raw data stored?
- [X] researcher PC
- [ ] research group file server:
- [X] other (specify): external memory hard drive personal back-up
* Who has direct access to the raw data (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): co-researcher

3b. Other files
---
* Which other files have been stored?
- [X] file(s) describing the transition from raw data to reported results. Specify: Please refer to methodology section in the article
- [X] file(s) containing processed data. Specify: R script (.R) and exel output files (.xlsx) with in-between steps of the analysis for the behavioral data,
- [X] file(s) containing analyses. Specify: Output files of spss analysis (.spv), and the Brain Vision Analyser export files with electrophysiological data
- [X] files(s) containing information about informed consent (.txt)

- [X] a file specifying legal and ethical provisions (.pdf)
- [ ] file(s) that describe the content of the stored files and how this content sho
uld be interpreted.
- [ ] other files. Specify: ...
* On which platform are these other files stored?
- [X] individual PC
- [ ] research group file server
- [X] other: external hard drive of responsible ZAP
* Who has direct access to these other files (i.e., without intervention of another pe
rson)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...
4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [x] NO
* If yes, by whom (add if multiple):
- name:
- address:
- affiliation:
- e-mail:

Data Storage Fact Sheet (22 June 2015)
Name/identifier study: "Preserved reinforcement learning mechanisms in major depressive disorder: evidence from behavioral,
computational modeling and electrophysiological data" (Chapter 5)
Author: Jasmina Bakic
Date: 22 June 2015

1. Contact details
==========
1a. Main researcher

- name: Jasmina Bakic
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Jasmina.Bakic@UGent.be
1b. Responsible Staff Member (ZAP)

- name: Gilles RC Pourtois, PhD.
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: gilles.pourtois@ugent.be

2. Information about the datasets to which this sheet applies
==========
* Reference of the publication in which the datasets are reported:
Bakic, J., Pourtois, G., Jepma, M., Duprat, R., De Raedt, R., & Baeken, C. (2014) Preserved reinforcement learning mechanisms in major depressive disorder:
evidence from behavioral, computational modeling and electrophysiological data. Manuscript in preparation
* Which datasets in that publication does this sheet apply to?: the sheet applies to a ll the data used in the publication

3. Information about the files that have been stored
==========
3a. Raw data
---
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO
If NO, please justify:
* On which platform are the raw data stored?
- [X] researcher PC
- [ ] research group file server:
- [X] other (specify): external memory hard drive personal back-up
* Who has direct access to the raw data (i.e., without intervention of another person)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): co-researcher

3b. Other files
---
* Which other files have been stored?
- [X] file(s) describing the transition from raw data to reported results. Specify: Please refer to methodology section in the article
- [X] file(s) containing processed data. Specify: R script (.R) and exel output files (.xlsx) with in-between steps of the analysis for the behavioral data,
as well as Brain Vision Analyzer files (.cnt, .ehst2, .hfinf) with analysis steps.

- [X] file(s) containing analyses. Specify: Output files of spss analysis (.spv), and the Brain Vision Analyser export files (.txt) with electrophysiological data
- [X] files(s) containing information about informed consent. Specify: files are stored in the hard copy in the main researcher's office
- [ ] a file specifying legal and ethical provisions.
- [ ] file(s) that describe the content of the stored files and how this content sho
uld be interpreted.
- [ ] other files. Specify: ...
* On which platform are these other files stored?
- [X] individual PC
- [ ] research group file server
- [X] other: external hard drive of responsible ZAP
* Who has direct access to these other files (i.e., without intervention of another pe
rson)?
- [X] main researcher
- [X] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...
4. Reproduction
==========
* Have the results been reproduced independently?: [ ] YES / [x] NO
* If yes, by whom (add if multiple):
- name:
- address:
- affiliation:
- e-mail: