ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

DOTTORATO DI RICERCA

INTERNATIONAL PHD PROGRAM IN COGNITIVE NEUROSCIENCE

Ciclo XXVIII

Settore Concorsuale di afferenza: AREA 11/E1

Settore Scientifico disciplinare: M-PSI/02

ADAPTIVE AND MALADAPTIVE IMPLICATIONS OF REINFORCEMENT LEARNING PROCESSES: FRONTO-STRIATAL LOOPS AND BEHAVIOURAL CORRELATES

Dottoranda	Sara Garofalo
Coordinatore Dottorato	Prof.ssa Monica Rubini
Relatore	Prof. Giuseppe di Pellegrino

Esame finale anno 2016

"We are like the prisoners in the cave. There are platonic 'biological bases of behaviour' that we want to discover (the figures walking on the pathway behind the prisoners), but all we can observe are the shadows cast on the wall (empirical data) by the flame in the back of the cave (methods and technologies); our concept of the nature and shape of the figures (theories) are shaped by past experience, intuition, and, perhaps most importantly, how the light of the flame defines the shadows.

However, we have one important advantage over the prisoners in Plato's cave: We can, to some extent, control the flame. We can develop new technologies and methodologies, and we can combine methodologies in interesting, novel, and insightful ways. We can compare the shadows cast on the wall using different materials to fuel the fire."

Michael X Cohen

INDEX

INTRODUCTION

That humans and animals learn from interaction with the environment is a foundational idea underlying nearly all theories of learning and intelligence (Sutton and Barto, 1998). Learning that certain outcomes are associated with specific actions or stimuli (both internal and external), is at the very core of the capacity to adapt behaviour to environmental changes. It is such a basic ability to be observable even in mollusks like the Aplysia (Nargeot and Simmers, 2011), a sea slug with only about 20,000 neurons, but at the same time it carries a critical survival value, as it allows flexible adaptation of ongoing behaviour to volatile contingencies and unexpected outcomes (Kehagia et al., 2010; Cohen et al., 2011).

Reinforcement learning is a form of learning in which choices are guided by the presence of rewards and punishments (the *reinforcers*) in the environment. In this perspective, humans and animals are conceived as goal-directed agents who learn, by trial and error, how to maximize current and future rewards (or minimize punishments) while interacting with an uncertain environment (Sutton and Barto, 1998).

The two most basic forms of associative learning are widely known as Pavlovian Conditioning and Instrumental Conditioning.

In Pavlovian Conditioning, learning occurs when rewards or punishments (unconditioned stimuli - US) are paired with otherwise irrelevant environmental cues, which gain salience thanks to this association (thus becoming conditioned stimuli, CS). These CS-US pairings occur regardless of any action performed, which means that the probability of an outcome (US) given the presence

of a specific cue (CS) is independent of the individual performance. Once a CS is paired with a US, the CS acquires a motivational salience that can elicit a so-called conditioned response (CR), a psychophysiological reaction to the predicted outcome that reflects the existence and the strength of the associative learning process.

In Instrumental Conditioning, on the other hand, learning is represented by the association between an action and an outcome (e.g., a button press and the consequent occurrence of a reward). In this case, individual performance does influence the probability of getting rewards and punishments. For this reason, instrumental actions are consequently shaped by the occurrence of rewards and punishments, as they can be actively sought or avoided.

Understanding how associations between (past and present) events, actions and their related outcomes are build, is the principal goal of research in learning theory (Behrens et al., 2007).

From a computational point of view, learning has been conceptualized as a prefrontally mediated modulation of the synaptic weights connecting population of neurons coding for stimuli or contexts, goal-directed actions and the outcomes connected to them (Cohen et al., 2011; Hebb, 1949).

Studies on the neural mechanisms underlying learning and flexibility pointed to a major role played by fronto-striatal loops and modulated by monoamines, such as dopamine (DA), serotonin (5-hydroxytryp-tamine, 5-HT) and noradrenaline (NA) (Kehagia et al., 2010).

Such fronto-striatal loops are involved in behaviour monitoring by learning rules, forming expectations about outcome delivery and coding for prediction error (i.e., when the actual

outcome differs from the predicted outcome) (O'Doherty et al., 2003; Schultz et al., 1997; Histed et al., 2009; Miller, Nieder, Freedman, and Wallis, 2003; Miller and Cohen, 2001; Wallis, Anderson, and Miller, 2001; White and Wise, 1999).

Neurochemically, strong links have been reported in literature with dopaminergic activity (Kehagia et al., 2010). Many studies showed a direct relationship between striatal D2 and D1 receptor activity and frontal functions and dysfunctions (Pycock et al., 1980; Bach et al., 2008; Kellendonk et al., 2006; Castner et al., 2001).

The importance of further investigation into the behavioural and neural basis of reinforcement learning resides in its strong links with both adaptive and maladaptive human processes. Reinforcement learning is indeed believed to be at the core of human social and emotional abilities (Behrens et al, 2008; Brown and Brüne, 2012), as well as implicated in different forms of addiction - like drug or alcohol dependence (Baker, et al. 2011; Crean et al. 2011; Park. et al., 2010) - and psychiatric diseases - including schizophrenia, obsessive compulsive disorder (OCD), post-traumatic stress disorder and phobia (Everitt and Robbins, 2005; Gold, et al. 2008; Fletcher and Frith, 2009; Berridge et al., 1998).

In the present work, appetitive and aversive reinforcement learning paradigms have been used to investigate the fronto-striatal loops and behavioural correlates of adaptive and maladaptive reinforcement learning processes, aiming to a deeper understanding of how cortical and subcortical substrates interacts between them and with other brain systems to support learning.

A first study, carried out at the Centre for studies and research in Cognitive Neuroscience (CNC) of the University of Bologna, investigated brain responses to the omission aversive of aversive outcomes in healthy volunteers. This experiment investigated the role of a specific mediofrontal negative event-related potential (ERP), the mediofrontal negativity, a component implicated in prediction-error signalling. Research based on ERPs reported mediofrontal negativities following unexpected negative feedback or performance error. Some authors proposed that these signals reflect reward prediction error for worse than expected outcomes, while others suggested that mediofrontal negativities express medial prefrontal cortex (mPFC) coding for unexpected non-occurrence of a predicted outcome, whether worse or better than expected. Many studies found mediofrontal negativities coding for unexpected negative outcomes; however, few studies found them after unexpected positive outcomes. This research investigated ERP and skin conductance response (SCR) to the unexpected omission of electric shocks during Pavlovian aversive conditioning. To manipulate expectancies, participants were presented with visual stimuli paired with electric shocks on either 80% (CS+1) or 20% (CS+2) of trials. SCR analysis confirmed higher shock-delivery expectancy for CS+1, relative to CS+2. ERP analysis evidenced a stronger negative frontocentral ERP component after unexpected, relative to expected, shock-omission. The results of this experiment have been successfully published in Garofalo et al. (2014).

A second study, carried out at the CNC (University of Bologna), aimed at the investigation of ERP signals of Timing Prediction Error (TPE) during aversive reinforcement learning in healthy volunteers. The medial prefrontal cortex (mPFC) and the anterior cingulate cortex (ACC) have

been consistently implicated in learning predictions of future outcomes, and signaling prediction errors (i.e., unexpected deviations from such predictions). A recent computational model of the mPFC posits that these prediction errors should be modulated by outcomes occurring at unexpected times, even if the outcomes themselves are predicted. However, unexpectedness *per se* is not the only variable that modulates mPFC activity, as studies reported its sensitivity to the salience of outcomes. In the present study, mediofrontal negativity, a component of the event-related brain potential generated in the mPFC/ACC and coding for prediction errors, was measured in 48 participants performing a Pavlovian aversive conditioning task, during which aversive (thus, salient) and neutral outcomes were unexpectedly anticipated or delayed in time. Mediofrontal ERP signals of prediction error were observed for outcomes occurring at unexpected times, but were specific for salient (shock-associated), as compared to neutral, outcomes. These findings have important implications for theoretical accounts of the mPFC/ACC and suggest a critical role of timing and salience information in prediction error signaling. The results of this experiment are in preparation for submission to international scientific journals.

A third study, in collaboration with the Department of Engineering of the University of Bologna, aimed at investigating the cortical generators of the ERP component found to code for salient timing prediction error in the second experiment here described. Cortical generators were estimated using a source localization method known as sLORETA. Results showed that some regions in the frontal and prefrontal cortex may be responsive to unexpected timing of feedback,

to a larger extent in case of aversive than neutral feedback. The present study can contribute to inform current theories of performance monitoring in frontal/prefrontal cortex and to drive future investigation and results are discussed with respect to findings of previous works. These results have been successfully published in the Conference Proceedings of the 2015 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Magosso et al., 2015).

A fourth study, carried out at the Department of Psychiatry of the University of Cambridge, used fMRI and computational modelling techniques to investigate the presence of deficits in fronto-striatal signals of reward processing in Parkinson's Disease associated psychosis. Psychotic symptoms frequently occur in Parkinson's Disease, but their pathophysiology is poorly understood. According to cognitive neuropsychiatric theory, and the National Institute of Health RDoc programme, the pathophysiological basis of psychiatric symptoms may be better understood in terms of dysfunction of underlying domains of neurocognition, rather than simply according to diagnosis. Abnormal frontal striatal reinforcement learning signalling, associated with dysregulated dopamine, has been proposed as a key process contributing to the pathogenesis of psychosis. This theory has received empirical support in the study of schizophrenia spectrum disorders and preclinical models of psychosis, but has not been tested in psychosis associated with a degenerative neurological condition such as Parkinson's Disease. This study aimed to investigate brain responses associated with reward feedback and prediction error signalling during reinforcement learning in Parkinson's Disease associated psychosis. An instrumental

reinforcement learning task was performed by three groups of subjects during fMRI scanning: patients with Parkinson's Disease with a history of psychotic symptoms (n=12), patients with Parkinson's Disease without a history of psychotic symptoms (n=17), and healthy controls (n=24). Results confirmed the presence of altered activation in the medial prefrontal cortex and the ventral striatum in PD patients with psychosis in different stages of reward processing. These results extend the link between psychotic symptoms and abnormalities in fronto-striatal reward processing beyond schizophrenia spectrum conditions into psychosis as manifest in Parkinson's Disease. The results of this experiment are in preparation for submission to international scientific journals.

A fifth study, carried out at the CNC (University of Bologna), focused on individual differences in the influence of task-irrelevant Pavlovian cues on human behaviour. Pavlovian-to-instrumental transfer (PIT) refers to the process of a Pavlovian reward-paired cue acquiring incentive motivational proprieties that drive choices. It represents a crucial phenomenon for understanding cue-controlled behaviour, and it has both adaptive and maladaptive implications (i.e., drug-taking). In animals, individual differences in the degree to which such cues bias performance have been identified in two types of individuals that exhibit distinct Conditioned Responses (CR) during Pavlovian conditioning: Sign-Trackers (ST) and Goal-Trackers (GT). Using an appetitive PIT procedure with a monetary reward, the present study investigated, for the first time, the extent to which such individual differences might affect the influence of reward-paired cues in humans. In a first task, participants learned an instrumental response

leading to reward; then, in a second task, a visual Pavlovian cue was associated with the same reward; finally, in a third task, PIT was tested by measuring the preference for the reward-paired instrumental response when the task-irrelevant reward-paired cue was presented, in the absence of the reward itself. In ST individuals, but not in GT individuals, reward-related cues biased behaviour, resulting in an increased likelihood to perform the instrumental response independently paired with the same reward when presented with the task-irrelevant reward-paired cue, even if the reward itself was no longer available (i.e., stronger PIT effect). This finding has important implications for developing individualized treatment for maladaptive behaviours, such as addiction. The results of this experiment have been successfully published in Garofalo et al. (2015).

A sixth study, carried out at the Department of Psychology of the University of Cambridge, aimed at testing the influence of Pavlovian aversive stimuli on avoidance instrumental responses by using a Pavlovian-to-Instrumental Transfer (PIT) task. PIT effect reflects the learned motivational influence of the CS over the instrumental response (Blundell, Hall, and Killcross, 2001; Corbit, Muir, and Balleine, 2001; Dickinson, Smith, and Mirenowicz, 2000; Hall, Parkinson, Connor, Dickinson, and Everitt, 2001; Holland and Gallagher, 2003). Two forms of transfer have been object of investigation: outcome-specific transfer (which is thought to be mediated by the sensory-specific properties of the reinforcer, deriving from a S-O-R association) and general transfer (which is thought to be mediated by a direct S–R association, reflecting a more general motivational response). Additionally, in order to clarify the nature of the transfer

effect (goal-directed/habit), the impact of reinforcer devaluation after different amounts of instrumental training has been investigated. A double dissociation between specific and general aversive transfer effects was reported for the first time in humans. Specific transfer was expressed in the percentage of responses but not in the vigour of such responses; whereas the opposite pattern was observed for general transfer, where the effect was captured by the vigour of the responses but not percentage of responses. Moreover, specific transfer was enhanced by instrumental overtraining, but not by reinforcer devaluation. General transfer was not affected by neither instrumental overtraining nor reinforcer devaluation. The results of this experiment are in preparation for submission to international scientific journals.

PART I - FRONTO-STRIATAL LOOPS INVOLVED IN REINFORCEMENT LEARNING

Adapting behaviour to a changing environment

Prediction Error signals

In its most basic form, learning is the process by which we become able to use past and current experience to predict future events (Niv and Schoenbaum, 2008). According to associative learning theories (Sutton and Barto, 1981; Rescorla and Wagner, 1972), learning occurs when the actual outcome differs from the predicted outcome, resulting in a prediction error. Therefore, prediction error represents the extent to which an outcome occurs surprisingly or unpredictably and inform about changes in the environment (Schultz and Dickinson, 2000; Sutton and Barto, 1998). Prediction error signals are thought to drive learning as are used to update expectations in order to make predictions more accurate (Gläscher et al., 2011).

People form representations of events based on experience and, consequently, create predictions about what is going to happen in specific situations. The actual event that ultimately occurs is then compared with the expectancy and, if different, a prediction error signal is generated and used to update knowledge.

Neural underpinnings of Prediction Error

This ability to acquire associations useful to adapt ongoing behaviour, carries an important adaptive value that has been mainly attributed to the prefrontal cortex (Kehagia et al., 2010).

The medial prefrontal cortex (mPFC) and the anterior cingulate cortex (ACC) have been reported as a key areas in acquiring and updating outcome expectation (Rushworth and Behrens, 2008), being involved in the evaluation of both external cues and actions (Matsumoto et al., 2003; Amiez et al., 2005; Amiez et al., 2006; Matsumoto et al., 2007; Seo and Lee, 2007).

Prediction error signals have been reported as expressed in midbrain and subsequently conveyed to mPFC and ACC areas (Schultz and Dickinson, 2000; Schoenbaum et al., 2010; Holyd et al., 2004). An hypothesis proposed by Holroyd and Coles (2002) suggests that the mesencephalic dopamine system conveys reinforcement learning signals to the frontal cortex in corresponsence of prediction errors, which generates mediofrontal negativities by disinhibiting the apical dendrites of motor neurons in the anterior cingulate cortex, ensuring control over future actions.

Electrophysiological correlates of prediction error signals, deriving from the dopaminergic projections to mPFC/ACC, can be identified in mediofrontal negative Event-Related Potentials (ERP) observed after performance error (Error Related Negativity - ERN) or following unexpected feedbacsa (Feedback Related Negativity - FRN) (Sambrook and Goslin, 2016; Hauser et al., 2014; Holroyd and Coles., 2002; Gehring et al., 1993; Falkenstein et al., 1990; for reviews see Nieuwenhuis et al., 2004 and Walsh and Anderson, 2012).

Models

Currently, there is still an open debate about if learning via prediction error is primarily driven by an event's valence or simply its unexpectedness (Ferdinand et al., 2012).

The Reinforcement Learning account of mediofrontal negativity (RL-ERN) states that mPFC is involved in the elaboration of worse than expected outcome (Holroyd and Coles, 2002; Hajcak et al 2006), supporting the idea of mPFC coding for valence. This signal is thought to be triggered by either internal information (e.g., when a performance error occurs instead of the expected correct response) or external information (e.g., when a choice is followed by negative feedback). Following this line, many studies reported mPFC activation for aversive representation of prediction error (Seymour et al., 2005) responding to negative affect and pain (Shackman et al., 2011) and negative outcomes (Hajcak et al 2006; Gehring and Willoughby, 2002; Holroyd et al., 2004).

A more recent hypothesis comes from the Predicted Response-Outcome (PRO) model, which proposed a comprehensive model of the role of mPFC and the adjacent anterior cingulate cortex (ACC) in learning (Alexander and Brown, 2011; 2014) and their interaction with the dorsolateral prefontal cortex (Alexander and Brown, 2015).

The authors proposed that mPFC codes the unexpected non-occurrence of a predicted outcome, regardless of its affective valence. When an expected outcome is surprisingly omitted (i.e., an expected reward or punishment are not delivered) an unexpected non-occurrence signal is reflected by mPFC activation, regardless of whether outcome is worse or better than expected.

This signal should be maximal when an expected outcome fails to occur, while it is inhibited when the predicted outcome actually occurs (Alexander and Brown, 2011).

The PRO-model holds that mediofrontal negativities, reflect mPFC activation for the unexpected non-occurrence of a predicted outcome. Consequently, both positive and negative unexpected non-occurrence events should be able to generate mediofrontal negative component. It has to be noticed that the omission of a positive outcome has a negative valence (e.g., frustration if you don't get a monetary win), while the omission of a negative outcome has a positive valence (e.g., relief if you don't get an electric shock) (Rolls, 2000). In line with both hypotheses, many findings reported frontal negative ERP signals for unexpected negative outcomes (Crowley et al., 2009; Holroyd et al., 2003; Holroyd and Krigolson, 2007; Chase et al., 2007); but few studies also reported similar mediofrontal negativities to unexpected positive outcomes (Talmi et al., 2013; Oliveira et al., 2007; Ferdinand et al., 2012), a finding that is coherent with the PRO-model but not with the Reinforcement Learning account. Even if these experiments did not aim to directly test the unexpected non-occurrence hypothesis postulated by the PRO-model, their findings are coherent with it as mediofrontal negative ERPs were reported after the presentation of a different stimulus than the one expected (i.e., a stimulus with the opposite valence or a neutral stimulus), rather than after the omission of the expected stimulus itself. For example, in a study from Talmi and colleagues (Talmi et al., 2013), participants explicitly knew the outcome probability and a "truth cue" informed about the subsequent presence or absence of the outcome. In this paradigm, the mediofrontal negativity was related to "truth cue" onset rather than to the omission of the outcome itself. In a different study, Oliveira and colleagues (Oliveira et al., 2007), reported mediofrontal negativity after a better than expected performance-feedback

during an anticipation-timing task. Again, rather than the unexpected omission of the predicted feedback, participants were presented with a positive feedback when a negative one was expected. In a further study by Ferdinand and colleagues (Ferdinand et al., 2012), mediofrontal negativity was found after a rare, and thus unexpected, positive or negative feedback (relative to an intermediate feedback of high frequency) during a time-estimation task.

Another interesting aspect hypothesized by the PRO-model concerns the prediction of the timing of outcomes. ACC/mPFC activation should be registered for outcomes occurring at unexpected times, even if the outcomes themselves are expected (Alexander and Brown, 2011), therefore signalling a Timing Prediction error (TPE).

The PRO-model suggests that some ACC/mPFC neurons are reliably activated as a consequence of specific events and increase their firing to anticipate them, peaking around the time when the event is most likely to occur. The more likely an event becomes, the stronger becomes also the anticipatory neural firing. These signals reflect a learned prediction of the probability and timing of possible outcomes and are inhibited when the outcome actually occurs. Maximal activity will be consequently registered when an expected outcome fails to occur. Cells that are excited by predictions and inhibited by outcomes will generate Prediction error signals, computed as the difference between a prediction and an outcome. Predictions about the timing of outcome delivery are the result of implicit learning processes in which relevance is given not only to what event will occur or how likely its occurrence is, but also when it will occur. In this context, Prediction error signals are used to update previous expectations, when new contradictory and salient events occur.

This assumption of the PRO-model is largely coherent with evidences obtained in electrophysiological animal studies where dopamine increases corresponding to TPEs have been reported (Hollerman and Schultz, 1998). It is, indeed, now widely accepted that the dopaminergic system plays an important role in signalling Prediction errors (Holroyd and Cole, 2002; Schultz et al., 1997; Schultz, 2007) and electrophysiological studies with animals showed dopaminergic activity coding for unpredicted timing of outcome, when upcoming events are either delayed or anticipated in time (Kobayashi and Schultz, 2008; Schultz and Dickinson, 2000; Hollerman and Schultz, 1998; Fiorillo et al., 2003; 2008).

Evidences obtained by functional magnetic resonance imaging (fMRI) studies identified the ventral tegmental area (VTA) and ventral striatum (VS) - key dopaminergic areas – as coding for reward prediction errors and timing-related performance respectively (Klein-Flügge et al., 2011). Moreover, Forster and Brown (2011) investigated how the delivery of rewards at unexpected timings modulates activity in the ACC. Using a probabilistic learning task, participants were presented with outcomes at variable timings. Outcomes unexpectedly delayed in time correlated with activations in the caudate cingulate zone and dorsomedial prefrontal cortex.

Few previous studies reported inconsistent findings about the relationship between timing of outcome delivery and mediofrontal negativity (Weinberg et al., 2012; Wang et al., 2014; Peterburs et al., 2015), but none of these studies ever considered how expectation and violation of expectation (thus, prediction error) can modulate that.

The idea proposed by the PRO-model of ACC/mPFC general role in predicting outcomes and signalling prediction errors, irrespective of their valence, has been supported by many studies on

mediofrontal negativity (Garofalo et al., 2014; Ferdinand et al., 2012; Jessup et al., 2010; Oliveira et al., 2007; Talmi et al., 2013). However, unexpectedness per se is not the only variable that accounts for ACC/mPFC activation. Mediofrontal negative ERP components have been reported to be sensitive to the salience of outcomes and errors, being modulated by the amount of punishment (Maier et al., 2013; Talmi et al., 2013; Ganushchak et al., 2008; Hajcak et al., 2005), the attentional significance (Maier et al., 2011; 2012) or the perceptual salience (Lou et al., 2015) of the events regardless of their likelihood.

In scientific literature, the concept of salience has been used with many different meanings.

For example, an extension of the PRO-model (Alexander and Brown, 2014) proposes that ACC/mPFC responds to any salient sensory input, rather than being restricted only to action outcomes, as stated in the first formulation of the model. However, what is intended for "salient" is not really clarified and it seems to be used just as a general term to indicate any non-action-related event. In other studies, salience has been conceptualized as higher versus lower rewards and punishments (Talmi et al., 2013) or as perceptual salience (Lou et al., 2015).

Commonly, the term salience is attributed to something that has a motivational significance (e.g., drive to get rewards and/or to avoid punishments), and a behavioural component of such significance can possibly be identified and attributed to ACC/mPFC role in reinforcement learning (e.g., when acquiring relevant information to optimize imminent and future behaviour to gain rewards or when anticipatory preparation to receive an inevitable punishment is at stake). In appetitive contexts, for example, it might be necessary to grasp food or to swallow juice right in

time to avoid unwanted loss (Hollerman and Schultz, 1998). Similarly, predicting an aversive event can favour bodily adaptation to something potentially annoying, painful or even dangerous.

Maladaptive implications

The case of Parkinson's Disease associated psychosis

When altered, stimulus-reward associations are likely to favour assignment of salience to otherwise irrelevant external and internal states, thus favouring abnormal salience processing (Miller at al., 1976; Berridge et al., 1998). Maladaptive forms of reward processing, mediated by the dopaminergic system, are believed to be at the core of the formation of psychotic symptoms (Fletcher and Frith, 2009; Kapur, 2003; Heinz, 2002). A large number of studies have shown that activity in the ventral striatum (VS) (Bernacer et al., 2013; Murray et al., 2008; Fusar-Poli and Meyer-Lindenberg, 2012; Juckel et al., 2006a; de Leeuw et al., 2015; Schlagenhauf et al., 2014), medial prefrontal cortex (mPFC) (Schlagenhauf et al., 2009; Bernacer et al., 2013) and inhibitory functions of the medial temporal lobe (Lodge and Grace, 2011) are altered in patients with psychosis or at risk of psychosis.

In particular, studies reported impairments in various different stages of reward-processing: Reward Anticipation (Grimm et al., 2014; Esslinger et al., 2012; Juckel et al., 2006b), Reward Feedback (Waltz et al., 2010; Schlagenhauf et al., 2009) and Prediction Error (Murray et al., 2008b; Griffiths et al., 2014; Morris et al., 2012; Corlett and Fletcher, 2012).

Psychiatric symptoms - such as, psychosis, depression, anxiety and apathy - are a frequent comorbidity in Parkinson's Disease, and for long have been regarded as mainly side effects of the dopaminergic treatments (Gallagher et al., 2012). However, the role of dopaminergic medication is still controversial and some evidences point to a more complex interaction between disease and treatment-related effects (Park and Stacy, 2009). PD associated psychosis is present in about 15-40% of treated PD patients and has a major influence on quality of life and likelihood of nursing home placement (Gallagher et al., 2012; Park and Stacy, 2009).

To date, there have been a paucity of attempts to integrate the pathophysiology of psychotic experiences in Parkinson's Disease with the current models of their pathogenesis in schizophrenia. Leading candidates to understand the presence of psychosis in Parkinson's Disease include medication effects, perhaps operating in the basal ganglia and cortex, and/or disruption of cortical function secondary to pathology. Dopaminergic treatment is a strong risk factor for the manifestation of psychosis in Parkinson's Disease, and elevation of dopamine levels is strongly implicated in the psychosis of schizophrenia. However, the explanation cannot be that simple, as less than half the patients with Parkinson's Disease manifest psychosis.

According to cognitive neuropsychiatric theory and the National Institute of Health RDoc programme, the pathophysiological basis of psychiatric symptoms may be better understood in terms of dysfunction of underlying domains of neurocognition, cutting across traditional diagnostic boundaries, rather than simply according to diagnosis. Abnormal frontal striatal reinforcement learning signalling, associated with dysregulated dopamine, has been proposed as a key processes contributing to the pathogenesis of psychosis (Miller 1976; Kapur 2003). This

theory has received empirical support in the study of psychosis in schizophrenia spectrum disorders and preclinical models of psychosis, but has not been tested in PD associated psychosis.

Driving behaviour

Influences between Pavlovian and Instrumental forms of reinforcement learning

Goal-directed behaviour can be variably influenced by external and internal factors which impact the values and priorities assigned to rewards and goals (Doya, 2008). One of the most simple and effective mechanisms for influencing choice is reinforcement learning. Reinforcement learning allows animals to connect spatially and/or temporally related events in order to predict future events. Given the complexity of the animal's environment, learning that an arbitrary cue (e.g., a sound) is predictive of a certain goal (e.g., obtain a reward, such as food), allows the animal to learn a flexible response that facilitates achievement of the goal itself. In most cases such *cue-controlled behaviour* is adaptive; for example it helps one obtain food when hungry (Holmes, Marchand, and Coutureau, 2010; Perks and Clifton, 1997). However, an inflexible association can lead to perseverance in the same choice even if the goal itself is no longer available, or has negative long-term consequences (Holmes et al., 2010). For example, a cue associated with drugs can induce relapse even when the drug is not voluntary sought, and a sign associated with food can induce craving in the absence of hunger, leading to compulsive over-eating (Volkow, Wang, Fowler, and Telang, 2008). These biases on voluntary choice are also implemented in

marketing strategies, such as advertisements, to influence consumer behaviour (Bray, Rangel, Shimojo, Balleine, and O'Doherty, 2008; de Wit and Dickinson, 2009; Smeets and Barnes-Holmes, 2003). Cue-controlled behaviours have been interpreted as the endpoint of an initial intentional seeking behaviour (of a reward), which leads to habitual, and ultimately compulsive, conduct characterized by a loss of control over behaviour (Everitt and Robbins, 2005). This interesting framework proposes that the transition from intentional volition to habit and compulsion can be explained by interactions between Pavlovian and instrumental learning processes: a reward acts as an instrumental reinforcer by enhancing actions that are able to produce it, while Pavlovian learning confers incentive salience to cues (Conditioned Stimuli or CS) closely associated with the reward (Everitt and Robbins, 2005). Such cues can elicit craving and motivation towards the associated reward, thus biasing choice. Well-known evidence of this effect can be found in the so-called Pavlovian-to-Instrumental Transfer (PIT) effect (Estes, 1943, 1948). PIT captures the ability of a Pavlovian cue (i.e., a CS associated with a reward) to increase the likelihood of an instrumental response independently paired with the same (specific transfer), or a similar (general transfer), reward (de Wit and Dickinson, 2009; Holmes et al., 2010; Rescorla and Solomon, 1967). This effect emerges without any formal association between Pavlovian and instrumental contingencies, and even when the reward itself is no longer available (Talmi, Seymour, Dayan, and Dolan, 2008). PIT has been mainly studied in non-human animals (Balleine, 1994; Colwill and Rescorla, 1988; Corbit and Balleine, 2003; Delamater and Holland, 2008; Delamater, 1995, 1996; Holland and Gallagher, 2003; Holland, Petrovich, and Gallagher, 2002; Holland, 2004; Lovibond, 1981; Rescorla and Solomon, 1967; Rescorla, 1994a, 1997, 2000) (for review, see Dickinson and Balleine, 1994, 2002; Holmes et al., 2010), but some recent

studies have also reported this effect in humans (Allman, DeLeon, Cataldo, Holland, and Johnson, 2010; Bray et al., 2008; Hogarth, Balleine, Corbit, and Killcross, 2013; Hogarth, Dickinson, and Duka, 2010; Hogarth, Dickinson, Wright, Kouvaraki, and Duka, 2007; Hogarth, Field, and Rose, 2013; Lovibond and Colagiuri, 2013; Nadler, Delgado, and Delamater, 2011; Paredes-Olay, Abad, Gámez, and Rosas, 2002; Prévost, Liljeholm, Tyszka, and O'Doherty, 2012).

The importance of a deeper understanding of the interactions between Pavlovian and instrumental learning processes comes from evidences suggesting that transfer effect is a candidate mechanisms underlying addiction (e.g., drug seeking and relapse) (Hogarth at al., 2013; 2010; 2007; Watson et al., 2012; 2013; 2014), compulsions (Everitt and Robbins, 2005) and other neuropsychiatric disorders (Dayan et al., 2006; Dayan and Huys, 2008; Boureau and Dayan, 2011).

Forms of transfer

During Pavlovian conditioning a CS is connected to the reinforcer by both a direct motivational representation of its value and an indirect representation of its sensory features (Dickinson and Balleine, 2001). This differentiation is thought to be reflected in two kinds of transfer effects: in a *general* form of transfer, the CS invigorates instrumental responses paired with motivationally similar reinforcers; whereas, in an *outcome-specific* form transfer, the CS exerts its influence selectively on instrumental responses associated with the exact same reinforcer.

Some authors suggested that the difference between outcome-specific and general forms of transfer can be reflected in a different underlying learning model. Outcome-specific transfer - depending on a detailed match of specific reinforcers with the respective CS and instrumental response – should reflect a model-based process; conversely, general transfer - depending solely on the motivational valence of the reinforcer - should reflect a model-free process (Dolan and Dayan, 2013).

However, this theoretical differentiation does not inform about the nature of such behaviours. From a clinical perspective, it would be more interesting to understand if specific and general transfer reflect goal-directed or habit-like behaviours.

Instrumental behaviour is considered goal-directed when it reflects knowledge of the relationship between response and outcome (R-O control) and if it is motivationally relevant (desirable). Conversely, it is considered habitual when it depends on past reinforcement and is divorced from the current value of the outcome (Dolan and Dayan, 2013). Goal-directed is highly flexible, but computationally expensive; habitual is automatic and computationally efficient, but inflexible to the requirements of a changing environment (Dolan and Dayan, 2013; Dayan, 2009).

Since transfer effect implies an external Pavlovian control on instrumental responses (potentially governed by both model-based and model-free mechanisms), it can hardly reflect a goal-directed behaviour. In fact, by definition, transfer is observed in extinction, which means in absence of the actual reinforcer (the goal).

A widely used procedure in literature to directly test the goal-directed nature of an instrumental response is the so-called devaluation (Dolan and Dayan, 2013). Devaluation procedures aim to

weaken the value of a reinforcer by using a manipulation that should determine a reduced instrumental performance. For example: in appetitive contexts, food can be paired with a toxin (e.g., Colwill and Rescorla, 1988) or satiated (e.g., Balleine and Dickinson, 1998); in aversive contexts, an electric-shock stimulator can be visibly disconnected from the participant to ensure absence of pain (e.g., Gillan et al., 2014). Devaluation procedures reduce goal-directed instrumental responses oriented to obtaining rewards or avoiding potential punishments.

Another variable that can favour the formation of habits is the continuative and repetitive practice of an action, which can be experimentally reproduced with instrumental overtraining (Yin and Knowlton, 2002; Tricomi, Balleine, and O'Doherty, 2009; Dayan and Berridge, 2014). Instrumental overtraining reduces the impact of the reinforcer on the instrumental responses, so that the performance is not anymore guided by the current value of the outcome of the action (goal-directed), but it's rather habitual. For example, in humans, instrumental overtraining can determine unnecessary avoidance responses to previously aversive stimuli, but now safe (Gillan et al., 2014).

Individual differences

An important, but still neglected, aspect in the human literature about PIT concerns individual differences. In the animal literature, the extent to which a Pavlovian cue becomes attractive and exerts a biasing effect varies between individuals. In particular, Sign-Trackers (ST) and Goal-Trackers (GT) have been shown to have different learning styles, consisting of a tendency to attribute more or less incentive salience to Pavlovian reward-associated cues. In a typical

Pavlovian conditioning paradigm, a CS (e.g., lever presentation) is paired with a reward (e.g., food pellet), which is delivered in a different spatial position. In such a situation, two different Conditioned Responses (CR) (i.e., learned responses to a previously neutral stimulus) might be expressed. Some animals approach and engage the CS (the Sign) itself and, only after its termination, reach the location of reward delivery; other animals, upon CS presentation, immediately engage the location of reward delivery (the Goal), even if it is not yet available. The first CR has been categorized as Sign-Tracking behaviour, while the second CR has been categorized as Goal-Tracking behaviour. ST and GT can be conceived of as different learning styles, expressed through a specific CR during Pavlovian learning. ST behaviour is thought to arise from the attribution of incentive salience to Pavlovian reward-paired cues, which consequently become a powerful source of motivation for future behaviour (Flagel et al., 2011). In ST, incentive stimuli become attractive, eliciting approach towards them and promoting potentially maladaptive cue-controlled behaviours; ST individuals, indeed, are generally more vulnerable to addiction and relapse (Flagel, Watson, Akil, and Robinson, 2008; Robinson and Flagel, 2009; Tomie, Aguado, Pohorecky, and Benjamin, 1998). The ST and GT profiles do not seem to be limited to the CR expressed, but are also associated with differences in traits such as impulsivity; ST individuals are characterized by higher levels of impulsive behaviour compared to GT individuals (Flagel et al., 2009; Tomie, Aguado, Pohorecky, and Benjamin, 2000).

A deeper investigation into individual differences in attributing incentive salience to reward-paired stimuli would thus be important for understanding and reducing the propensity to develop maladaptive behaviours.

PART II - EXPERIMENTAL STUDIES ON PREDICTION ERROR

Study 1 - Mediofrontal negativity signals unexpected omission of aversive events.

Abstract

Research based on event-related potential (ERP) reported mediofrontal negativities following unexpected negative feedback or performance error. Some authors proposed that these signals reflect reward prediction error for worse than expected outcomes, while others suggested that mediofrontal negativities express medial prefrontal cortex coding for unexpected non-occurrence of a predicted outcome, whether worse or better than expected. Many studies found mediofrontal negativities coding for unexpected negative outcomes; however, few studies found them after unexpected positive outcomes. The present study investigated ERP and skin conductance response (SCR) to the unexpected omission of electric shocks during Pavlovian aversive conditioning. To manipulate expectancies, participants were presented with visual stimuli paired with electric shock on either 80% ($CS+_1$) or 20% ($CS+_2$) of trials. SCR analysis confirmed higher shock-delivery expectancy for $CS+_1$, relative to $CS+_2$. ERP analysis evidenced a stronger negative frontocentral ERP component after unexpected, relative to expected, shock-omission. Methodological and theoretical implications are discussed.

Experimental Design

The aim of the present study was to investigate ERPs for the unexpected omission of a physical pain (an electric shock); thus, a prediction error signal for a positively valenced event (in this case, a relief condition). More specifically, this study aims to directly test the prediction made by the PRO-model (Alexander and Brown, 2011; 2014) that the unexpected non-occurrence of a salient negative event (i.e., a relief event) triggers mediofrontal negativity. To this end, the unexpected non-occurrence has been operationalized as the absence of a predicted electric shock, rather than as the presentation of a different stimulus than the one expected. A mediofrontal negativity triggered by the unexpected omission of a predicted negative outcome would lead to the conclusion that an event was expected and that its omission produced a prediction error signal.

For this purpose, a Pavlovian aversive conditioning paradigm with a partial reinforcement schedule was used. On each trial, participants were presented with a visual stimulus (Japanese kana) displayed for 4 seconds, with a variable 7-9 seconds inter-trial interval (ITI). On some trials, a mild electric shock was delivered during the final 200ms of stimulus presentation (Fig. 1.1). Participants performed a total of 640 trials divided in 16 blocks. During a single block, 2 different stimuli were presented 20 times each. Subjects were required to press a keyboard button correspondent to the left or right presentation of the stimulus and they were informed that this response had no effect on shock-delivery. In order to manipulate expectancies, one of the two stimuli presented in each block was paired with an electric shock on 80% of trials (CS+1), whereas the other was paired with an electric shock on 20% of trials (CS+2). Thus, by learning associations between stimuli and shock-delivery probability, participants should build up a high

shock-delivery expectancy for CS+1 stimuli and a low shock-delivery expectancy for CS+2 stimuli. Two critical conditions were compared: unexpected shock-omission (CS+1 trials without shock-delivery: CS+1ws) and expected shock-omission (CS+2 trials without shock-delivery: CS+2ws). On CS+1ws trials, not receiving a shock represents an unexpected omission, while on CS+2ws trials, not receiving a shock represents a predicted event. Skin Conductance Response (SCR) and electroencephalogram (EEG) were recorded during the task: SCR analysis was used as a somatic indicator of expectancy, thus it was used to test the presence of a different shock-delivery expectancy between CS+1ws and CS+2ws (Schiller et al., 2008). The EEG was recorded to examine ERPs associated with expected and unexpected shock-omission. If unexpected omission of a negative event (i.e, a positively valenced relief event) triggers mediofrontal negativity as predicted by the PRO-model, then a stronger mediofrontal negativity for unexpected shock omission relative to expected shock omission should be observed. Moreover, in order to investigate possible relationships between personality traits related to behavioural responsiveness to reward (Gray, 1990) and ERPs associated with positively valenced unexpected omission, Behavioural Activation System (BAS) inventory (Carver and White, 1994) was filled by participants at the end of the task.

Materials and Methods

Participants

Twenty volunteers with no history of neurological diseases were randomly recruited from the student population of the University of Bologna. Two participants completed only one of the three testing sessions and three other participants were excluded for excessive eye artifacts. Thus, the final sample included fifteen (7 female, 1 left-handed) subjects between 19 and 29 years of age (mean =23.93; sd =2.37). All participants gave informed consent. The study was conducted in accordance with institutional guidelines and was approved by the Department of Psychology's ethical committee.

Stimuli and materials

The task consisted in a Pavlovian aversive conditioning paradigm with a partial reinforcement schedule. Visual stimuli were used as conditioned stimuli (CSs) and mild electrical shocks served as unconditioned stimuli (USs). A CS consisted in a 3cm white square with a Japanese kana on it. All the stimuli were balanced for luminance, complexity and colour saturation and were displayed on a 17-inch colour monitor with a black background, at a viewing distance of 80 cm. Electrical shocks were pulses of 200 ms duration generated by a Digitimer Stimulator (Model DS7, Digitimer Ltd., Hertfordshire, United Kingdom). Mild shocks were administered to the inner wrist of the dominant hand, to which two SU15N1 electrodes (SEI EMG, Padova) were attached. The shock intensity was individually set before the task: stimulation was initially set at 5 and the intensity was gradually increased (1 mA increments) to a level the participant indicated

as “uncomfortable, but not painful” (Schiller et al., 2008). The mean shock intensity was 4.6 mA (sd =2.1 mA; min =1.8 mA; max =8.9 mA).

A single trial consisted in the presentation of one of two possible CSs, on the left or the right side of the screen: one was paired with an electric shock on 80% of trials ($CS+_1$) and one was paired with shock on 20% of trials ($CS+_2$). The CSs were presented for 4 seconds, with a 7-9 seconds inter-trial interval (ITI) during which a fixation point was presented (Fig. 1.1). Stimuli were presented 8.5 cm to the left or to the right of fixation, and participants were asked to press a left or a right button corresponding to the side of stimulus presentation. This procedure was introduced in order to keep subject’s attention focused on the stimuli. Participants performed a total of 640 trials divided in 16 blocks. Each block consisted of 40 trials, equally divided between the two conditions ($CS+_1$ and $CS+_2$). In each block, 2 out of 32 different Japanese kana were used as CSs and their assignment to a specific category ($CS+_1$ and $CS+_2$) was counterbalanced across subjects. Each block was internally divided in two identical hemiblocks containing half of the stimuli. At the very beginning of each block, there was a brief learning phase during which 5 stimuli for each category (10 stimuli total) were visualized. This phase was not considered in the subsequent analysis. The 16 blocks were spread across three testing days: 5 on the first day, 6 on the second day and 5 on the last day. Between individual testing sessions there was a variable interval lasting from 1 to 4 days. A PC running Presentation software (Neurobehavioural Systems, Albany, CA) controlled stimulus presentation.

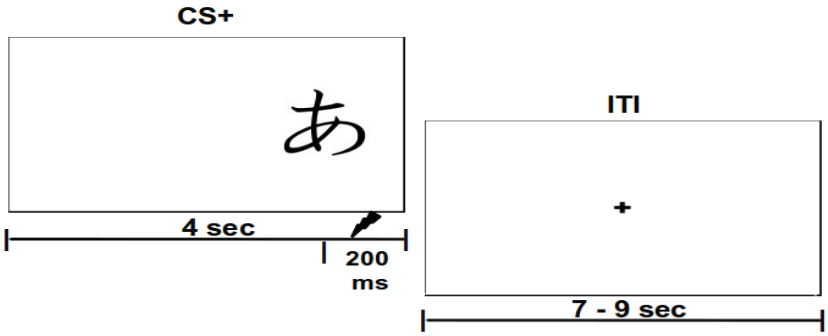

Figure 1.1
Schematic representation of a single trial.

Procedure

On arrival, participants were comfortably seated in a silent room and their position was centered relative to the screen. Both SCR and EEG signals were collected continuously during the task and stored for off-line analysis. Participants were required to remain as quiet and still as possible, in order to avoid confounding effects on measurements. The experimental session began with a 10 minutes rest period during which the participant acclimated to the environment and a correct attachment and conductance of the electrodes - for both SCR and EEG - was ensured. Then, the participants' ability to generate SCRs to external stimuli was tested by generating loud sounds (i.e., clapping of the hands). Subsequently, the intensity of the electrical stimulation was individually set as previously described. Before the beginning of the task, participants were told that they would see visual images, on the right or left side of a computer screen, that might be followed by a mild electrical shock. They were required to pay attention to the screen and press a keyboard button correspondent to the left or right presentation of the stimulus. Participants were informed that their response about side of presentation had no effect on shock-delivery. Subjects

were also told that they could take a short break between each block, if they wished. At the end of each day, all participants underwent an extinction session in order to eliminate the effects of the aversive conditioning. At the end of the last session, all subjects filled the Italian version (Leone et al., 2002) of BAS inventory (Carver, C. S. and White, 1994).

SCR recording and analysis

The skin conductance response (SCR) was recorded using Ag/AgCl electrodes (TSD203 Model; Biopac Systems, Goleta, CA), filled with isotonic hyposaturated conductant. Electrodes were attached to the volar surface of the middle and index fingertip of the nondominant hand and held with Velcro straps. The signal was recorded using a DC amplifier (Biopac GSR100) with a gain factor of 5 μS/V and low-pass filter set at 10 Hz. The analog signal was digitized using the MP-150 digital converter (Biopac Systems) at a rate of 200 Hz and fed into AcqKnowledge 3.9 recording software (Biopac Systems).

SCR data acquired during the task were offline analysed by using custom made scripts realized on MATLAB (2011a, The MathWorks, Inc., Natick, Massachusetts, United States). Statistical analyses were performed with SPSS 19.0 (IBM Corp. Released 2010. IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp.). The skin conductance response was transformed into microsiemens and calculated for each trial. SCR was calculated as the peak-to-peak amplitude difference of the largest deflection in the 0.5-4.5 sec latency window after stimulus onset (Schiller et al., 2008). The minimal response criterion was 0.02 μs and smaller responses were encoded as zero; raw skin conductance scores were square root transformed to normalize the

distributions and scaled to each subject's maximal US response to account for interindividual variability (Schiller et al., 2008). Only trials without shock occurrence were analysed.

EEG recording and analysis

Electroencephalographic (EEG) signal was recorded with Ag/AgCl electrodes (Fast'n Easy Electrodes, Brain Products, Gilching, Germany) from 26 electrode sites (Fp1, F3, F7, FC1, FC5, C3, T7, CP1, P3, P7, O1, Fz, FCz, Cz, Pz, Fp2, F4, F8, FC2, FC6, C4, T8, CP2, P4, P8, O2), as well as from the right mastoid. The left mastoid was used as reference electrode, and the ground electrode was placed on the right cheek. All electrodes were offline rereferenced to the average of both mastoids. Vertical and horizontal electrooculogram (EOG) was recorded from above and below the left eye and from the outer canthi of both eyes. Both the EEG and EOG were recorded with 0.01-100 Hz band-pass of and amplified by BrainAmp DC amplifier (Brain Products, Gilching, Germany). The amplified signals were digitized at a sampling rate of 500 Hz and were offline filtered with a 25 Hz low-pass filter.

EEG data acquired during the task were offline analysed by using MATLAB (2011a, The MathWorks, Inc., Natick, Massachusetts, United States) and EEGLAB 11.0.5.4b free toolbox (Delorme and Makeig, 2004). Statistical analyses were performed with SPSS 19.0 (IBM Corp. Released 2010. IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp.). Epochs of 400 ms before and 800 ms after stimulus offset were extracted from the continuous EEG. The average voltage during the time interval from 400 ms to 200 ms before stimulus offset was used as baseline, as the shock was delivered 200 ms before stimulus offset on trials with shock-delivery.

In this way, no influence of expectancy violation or confirmation on the ERP was possible during the baseline window. First, epochs presenting large artifacts were identified and excluded using a rejection method implemented in EEGLAB (pop_autorej, Onton and Delorme, SCCN/INC/UCSD, 2007 with default parameters). Second, EOG artifacts were corrected using a regression approach (Gratton et al., 1983). The mean percentage of excluded trials was 12% (sd =0,4%). Epochs were then averaged separately for each participant and each condition. ERPs were calculated as the local peak amplitude in a 250-350ms interval after stimulus offset.

BAS Inventory

The Behavioural Activation System (BAS) inventory is a personality self-reported questionnaire proposed by Carver and White (1994). And based on Gray's personality theory (1990). It is composed of 13 items, each one is a statement that a person may either agree or disagree with. It is required to indicate how much you agree or disagree with that item on a 4 point likert scale (from "totally agree" to "totally disagree"). The inventory can be divided in three subscales: Drive (BAS_D, reflecting the strength with which reward outcome guides subsequent behaviour), Reward Responsiveness (BAS_{RR}, indexing the degree to which a person derives pleasure from reward) and Fun Seeking (BAS_{FS}, reflecting a novelty-seeking trait). The BAS is generally believed to control appetitive motivation and is sensitive to signals of reward and escape from punishment. Activity in this system elicits movement toward goals and to engage in goal-directed efforts. Greater BAS sensitivity is reflected in greater experience for positive feelings and is supposed to be unrelated to negative affect.

BAS scales were scored according to the procedures indicated from the authors and final scores (BAS_{TOT} mean =40.47, sd =3.51; BAS_{D} mean =12.14, sd =1.88; BAS_{RR}, mean =18.4, sd =1.68; BAS_{FS} mean =9.94, sd =2.40) were consistent with those reported in literature (Carver and White, 1994).

Results

SCR

SCR analysis was performed to check if there was a higher shock-delivery expectancy in the $CS+_{1ws}$ condition relative to the $CS+_{2ws}$. SCR was calculated as the peak-to-peak amplitude difference of the largest deflection in the 0.5-4.5 seconds latency window after stimulus onset (Schiller et al., 2008). To analyze how rapidly expectancies were build up and to determine possible habituation effects on SCR level (Dawson et al., 1990), each block was divided in two hemiblocks, containing 20 stimuli each. Assumptions of normal distribution, independence of residuals and sphericity were verified. A 2x2 repeated measure ANOVA with Time (first/second hemiblock) and Condition ($CS+_{1ws}$ /$CS+_{2ws}$) as within factors was performed. A significant main effect of Time ($F(1, 14) =26.27$; two-tailed $p <.0001$; partial $\eta^2 =.68$; $N =15$) was found, with the first hemiblock (mean =.48 µS) presenting a significantly higher SCR level than the second hemiblock (mean =.37 µS). This indicates that the general electrodermal responsiveness to the task is reduced in the second hemiblock. Furthermore, a significant interaction Time X Condition ($F(1, 14) =11.06$; two-tailed $p =.006$; partial $\eta^2 =.48$; $N =15$) was found. Bonferroni-corrected

post-hoc analysis on the interaction effect revealed a significant difference (p =.04) between $CS+_{1ws}$ (mean =.52 µS; sd =.05 µS) and $CS+_{2ws}$ (mean =.44 µS; sd =.04 µS) in the first hemiblock; no significant effect emerged for the second hemiblock (p =.90) (Fig. 1.2). This indicates that subjects showed a higher arousal level when presented with the $CS+_{1ws}$ stimulus relative to $CS+_{2ws}$ in the first hemiblock consistent with a higher shock-delivery expectancy on these trials (Schiller et al., 2008). No further effects reached significance (ps >.27).

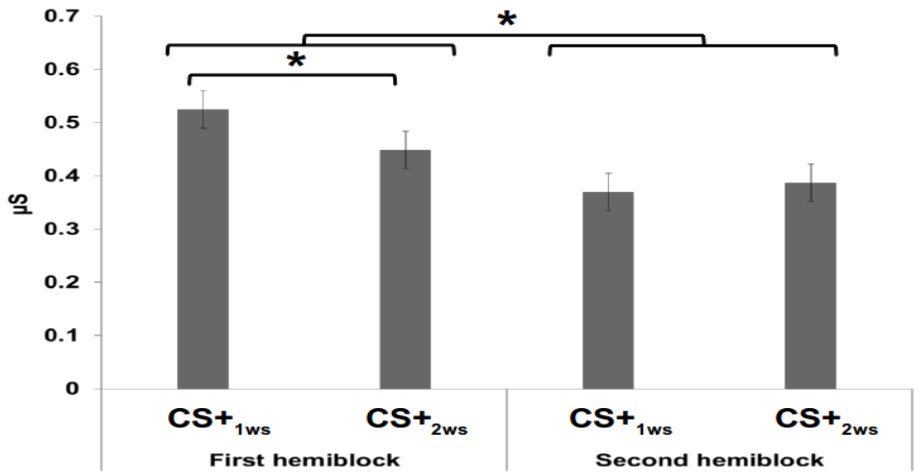

Figure 1.2
SCR levels during the two hemiblocks. Bars indicate standard error. $CS+_{1ws}$ corresponds to the expected shock-delivery condition; $CS+_{2ws}$ corresponds to the unexpected shock-delivery condition.

The analysis of SCR data confirmed that different expectancies were successfully generated in the first hemiblock as subjects presented a higher SCR in the $CS+_{1ws}$ condition, as compared with $CS+_{2ws}$ condition. In sum, the SCR data suggest that the build-up of expectancy was very rapid

during the first hemiblock and then habituation hid the effect in the second hemiblock (Dawson et al., 1990).

Behavioural data

Percentage of errors and reaction times were also analysed. Assumptions of normal distribution, independence of residuals and sphericity were verified. Two separate 2x2 repeated measure ANOVAs with Time (first/second hemiblock) and Condition ($CS+_{1ws}$ /$CS+_{2ws}$) as within factors were performed.

Neither main effects, nor interaction effect emerged from the analysis on the percentage of errors (ps >.17). The average percentage of errors was 1.24 (first hemiblock: $CS+_{1ws}$ = 0.96 and $CS+_{2ws}$ =1.20; second hemiblock: $CS+_{1ws}$ = 1.37 and $CS+_{2ws}$ =1.42).

The analysis of reaction times showed a significant main effect of Condition ($F(1, 14) = 8.63$; two-tailed $p = .01$; partial $\eta^2 = .38$; $N = 15$), with $CS+_{1ws}$ (mean =520.86 ms; sd =150.87 ms) presenting faster reaction times than $CS+_{2ws}$ (mean =534.24 ms; sd =159.43 ms). A marginal effect of Time ($p = .06$) and a significant Time X Condition interaction were found. Bonferroni-corrected post-hoc analysis on the interaction effect revealed a significant difference ($p = .03$) between $CS+_{1ws}$ (mean = 523.77 ms; sd = 149.42 ms) and $CS+_{2ws}$ (mean = 540.63 ms; sd = 161.03 ms) in the second hemiblock; a marginally significant effect emerged also for the first hemiblock ($p = .058$), with $CS+_{1ws}$ (mean = 517.36 ms; sd = 142.10 ms) presenting slightly faster reaction times than $CS+_{2ws}$ (mean = 527.55 ms; sd = 147.02 ms). These results suggests that, although shock-delivery was not relevant for the task, subjects were faster in indicating the side

of presentation of the stimulus when presented with the cue which was highly associated with shock. This effect was stronger in the second hemiblock, relative to the first, when stimulus contingencies were better learned and expectations were stronger.

EEG

The aim of EEG analysis was to investigate the presence of an mediofrontal negativity-like component during unexpected, relative to expected, shock omission. mediofrontal negativity has been reported to present its maximum negative peak between 250 and 350ms following the event indicating the prediction error (Hajcak et al., 2006). In the present task, participants can be completely sure about shock omission and consequently, about whether their expectancy is confirmed or violated by the actual course of events only upon stimulus offset. Thus, this event represents the point in time where the prediction error occurs. Therefore, ERPs were calculated as the most negative peak in a 250-350ms interval after stimulus offset, in unexpected shock-omission trials ($CS+_{1ws}$) and expected shock-omission trials ($CS+_{2ws}$). Grandaverage scalp topographies showed a frontal maximum activation for the difference between these two conditions (Fig. 1.3).

Statistical analysis were conducted on Fz electrode, where mediofrontal negativity has been previously reported (Talmi et al., 2013; Hajcak et al., 2006). Assumptions for a correct use of t-test (normal distribution, independence of residuals and sphericity) were verified. A paired t-test to compare $CS+_{1ws}$ (mean trial number =58; sd =4) and $CS+_{2ws}$ (mean trial number =288; sd =16) was performed. A significant difference between $CS+_{1ws}$ (mean = -0.31 μV; sd =2.40 μV) and

$CS+_{2ws}$ (mean =.95 μV; sd =1.82 μV) was found (t(14) = -2.73; two-tailed p =.016; partial η^2 =.34; N =15). These results indicate that the unexpected omission of a shock ($CS+_{1ws}$ condition) generates a stronger frontal negative ERP component relative to the expected omission of a shock ($CS+_{2ws}$ condition) (Fig. 1.4).

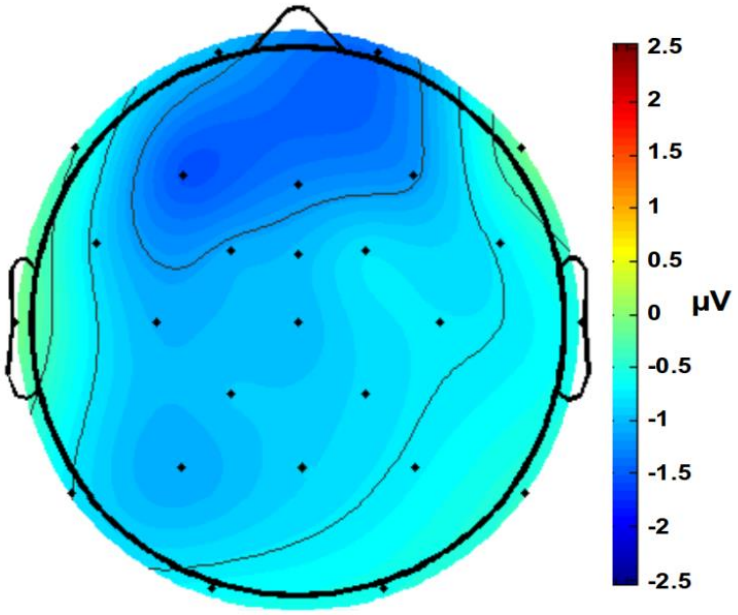

Figure 1.3
Scalp distribution of the difference between expected and unexpected shock-omission in a time window of 250-350ms after stimulus offset.

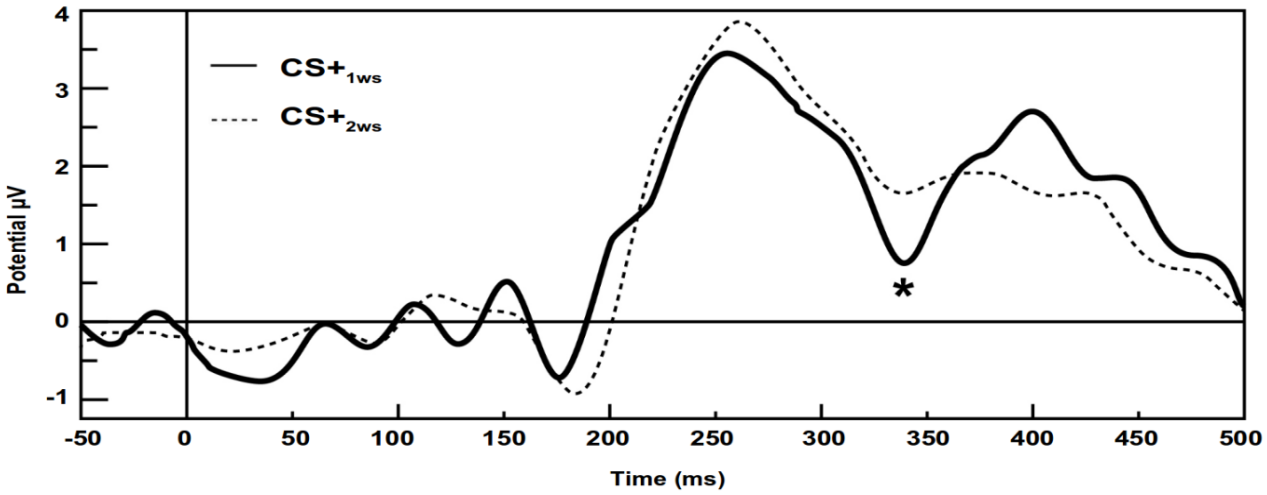

Figure 1.4
Grandaverage ERP waveforms from channel Fz. The asterisk shows the maximum peak amplitude reached in the considered time window (250-350 ms). $CS+_{1ws}$ corresponds to the unexpected shock-omission condition; $CS+_{2ws}$ corresponds to the expected shock-omission condition. 0 represents stimulus offset.

Moreover, to further investigate the relationship between shock-delivery expectancy and neural activity consequent to its violation, a correlation between shock-delivery expectancy (as expressed by the difference in SCR signal during the two expectancy conditions $CS+_{1ws}$ and $CS+_{2ws}$) and unexpected omission neural signal (as indexed by the difference in ERPs originated in Fz during $CS+_{1ws}$ and $CS+_{2ws}$) was analysed. A significant positive correlation ($r = .61$, one-tailed $p = .01$; $N = 15$) was found. Critically, this effect indicates that the stronger shock-delivery expectancy (as measured by psychophysiological arousal), the stronger prediction error for unexpected shock-omission (as reflected by cortical ERPs) (Fig. 1.5).

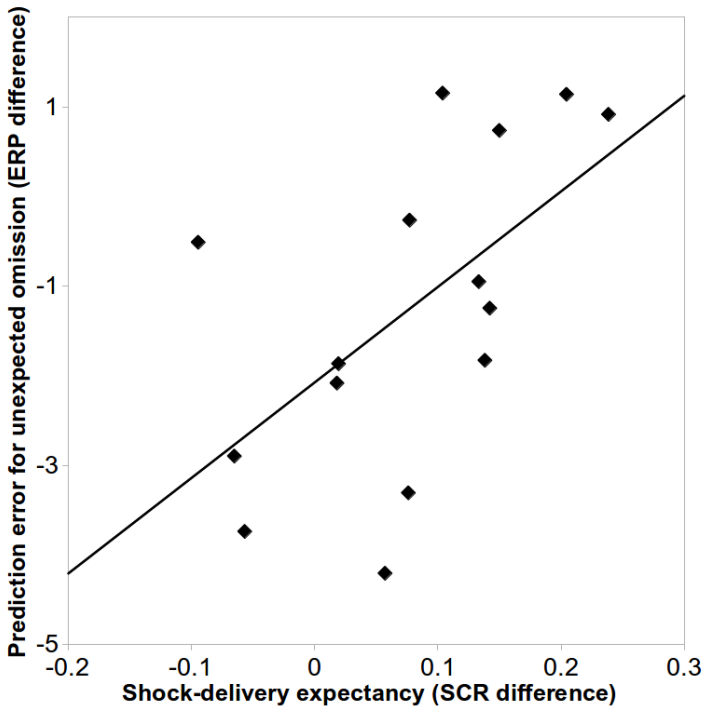

Figure 1.5
Scatter-plot of the correlation between shock-delivery expectancy (SCR difference) and prediction error neural signal for unexpected shock-omission (ERP difference).

BAS Inventory

To examine the relationship between reward responsiveness and neural activation consequent to the unexpected omission of shock-delivery, participants were required to fill the Behavioural Activation System (BAS) inventory. Correlation between BAS (total score and subscales) and ERP peak amplitude during unexpected shock-omission condition was calculated. A marginally significant (for Bonferroni-corrected α =.0125) negative correlation emerged with BAS-Drive (BAS_D) subscale (r =-.54; one-tailed p =.019; N =15). This negative correlation indicates that the higher BAS_D is, the more negative is the peak for unexpected omission. Thus, people with high levels of reward responsiveness tended to present a more negative ERP peak when a shock

unexpectedly did not occur, a situation that is better than expected (i.e., a relief) (Fig. 1.6). No correlations with other BAS subscales or the total score reached significance (ps >.08).

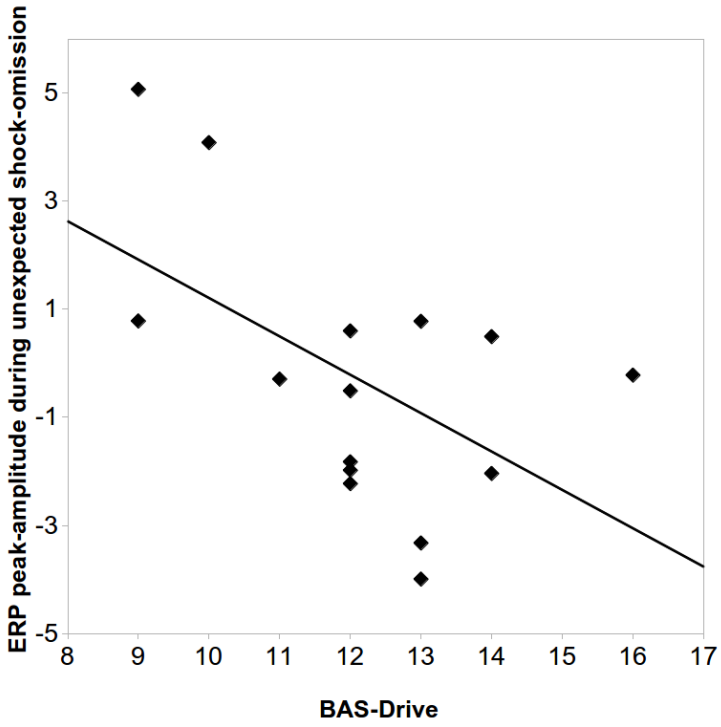

Figure 1.6
Scatter-plot of the correlation between BAS_D score and ERP peak-amplitude in the unexpected shock-omission condition.

Conclusion

EEG and SCR were recorded during a Pavlovian aversive conditioning procedure, in which participants learned to associate visual stimuli to a high ($CS+_1$) or a low ($CS+_2$) shock-delivery probability. This made it possible to contrast and compare two critical conditions: unexpected shock-omission ($CS+_{1ws}$) and expected shock-omission ($CS+_{2ws}$). SCR analysis allowed ensuring that subjects actually differentiated between the two stimuli and that different shock-delivery expectancies were created. This evidence was also supported by reaction time analysis. A larger

SCR response indicating a higher arousal level at $CS+_{1ws}$, relative to $CS+_{2ws}$, was found. From this result, it can be inferred that there was a higher shock-delivery expectancy for the $CS+_{1ws}$ condition relative to the $CS+_{2ws}$ condition. ERPs in the 250-350ms interval after stimulus offset, i.e., the point in time when subjective expectancy was confirmed or violated, were calculated for the unexpected ($CS+_{1ws}$) and the expected ($CS+_{2ws}$) shock-omission conditions. Analysis showed the presence of a stronger frontocentral negative ERP component in the $CS+_{1ws}$ condition, relative to $CS+_{2ws}$. Thus, the unexpected omission of a predicted shock generated a stronger frontocentral negativity, signalling unexpected omission.

Time window, frontocentral localization and negative deflection detected in the unexpected omission condition are coherent with an mediofrontal negativity-like component (Hajcak et al., 2006). Moreover, previous fMRI studies revealed mPFC activity after unexpected shock-omission during Pavlovian aversive conditioning paradigm, which is in line with ERN and mediofrontal negativity source localization reported in the literature (Nieuwenhuis et al., 2004).

Since a mediofrontal negativity component was found after a positively valenced unexpected omission event, the present results are not in line with RL-ERN theory account of mediofrontal negativities as selectively coding for worse than expected outcomes (Holroyd and Coles, 2002). The present findings are, instead, coherent with the view that mediofrontal negativities can also be triggered by a positive event (Oliveira et al., 2007; Talmi et al., 2013; Ferdinand et al., 2012) thus supporting PRO-model conception that the mPFC detects both positive and negative unexpected omissions (Alexander and Brown, 2011). This view seems to be able to account for a wide range of findings, thus representing a promising unifying framework for mPFC functioning.

It assumes that the mediofrontal negativity reflects general outcome monitoring on the basis of expectancy violations irrespective of outcome-valence aspects, thus explaining that positive and negative unexpected feedback elicit similar mediofrontal negativities. Similarly, also the ERN (Holroyd et al., 2003; 2007) following performance errors in rapid choice tasks can be explained by the fact that an error violates the expectancy of a correct response (Holroyd et al., 2002; San Martìn et al., 2010). fMRI studies also reported enhanced mPFC activation to unexpected wins as compared with losses, when losses are more frequent than wins; thus, suggesting that error effect reflected in ERN results from a comparison between actual and expected outcomes, rather than by the error feedback per se (Jessup et al., 2010).

According to the PRO-model, the size of the mediofrontal negativity is proportional to the unexpectedness of the outcome (Alexander and Brown, 2011). Critically, as a demonstration of the link between expectancies and the observed negative ERP component, correlation analysis in the present study showed that higher shock-delivery expectancy, as expressed in SCR, is linked to a greater neural unexpected omission signal, as expressed in frontal negative ERPs at stimulus offset. Thus, in accordance with the predictions of the PRO-model, the frontal unexpected omission signal is larger when a stronger expectancy is disconfirmed, i.e., the higher expectancy, the higher mPFC response to expectancy-violations.

A further observation concerns the correlation between BAS measures (Carver and White, 1994) and frontocentral negative component for positive unexpected omission. The observed trend indicates that this mediofrontal negativity-like component triggered by an unexpected positive event is stronger for people who are more responsive to rewards. Indeed, as compared with

shock-delivery, shock-omission should be perceived as a positively valenced event (Konorski et al., 1987). The influence of reward responsiveness to ERP components (such as ERN) between subjects that are highly sensitive to punishment or to rewards has already been reported (Kreussel et al., 2012; Boksem et al., 2006; 2008), but these differences mainly concerned high-BAS as more responsive to ERPs associated with reward omission. To our knowledge, the present study for the first time reports the presence of a relation between BAS responsiveness and mediofrontal negativities coding for positive unexpected omission of a predicted punishment. However, this correlation could reflect a relationship between reward responsiveness and either positive events or omission events. Although this latter possibility seems extremely unlikely, future studies could specifically aim at disentangling these alternative explanations.

Study 2 - When timing matters: mediofrontal negativities signal the timing of salient outcomes

Abstract

Recent models, such as the predicted response-outcome (PRO) model, proposed a comprehensive computational model of the medial prefrontal cortex (mPFC) and anterior cingulate cortex (ACC) as primarily involved in encoding the likelihood of outcomes and signalling unexpected events (prediction error). An interesting hypothesis concerns the prediction of expected outcomes unexpectedly shifted in time, thus signalling Timing Prediction Error (TPE). However, unexpectedness per se is not the only variable that modulates mPFC activity, as studies reported its sensitivity to the salience of outcomes. In the present study mediofrontal negative event-related potentials (ERP), coding for prediction error, were obtained from 48 participants performing a Pavlovian aversive conditioning task, during which conditioned aversive (thus, salient) and neutral stimuli were unexpectedly anticipated or delayed in time. Results partially supported the PRO-model, as mediofrontal ERP signals of prediction error were observed for outcomes occurring at unexpected timings, but were specific for behaviourally salient outcomes (shock), as compared to neutral outcomes.

Experimental Design

The aim of the present study is to test whether the unexpected delay or anticipation in time of salient (i.e., aversive) and neutral outcomes can differently trigger ACC/mPFC activity, as expressed by mediofrontal negativities.

To this purpose, electroencephalographic (EEG) signal and galvanic skin response (GSR) were recorded from 48 participants performing a Pavlovian aversive conditioning task. Visual conditioned stimuli (CS) were associated with aversive (i.e., shock-associated) or neutral outcomes, but probabilistically delivered at different timings: in 80% of trials the outcome was presented at a predicted timing; in 20% of trials the outcome was unexpectedly shifted in time (either anticipated or delayed).

If only the likelihood of events accounts for ACC/mPFC activity (as proposed by the PRO-model), then salient and neutral stimuli delivered at unexpected timings should be equally able to elicit mediofrontal negativities. Whereas, if behaviourally salient outcomes selectively trigger ACC/mPFC activity, these should be able to elicit a stronger mediofrontal negativity, as compared to neutral outcomes, when unexpectedly shifted in time.

Materials and Methods

Participants

Forty-eight volunteers (24 female; 8 left-handed; age mean = 23.62, sd = 2.39; education mean = 16, sd = 1.2) were recruited from the student population of the University of Bologna and enrolled for the study. The participants had no history of neurological diseases and normal or corrected-to-normal vision. All gave written informed consent before the beginning of the experiment. The study was conducted in accordance with institutional guidelines of the University of Bologna and was approved by the Department of Psychology's ethical committee.

Stimuli

A Pavlovian aversive conditioning paradigm was used for the task. Japanese Kanji symbols served as conditioned (CS+) and neutral stimuli (CS-). Mild electric shocks were used as unconditioned stimuli (US). Coloured frames surrounding the CSs served as visual outcomes. The visual outcomes were of two possible colours (blue or orange) and each colour was uniquely associated with either aversive (CS+), thus shock-associated, or neutral (CS-) outcome. Colours were counterbalanced across subjects. The stimuli were displayed at a viewing distance of 80 cm on a 17-inch colour monitor with a black background. A PC running Presentation software (Neurobehavioural Systems, Albany, CA) controlled stimulus presentation. All stimuli were balanced for luminance, complexity and colour saturation. A Digitimer Stimulator (Model DS7, Digitimer Ltd., Hertfordshire, United Kingdom) generated the electrical shocks, consisting of pulses of 200 ms. Two SU15N1 electrodes (SEI EMG, Padova), attached to subjects' left hand were used for delivering the mild shock. Shock intensity was adjusted for each participant before each block: stimulation intensity was initially set at 0.5 and gradually increased (1 mA increments) to a level perceived as "annoying, but not painful" by the participant. The mean shock intensity was 3.58 mA (sd =3.08 mA; min =0.32 mA; max =9.27 mA).

Task Description

Anticipation and *Delay* versions of the task were created, differing solely on the timing in which the outcomes were presented (Fig. 2.1). Participants had to complete only one version.

The task began displaying a black screen for a random interval (1-2 sec.), followed by a fixation cross. Participants were required to press a button when presented with the fixation cross, in order to start the trial. For each trial, a stimulus (CS+/CS-) was displayed for a variable amount of time (depending on the Anticipation/Delay condition) and then followed by the corresponding visual outcome, consisting in a coloured frame (850 ms). During Shock Trials (CS+) a shock (US) was delivered in the last 200 ms of visual outcome. During Neutral Trials (CS-), the visual outcome was not paired with anything. As the FRN occurs within about 400 ms after outcome onset, an interval of 600 ms between the onset of the visual outcome and US delivery allowed to avoid confounding effects due to the overlap between the shock delivery and the epoch of interest for the EEG analysis.

In the Anticipation Task the CSs were followed by the visual outcome after 2500 ms in 80% of trials (Expected Timing) and after 1000 ms in 20% of trials (Unexpected Timing). Consequently, in a lower number of trials the outcome was unexpectedly anticipated in time. In the Delay Task the CSs were followed by the corresponding outcome after 1000 ms in 80% of trials (Expected Timing) and after 2500 ms in 20% of trials (Unexpected Timing). Consequently, in a lower number of trials the outcome was unexpectedly delayed in time.

The task consisted of 858 trials, divided in 11 blocks. Each block began with a learning phase, during which only the Expected Timing condition was presented for the first 10 trials. The subsequent 60 trials followed the 80-20 partial reinforcement schedule explained earlier. The trials in the learning phase were not considered for the EEG analysis. Galvanic Skin Response

was analysed from 88 catch trials (total) consisting in CS+ stimuli followed by no US and larger inter trial intervals (5-6 sec.).

In order to assess explicit learning, at the end of each block participants were required to indicate which of the two stimuli presented in that block was associated to the shock.

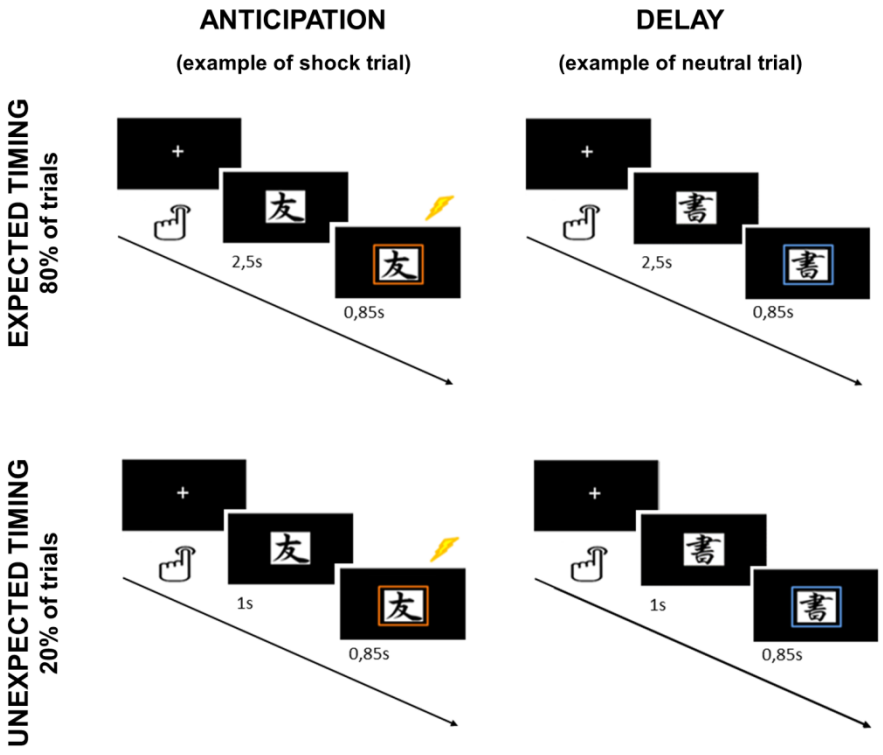

Figure 2.1 – Pavlovian aversive conditioning task
Participants were required to press a button to start a new trial and were subsequently presented with one of two possible conditioned stimuli (CS+/CS-). The stimuli were followed by a visual outcome indicating an imminent shock delivery (Shock Trials - CS+) or neutral outcome (Neutral Trials - CS-). Such outcomes occurred on 80% trials at an Expected Timing (up), and were shifted in time on 20% of trials (down). Two task versions were used between-subjects: Anticipation (left) and Delay (right), differing solely on the timing in which the outcomes were presented.

Procedure

Subjects were comfortably seated in a silent room and their position was centred relative to the screen. Electroencephalogram (EEG) and skin conductance response (SCR) were recorded continuously while participants completed the task and data were stored for off-line analysis. Participants were asked to remain as quiet and still as possible during task completion and keep their gaze strait towards the centre of the screen. After verifying that EEG and SCR were being properly recorded, the intensity of shock delivery was adjusted for each participant (see *Stimuli and Materials*). Written instructions for completing the task appeared on the computer screen. Participants were informed that their responses and actions had no effect on shock administration and were allowed to take short breaks between blocks.

EEG recording and analysis

The EEG signal was recorded using Ag/AgCl electrodes (Fast'n Easy Electrodes, Brain Products, Gilching, Germany) at 59 electrode sites and right mastoid. The left mastoid was used as reference site and a ground electrode was placed on subjects' right cheek. All electrodes were re-referenced offline to the average of both mastoids. Two electrodes placed above and below the left eye, as well as two electrodes placed in the outer canthi of both eyes were used to record vertical and horizontal electrooculogram (EOG). The recorded signals were amplified using a BrainAmp DC amplifier (Brain Products, Gilching, Germany)and recorded at a sampling rate of 500 Hz.

MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States) and EEGLAB 13.02.2b free toolbox (Delorme and Makeig, 2004) were used for offline analysis.

EEG data were re-referenced off-line to the average of both mastoids and filtered with a 1-30 Hz pass-band. Epochs of 600 ms before and after the onset of the visual outcome were extracted from the continuous EEG. Baseline-correction was performed using the average voltage in a 100 ms time window preceding the outcome onset. Epochs contaminated with large artifacts were identified and excluded whenever the voltage on an individual channel exceeded 400 to remove epochs with extremely large voltage fluctuations and whenever the joint probability of a trial exceeded five standard deviations to remove epochs with improbable data (Delorme et al., 2007). To correct remaining artifacts, the data were subjected to a temporal ICA (Jutten and Herault, 1991; Makeig et al., 1996) using the infomax algorithm (Bell and Sejnowski, 1995). The resulting component matrix was screened for independent components (ICs) representing stereotyped artifact activity, such as horizontal (saccades) and vertical (blinks) eye movements. This was done using a multistep correlational template-matching process as implemented in CORRMAP v1.02 (Viola et al., 2009). Topographies of ICs labeled as artifacts by the CORRMAP procedure were visually inspected and then calculated out of the data using inverse matrix multiplication. To increase signal-to-noise ratio for the FRN analyses {Debener, 2005}, frontocentral ICs corresponding to the FRN were extracted from the data. To this end, we first chose an IC that showed a typical frontocentral distribution from one participant, and then correlated all ICs of all participants with this template (correlations above r = 0.90). Mean extracted ICs showed typical frontocentral distributions in all conditions. These ICs were selected and back-projected into channel space using inverse matrix multiplication. All FRN

analyses were performed using these back-projected data. The FRN was quantified as the peak-to peak amplitude in the 200-350 ms time-window after outcome onset (Hajcak et al., 2006). To obtain the peak-to-peak amplitude, the difference between the minimum peak in the 275-350 ms time-window after outcome onset and the maximum peak in the 200-275 ms time-window after outcome onset was computed. All statistical analyses were performed with RStudio v0.98.1062 (Boston, MA).

Galvanic skin response recording and analysis

Ag/AgCl electrodes (TSD203 Model; Biopac Systems, Goleta, CA), filled with isotonic hyposaturated conductant were used for recording galvanic skin response (GSR). These were attached to participants' volar surface of the index and middle fingertip in their left hand (which did not require any motor movements during the task). A DC amplifier (Biopac GSR100) was used while recording the GSR signal. A gain factor of 5 μS/V and low-pass filter set at 10 Hz were used for recording the analog signal, which was then passed through a MP-150 digital converter at a 200Hz rate. The digital signal was then fed into AcqKnowledge 3.9 (Biopac Systems) and transformed into microsiemens for offline analysis. Data were analysed offline using custom-made MATLAB scripts (The MathWorks, Inc., Natick, Massachusetts, United States) and all statistical analyses were performed with RStudio v0.98.1062 (Boston, MA). All trials were recorded, however for the analyses only 88 catch trials in which no shock was delivered were considered, in order to exclude shock artifacts.

Skin Conductance Response (SCR) were extracted from the continuous signal and calculated for each trial as the peak-to-peak amplitude of the largest deflection during the 0.5-4.5 sec time window following stimulus onset (Schiller et al. 2008). The minimal response criterion was 0.02 μS and smaller responses were encoded as zero. Raw SCR scores were square root transformed to normalize the distributions and scaled to each subject's maximal US response to account for interindividual variability (Schiller et al., 2008). Three participants were excluded due to a malfunctioning of the SCR recorder instrument and four participants had non-measurable level of skin conductance. Thus, the SCR analysis included 41 participants.

Results

SCR results

SCR response during Shock (CS+) and Neutral (CS-) trials were compared to test for implicit acquisition of the Pavlovian contingencies.

A 2x2 repeated measure ANOVA was performed, with Task (Anticipation/Delay) as between-subjects independent variable, CS (CS+/CS-) as within-subjects independent variables and SCR as dependent variable. A marginally significant main effect of CS ($F(1, 39)= 3.59$; $p= 0.06$; part.$\eta^2= 0.09$) was found, with CS+ trials (mean= 0.1 μS; sd= 0.06 μS) presenting higher SCR levels relative to CS- trials (mean= 0.092 μS; sd= 0.05 μS) (Fig. 2.2). All other effects were not significant (ps >.54).

Even if these results were only marginally significant, participants showed a trend that supports an implicit learning of the association between the conditioned stimuli (CS+ or CS-) and their respective outcome (shock or neutral), as higher levels of arousal were found during CS+ trials as compared CS- trials.

Furthermore, explicit learning was assessed by asking the participant to indicate the stimulus associated with the shock, at the end of each block. The CS+ was correctly indicated 99,79% of times.

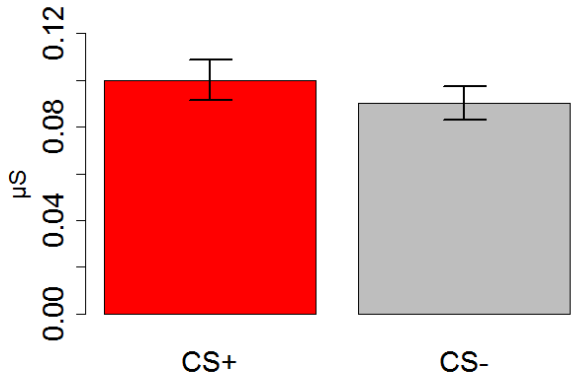

Figure 2.2 – Skin Conductance Response
SCR levels at stimulus presentation during Pavlovian aversive conditioning task. Bars represent standard error.

EEG results

The aim of EEG analysis was to contrast and compare the amplitude of mediofrontal negativities (FRN) elicited by shock and neutral outcomes delivered at expected or Unexpected Timings. Statistical analyses were performed on the data from the frontocentral electrode FCz, where FRN has been previously reported (Hajcak et al., 2006). A mixed-effect model was performed, using

Task (Anticipation/Delay) as between-subjects independent variable, Timing (Expected/Unexpected) and Salience (Shock/Neutral) as within-subjects independent variables. Peak-to-peak amplitude (see *EEG recording and analysis*) was used as dependent variable and calculated matching the number of trials in each condition. Results showed a significant main effect of Salience ($F(1, 46) = 18.88$; $p < .001$; part.$\eta^2 = .29$), qualified by a significant Task X Salience interaction ($F(1, 46) = 9.3$; $p < .001$; part.$\eta^2 = .17$) and a significant main effect of Timing ($F(1, 46) = 15.84$; $p < .001$; part.$\eta^2 = .26$), qualified by a significant Salience X Timing interaction ($F(1, 46) = 7.17$; $p = .01$; part.$\eta 2 = .13$) (Fig. 2.3 A-B). All other effects were not statistically significant (ps >.17).

Bonferroni-corrected post-hoc analysis on the Task X Salience interaction revealed an overall significant difference ($p < .0001$) between Shock Trials (mean = 3.62 μV; sd = 2.79 μV) and Neutral Trials (mean = 2.91; sd = 2.44) in the Delay Task, but not in the Anticipation Task ($p = 1$; Shock Trials mean = 2.46 μV, sd = 1.73 μV; Neutral Trials mean = 2.34 μV, sd = 1.61 μV).

Bonferroni-corrected post-hoc analysis on the Salience X Timing interaction revealed a significant difference (ps <.0001) between Shock Expected (mean = 2.77 μV; sd = 2.07 μV) and Shock Unexpected (mean = 3.30 μV; sd =2.65 μV), and between Neutral Unexpected (mean = 2.68; sd = 2.17 μV) and Shock Unexpected. No difference was found between Neutral Expected (mean =2.57 μV; sd =2.02 μV) and Neutral Unexpected ($p = 1$) and between Neutral Expected and Shock Expected ($p = .5$).

Critically, this analysis evidenced the presence of significantly higher ERP amplitude for salient (aversive) outcomes unexpectedly anticipated or delayed in time (Shock Unexpected), as

compared to both salient outcomes delivered at the expected timing (Shock Expected) and neutral outcomes unexpectedly anticipated or delayed in time (Neutral Unexpected) (Fig. 2.3 A-B). Thus, behaviourally salient outcomes unexpectedly shifted in time, selectively trigger stronger mediofrontal negativities.

Grandaverage scalp topographies showed a frontal activation for all Timing and Salience conditions, with a maximum activation for Unexpected Timing of shock outcomes (Fig. 2.3-C).

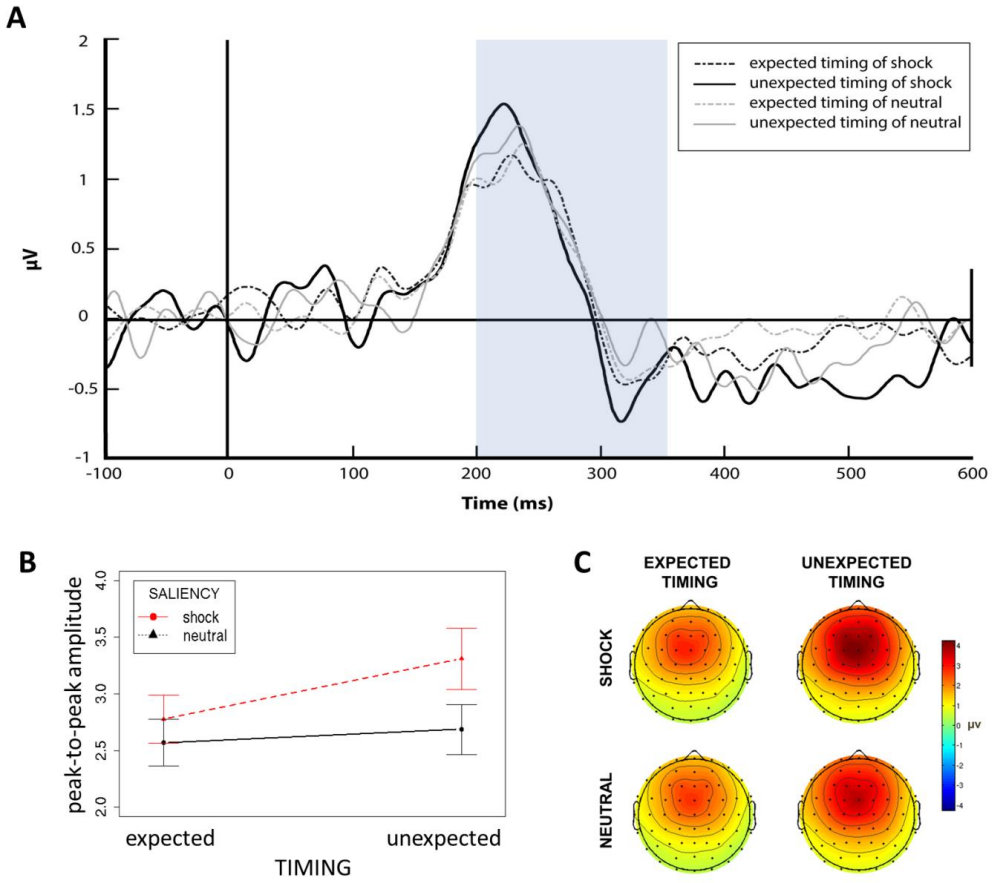

Figure 2.3 – Event Relates Potentials
Panel A shows the grandaverage ERP waveforms from channel FCz. The coloured band indicates the considered time-window (200-350ms). 0 represents the visual outcome onset; Panel B shows the average peak-to-peak amplitude in all Timing (Expected/Unxpected) and Salience (Shock/Neutral) conditions; Panel C shows the Scalp distribution of the peak-to-peak amplitude the 200-350ms time window after outcome offset.

Conclusion

The present study aimed at investigating whether expected salient (i.e., shock-associated) and neutral outcomes are able to elicit mediofrontal negativity when delivered at unexpected timings. Results showed that unexpected delays and anticipations in time of salient (i.e., shock-associated) outcomes selectively increase the amplitude of mediofrontal negativity, as compared to neutral outcomes.

These findings support the idea that mediofrontal negativity signal when salient, but not neutral, outcomes are likely to occur.

Analysis of both explicit and implicit measures of learning showed that participants correctly associated the stimuli (CS+/CS-) to their respective outcome (shock/neutral), thus confirming a successful manipulation of salience.

The present results constitutes a strong link between mediofrontal negative ERP components classically associated with prediction error (i.e., mediofrontal negativity) and the literature about midbrain dopaminergic signals coding information about the timing of outcome delivery (Schultz et al., 1997; Schultz and Dickinson, 2000; Hollerman and Schultz, 1998; Fiorillo et al., 2003; 2008).

Furthermore, it is in line withPrevious fMRI studies (Klein-Flügge et al., 2011; Forster and Brown, 2011) reporting activity in frontal and striatal areas (from which mediofrontal negativity is originated - Sambrook and Goslin, 2016; Hauser et al., 2014) correlated with prediction of the timing of outcome delivery (Klein-Flügge et al., 2011; Forster and Brown, 2011).

The results here presented partially support what is postulated by the PRO model (Alexander and Brown, 2011; 2014), as mediofrontal activation has been observed for outcomes occurring at unexpected times, even if the outcomes themselves are expected. However, these signals appear to be specific for salient events (Talmi et al., 2013; Lou et al 2015), which is in contrast with what predicted by the PRO model. Indeed, the PRO model postulates a dual role of ACC/mPFC in encoding the likelihood of outcomes - regardless of affective valence - and signalling all unexpected events (prediction error).

Nevertheless, in line with other studies (Talmi et al., 2013; Maier et al., 2012; 2013; Lou et al., 2015; Ganushchak et al., 2008; Hajcak et al., 2005), the current results showed how ACC/mPFC activation is sensitive to the salience of outcomes, as a higher amplitude of mediofrontal negativity was selectively observed for aversive salient outcomes occurring at unexpected timings, as compared to neutral outcomes similarly delivered at unexpected timings.

In conclusion, this study provides evidence that mediofrontal negativity is sensitive to the timing of outcome delivery of salient outcomes. This represents a step towards the link between midbrain dopaminergic activity following violations of the expected timing of outcome and mediofrontal ERP signals of prediction error.

Study 3 - Event-Related Brain Potential Signalling Unexpected Timing of Feedback: A Source Localization Analysis

Abstract

Mediofrontal event-related potential (ERP) components have been extensively reported following performance error or unexpected feedback, and are thought to reflect dopaminergic phasic changes conveyed to medial prefrontal cortex. Recent studies suggest a role for medial prefrontal cortex not only in signalling unexpected feedback but also unexpected timing of expected feedback. However, analyses of this aspect are still scarce. In this study, a task creating a condition of unexpected (anticipated) timing of a predicted feedback was designed and the associated ERP analysed. Both an aversive and a neutral feedback were considered. ERP analysis evidenced frontocentral ERP component elicited by unexpected, relative to expected, timing of feedback, more significant for aversive than neutral feedback. Cortical generators of this ERP component were estimated using sLORETA source localization method: some regions in the frontal and prefrontal cortex may be responsive to unexpected timing of feedback, to a larger extent in case of aversive than neutral feedback. Obtained results are discussed with respect to findings of previous works. The present study can contribute to inform current theories of performance monitoring in frontal/prefrontal cortex and to drive future investigation.

Experimental Design

The brain continuously makes predictions and compares outcomes or inputs with those predictions. Evaluation of the expectations against observed events is crucial to identify

discrepancies that may drive behaviour in an uncertain and changing environment. Regions in the medial prefrontal cortex (mPFC), especially anterior cingulate cortex (ACC), have repeatedly emerged as implicated in expectation evaluation and subsequent behavioural adjustments (Rushworth et al., 2004). In particular, event-related potential (ERP) studies, combined with cortical source localization techniques, reported the presence of frontocentral ERP components that originate in mPFC/ACC (for a review see, Walsh and Anderson, 2002), such as the feedback related negativity (FRN), usually observed after negative or unexpected feedback (Hajcak et al., 2006; Alexander and Brown, 2011; Ferdinand et al., 2012; Talmi et al., 2013; Garofalo et al., 2014). As mPFC/ACC receive a preponderance of projection from the midbrain dopamine system (Willians and Goldman-Rakic, 1993), these ERP components may reflect phasic changes in activity of dopaminergic neurons, that have been demonstrated to code discrepancy between predicted and actual outcome (Fiorillo eta l., 2003).

Recent fMRI studies (Forster et al., 2011) reported that mPFC/ACC not only signal violations of feedback expectancy but also violations of the expected timing of feedback (i.e., a correctly predicted feedback occurring at an unexpected time); this may reflect sensitivity to outcome timing observed in dopamine neurons (Hollerman and Schultz, 1998). However, to date, ERP components of timing prediction error have never been directly investigated, and information on how and to what extent they may be related to mPFC/ACC activity are lacking. Aim of the present study is to investigate whether predicted outcomes that occur at unexpected rather than expected timing produce differences in ERP, and to identify the cortical generators of these differences by means of an electrophysiological source localization technique (sLORETA). In particular, the anticipated timing condition was investigated; moreover, to assess whether

feedback valence may have an effect on timing prediction error, an aversive and a neutral feedback have been considered.

Materials and Methods

Participants and Procedure

Twenty healthy volunteers (9 females, mean age = 23.55, sd= 2.32) participated in the study. All participants gave informed consent and the study was approved by the institutional ethical committee. A Pavlovian aversive condition task was used. Each participant was presented with two visual stimuli (Japanese kanji), paired either with an aversive (shock-associated, CS+ trials) visual feedback or with a neutral (CS- trials) visual feedback, consisting in a coloured frame (red or blue, respectively), appearing around the image for 850 ms (Fig. 2.1). The aversive visual feedback was followed by a shock delivery during the last 200 ms, while during the neutral visual feedback presentation nothing happened until its termination. The shock intensity was individually set before the task to a level the participant indicated as "uncomfortable, but not painful". Each subject performed 350 CS+ trials and 350 CS- trials; in both conditions, the feedback was presented 2.5 s after stimulus (kanji) presentation on 80% of trials (*Expected Timing of feedback,* ET), while it was unexpectedly anticipated in 20% of trials and presented 1 sec after stimulus presentation (*Unexpected Timing of feedback,* UT) (Fig. 2.1).

EEG Recording and Analysis

During the task, electroencephalographic (EEG) signal was recorded from 59 electrode sites, referenced to left mastoid. Vertical and horizontal electro-oculogram (EOG) was recorded too. EEG signals were digitized at a sampling rate of 500 Hz, offline re-referenced to the average of both mastoids, filtered between 1-30 Hz and analysed by using MATLAB (MathWorks ®) and EEGLAB 11.0.5.4b free toolbox (Delorme and Makeig, 2004). Epochs of 100 ms before and 500 ms after feedback presentation were extracted from the continuous EEG (epochs did not include shock delivery). The average voltage during the 100 ms preceding feedback was used as baseline. Epochs presenting large artifacts were identified and excluded (mean percentage of excluded epochs = 3%) and EOG artifacts were corrected using a regression approach. Epochs were then averaged separately for each participant and each condition (aversiveET, aversiveUT, neutralET, neutralUT), to obtain the ERP. Previous studies extracted FRN component from the ERP at frontocentral sites (e.g., FCz) by computing the peak-to-peak difference in a time window included between 200 ms and 350 ms following feedback presentation [3],[5],[7]. Accordingly, in the present study, the presence of FRN-like components due to unexpected timing relative to expected timing condition, was assessed by computing the peak-to-peak amplitude recorded on FCz electrode in the window 220-340 ms after feedback presentation; based on grand average visual inspection, this temporal window included the peak-to-peak deflection.

Cortical Source Localization and Statistical Analysis

Standardized Low Resolution Electromagnetic Tomography free software (sLORETA, Key Institute for Brain-Mind Research, Zurich) was used to compute the cortical three-dimensional distribution from electrode scalp field. sLORETA method estimates current density magnitude at each of 6239 voxels in the gray matter of a reference brain (MNI 152 template, Mazziotta et al., 2001). This method finds a particular solution addressing the non-unique EEG inverse problem by assuming similar activation of neighboring neuronal sources, followed by an appropriate standardization of the current density, producing images of electrical neuronal activity without localization bias (Pascual-Marqui, 2001; Cannon, 2012), as validated by several simultaneous EEG/fMRI studies (Cannon, 2012). In this study, sLORETA was used to analyze source localization differences, due to unexpected compared to expected timing, in the temporal window 220 ms -340 ms. To this aim, for each subject and condition, the averaged ERP data within this window were directly imported and processed by sLORETA. Voxel-by-voxel paired t-test statistical analyses were performed using the built-in statistical non-parametric mapping: this method (Nichols and Holmes, 2002) uses randomization to estimate the exact significant threshold, regardless of any assumption of Gaussianity, and corrected for multiple testing (i.e., in this study, for the collection of tests performed for all voxels). Threshold at a significance level of $p<0.05$ was used.

Results

Fig. 3.2 (left panel) displays grand average ERP waveforms on FCz electrode in the 4 conditions (aversiveET, aversiveUT, neutralET, neutralUT). Mean peak-to-peak amplitudes computed on

the window 220 ms - 340 ms are displayed in the right panel. Paired t-test on peak-to-peak amplitudes revealed a significant difference only between aversiveUT and aversiveET (p=0.0053).

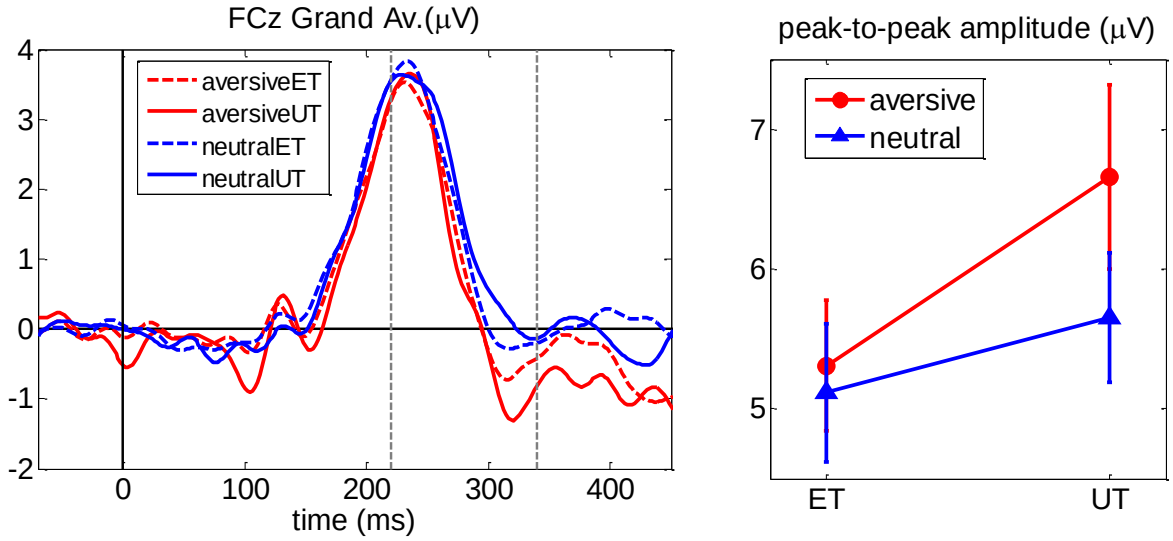

Figure 3.2
Left: GrandAverage ERP waveforms on electrode FCz in the four conditions. Black vertical line at time = 0 ms represents feedback onset. Gray dashed lines denote the time window for peak-to-peak computation. Right: Mean peak-to-peak amplitudes computed on the selected window. t-test: neutralUT vs neutralET, p=0.1882; aversiveET vs neutralET, p=0.3774; aversiveUT vs neutralUT, p=0.056.

Fig. 3.3 shows the statistical maps comparing sLORETA solutions between aversiveUT and aversiveET, revealing a significantly stronger activation of several areas of the brain in case of unexpected timing of the aversive feedback. The center of gravity for statistically stronger activated areas appears to be located in the fronto-temporal zones. In order to quantify this, voxels at the higher extreme of the map scale (i.e., with t-value higher than 80% of maximal t-value) were identified and clustered on the basis of the Brodmann Areas (BA) they belonged to.

Results are reported in Table 3.1, listing the first ten clusters of statistically stronger activation: they mainly belonged to middle/inferior frontal structures and middle/superior temporal structures.

Despite the absence of a statistically significant difference in the ERP between neutralUT and neutralET, cortical source analysis disclosed differences in source localization between the two conditions, as evidenced by the statistical maps in Fig. 3.4. Graphically, patterns resemble those obtained for aversive feedback (Fig. 3.3). For a quantitative comparison with aversive condition, Table 3.2 extracts the areas of statistically higher activation in unexpected compared to expected timing in case of neutral feedback, using the same procedure as Table 3.1; parieto-occipital areas appear to be involved in this case, besides fronto-temporal areas.

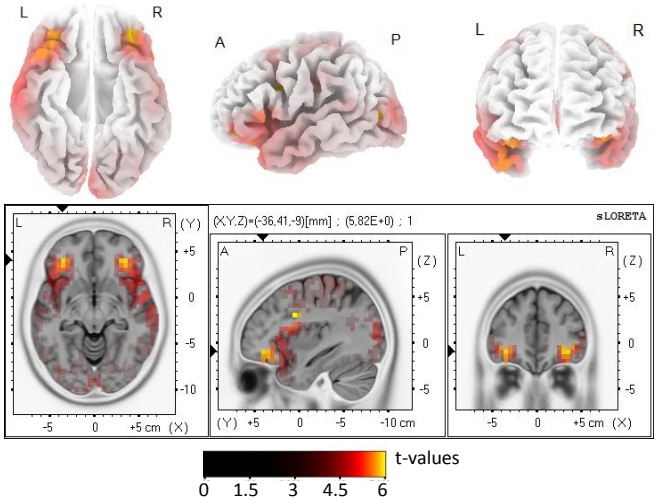

Figure 3.3
Grand average of sLORETA statistical maps comparing aversiveUT and aversiveET over the examined time window (220 ms-340 ms). Upper 3D representations: bottom, left, frontal view. Lower 2D representations: axial, sagittal and coronar slice. Colour indicates voxel-by-voxel t-value of the statistical comparison ($p<0.05$ for t-values above 2.69).

For completeness, sLORETA analyses were conducted to compare aversiveET with neutralET and aversiveUT with neutralUT, too. sLORETA comparison between aversiveET and neutralET (Fig. 3.5) revealed statistically significant higher activation in a few areas, still located in the frontal lobe but in regions complementary compared to those involved when contrasting conditions UT-ET; indeed, statistically significant voxels were in BA 10, 11, 46 in middle/superior frontal gyrus (Fig. 3.5). sLORETA comparison between aversiveUT and neutralUT did not produce statistically significant results. However, it is interesting to note that when contrasting these two conditions, differences in source localization (although not reaching significance) appear also in somatosensory areas 2 and 3 (Fig. 3.6).

number of voxel	*BA*	*Structures*	*Peak t-value*	*Mean t-value*
32	11	Middle frontal gyrus	5.84	5.43
6	9	Inferior frontal gyrus	5.93	5.25
4	6	Middle frontal gyrus Precentral gyrus	5.68	5.14
39	19	Middle temporal gyrus; Middle/superior occipital gyrus	5.93	5.09
17	45	Inferior frontal gyrus	5.37	5.09
9	39	Middle temporal gyrus	5.74	5.09
111	47	Middle/inferior frontal gyrus	5.66	5.08
3	10	Inferior frontal gyrus	5.16	5.05
68	38	Superior temporal gyrus	5.55	5.03
12	44	Inferior frontal gyrus Precentral gyrus	5.33	5.01

Table 3.1
Voxel clusters of statistcally higher activation in AVERSIVEut vs AVERSIVEet

Conclusion

In this pilot study, the aim was to investigate ERP components associated with feedbacks that were delivered at unexpected (anticipated) time, even if the feedbacks themselves were expected, and to estimate their cortical sources. Obtained results deserve some major comments.

1) ERP analysis evidenced the presence of a mediofrontal FRN-like components, suggesting that these components may reflect not only feedback prediction error, as reported by the vast majority of studies (Walsh and Anderson, 2002; Alexander and Brown, 2011; Ferdinand et al., 2012; Talmi et al., 2013; Garofalo et al., 2014), but also temporal prediction error. Furthermore, analysis at ERP level highlighted a salience effect, since shock associated feedback at unexpected time generated a stronger FRN-like component compared to neutral feedback (Fig. 3.2, right panel). This ERP result seems in line with previous work which reported evidence for ERP FRN-like component mainly signaling salience prediction errors (Talmi et al., 2013).

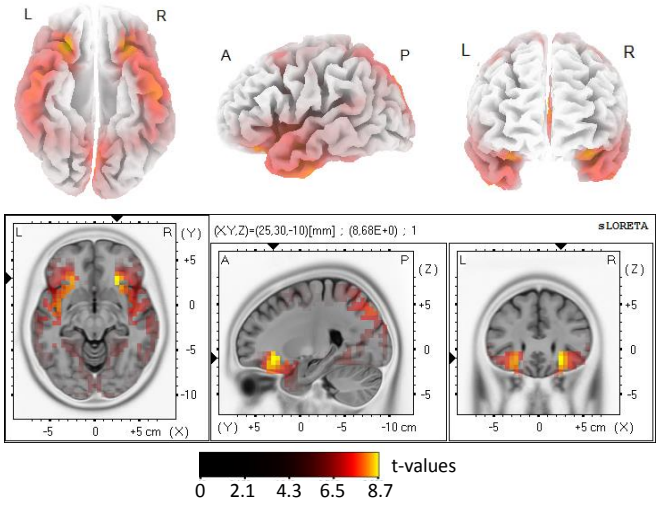

Figure 3.4
Grand average of sLORETA statistical maps comparing neutralUT and neutralET over the examined time window (220 ms-340 ms). $p<0.05$ for t-values above 2.86.

number of voxel	*BA*	*Structures*	*Peak t-value*	*Mean t-value*
87	19	Cuneus (occipital lobe) Precuneus (parietal lobe)	8.24	7.63
34	11	Middle/inferior frontal gyrus	8.36	7.61
115	47	Inferior frontal gyrus	8.68	7.60
97	18	Cuneus (occipital lobe)	8.16	7.53
76	13	Insula (sub-lobar)	8.37	7.51
48	31	Precuneus (occipital & parietal lobe)	7.91	7.50
88	20	Inferior temporal gyrus Fusiform gyrus (temporal lobe)	7.96	7.40
9	23	Limbic lobe (posterior cingulate) Cuneus (occipital lobe)	7.74	7.40
82	21	Middle/superior temporal gyrus	8.00	7.36
9	17	Cuneus (occipital lobe)	7.66	7.35

Table 3.2
Voxel clusters of statistcally higher activation in neutralut vs neutralet

2) sLORETA analysis provided interpretation, in terms of cortical source activation, of ERP analysis results. First, the significant effect induced by unexpected timing of aversive feedback on mediofrontal ERP, was explained by significantly higher brain electrical activity in several prefrontal and frontal areas (BA 11, 9, 6, 45, 47, 10, 44 Table 3.1). Second, at variance with the absence of a significant effect on scalp mediofrontal ERP, cortical source analysis revealed some cortical prefrontal areas (BA 11, 47) having statistically stronger activation for unexpected as compared to expected timing, even in case of neutral feedback. However, in neutral feedback condition, there was a minor involvement of prefrontal areas and a major engagement of parieto-occipital areas compared to aversive feedback condition. This may explain the absence of a statistically significant effect at mediofrontal scalp sites in case of neutral feedback.

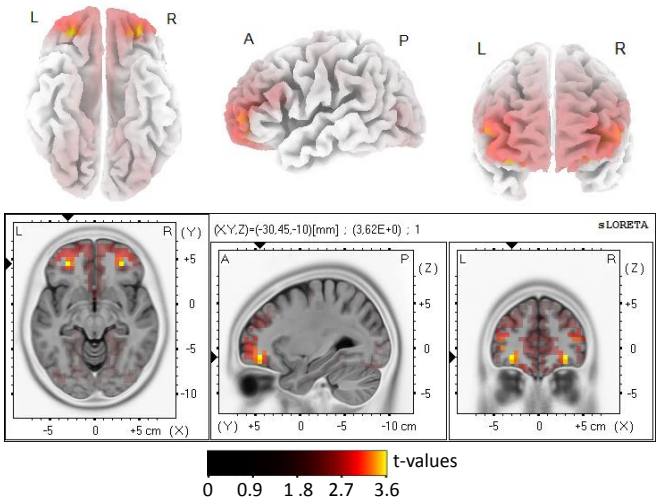

Figure 3.5
Grand average of sLORETA statistical maps comparing aversiveET and neutralET over the examined time window (220 ms-340 ms). $p<0.05$ for t-values above 3.

3) Despite higher activation in some prefrontal areas in unexpected timing conditions, sLORETA analysis failed to reveal a statistically significant role for the ACC in the examined task. Although this may be surprising as ACC has emerged as a primary locus coding unexpected feedback, our result is in line with a recent fMRI study (Forster et al., 2011) reporting a significant activation of ACC only for unexpected delayed feedback and not for unexpected anticipated feedback, which is the condition examined in our study. However, prefrontal and mesiotemporal areas that have emerged in this study as responsive to anticipated feedback may represent other targets of midbrain dopamine projections conveying outcome valance and timing information. In future, tasks involving both an anticipated timing of feedback and a delayed timing of feedback can be designed and analysed to further support the inferences of the present and previous studies.

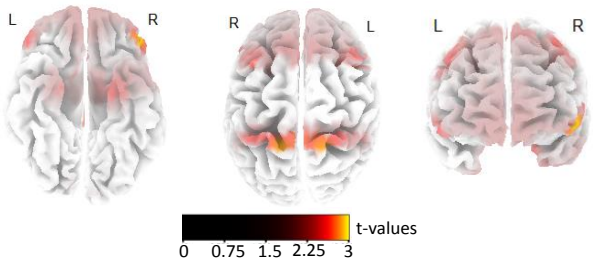

Figure 3.6
Grand average of sLORETA statistical maps (3D bottom, top, front view) comparing aversiveUT and neutralUT over the examined time window (220-340 ms). $p<0.05$ for t-values above 3.93. For sake of space, only 3D views are reported.

4) Secondary to the previous primary results (which targeted the main goal of our study), further interesting cues emerged from source localization analysis contrasting aversive and neutral feedbacks in each of the two timing conditions. First, different activation of superior prefrontal areas (figure 3.5 and figure 3.6, even if in figure 3.6 it did not reach significance) may represent coding of feedback valence. These specific regions, indeed, did not appear to be differently activated when contrasting the same feedback across the two timing conditions. Finally, in unexpected timing, stronger activation appears in somatosensory areas, although below significance threshold. Such activation can be related to shock mental imagery (Schmidt et al., 2014) induced by the anticipated aversive feedback.

In conclusion, results of these preliminary study represent a further step towards a possible link between mediofrontal ERP signals of timing prediction error, frontal/prefrontal cortex activity, and dopaminergic changes following violations of the expected timing of feedback.

It is important to underlie the main limitation of this preliminary study and its future developments. In particular, the main descriptive nature of this study, and the lack, at this level of analysis, of clear insights into the underlying operating mechanisms. Indeed, this work is intended to report only the results of a preliminary step towards more thorough and interpretative analyses that is going to perform using neural network modeling (a field of wide expertise of some of the authors, e.g. see Ursino et al., 2007) joined with methods of cortical connectivity estimation (Ursino et al., 2012) on the entire experimental dataset used for Study 2.

Study 4 - Altered fronto-striatal reward processing in Parkinson's Disease associated psychosis

Abstract

Psychotic symptoms frequently occur in Parkinson's Disease, but their pathophysiology is poorly understood. According to cognitive neuropsychiatric theory, and the National Institute of Health RDoc programme, the pathophysiological basis of psychiatric symptoms may be better understood in terms of dysfunction of underlying domains of neurocognition, rather than simply according to diagnosis. Abnormal frontal striatal reinforcement learning signalling, associated with dysregulated dopamine, has been proposed as a key process contributing to the pathogenesis of psychosis. This theory has received empirical support in the study of schizophrenia spectrum disorders and preclinical models of psychosis, but has not been tested in psychosis associated with a degenerative neurological condition such as Parkinson's Disease. The aim was to investigate brain responses associated with reward feedback and prediction error signalling during reinforcement learning in Parkinson's Disease associated psychosis. An instrumental reinforcement learning task was performed by three groups of subjects during fMRI scanning: patients with Parkinson's Disease with a history of psychotic symptoms (n=12), patients with Parkinson's Disease without a history of psychotic symptoms (n=17), and healthy controls (n=24). Results confirmed the presence of altered activation in the medial-prefrontal cortex and the ventral striatum in PD patients with psychosis during reward feedback (contrast of winning £1 versus neutral feedback). Also, lateral parietal deficits in reward prediction error signalling were found in Parkinson's Disease independent of whether they had psychotic symptoms. These results extend the link between psychotic symptoms and abnormalities in fronto-striatal reward

processing beyond schizophrenia spectrum conditions into psychosis as manifest in Parkinson's Disease.

Experimental Design

The aim of the present study was to investigate the neural correlates of reward processing in a population of patients with a diagnosis of Parkinson Disease (PD) with and without psychotic symptoms and a matched group of healthy Controls. In particular, three stages of reward-processing were investigated: Reward Anticipation, Reward Feedback and Prediction Error.

The hypothesis is that PD patients with psychosis should present differences in the neural activations associated with reward processing relative to both healthy controls and PD patients without psychosis.

Materials and Methods

Participants

The final sample was composed of 53 participants: 17 patients with a diagnosis of Parkinson Disease without any psychotic symptoms (PD); 12 patients with a diagnosis of Parkinson Disease and a history of current or previous psychotic symptoms (lifetime CAARMS scoring equal or over 3 in global and frequency scales) (PD psychosis); 24 healthy volunteers, with no history of neurological, psychiatric or medical disorders (Controls). The initial sample was composed of 66 participants. Ten participants (8 PD, 1 PD psychosis, 1 Control) were excluded

due to technical problems occurred to the task (responses not recorded) or to the fRMI scanning. Three additional subjects (2 PD, 1 Control) were excluded for low cognitive abilities that interfeared with the learning task. Patients were recruited via the PD research clinic at the John van Geest Centre for Brain Repair (BRC). All fulfilled the Queen Square Brain Bank Criteria for idiopathic PD (Gibb and Lees, 1988) and remained on their usual medications during testing. Each patient's dopaminergic drug regime was converted to an equivalent L-dopa dose (Tomlinson et al., 2010). Patients with dementia were excluded (operationalized as a mini-mental state score less than 24).Descriptive statistics of the sample are reported in Table 4.1. All subjects had normal or corrected-to-normal acuity and were without any contraindications for MRI scanning. The research was approved by the Cambridgeshire 2 NHS Research Ethics Committee; all participants provided informed consent.

Task

During the fMRI scan, subjects performed a computerized instrumental learning task based on previous studies (O'Doherty et al. 2004; Pessiglione et al., 2006; Seymour et al., 2007; Murray et al., 2008; Bernacer et al., 2013). During the task participants were presented with two fractal images and required to choose one of them, with a button press, in order to maximise their pay-offs. Each choice was followed by a visual feedback indicating the associated outcome. There were three possible trials: during Reward Trials one choice was associated with a £1 win in 80% of trials and to neutral feedback in 20% of trials (High-Likelihood - HL), whereas the other choice was associated with a neutral outcome in 80% of trials and to a £1 win in 20% of trials (Low-Likelihood - LL); during Bivalent Trials each choice was associated with a 50% chance of

Characteristics	Control	PD	PD psychosis	Test	p
Demographics					
Participants, n	24	17	12		
Age, mean (SD) yr	61.91 (5.83)	63.29 (9.94)	60.83 (6.6)	F(2, 50) =0.39	.68
Gender, M/F (% male)	10/14 (43)	10/7 (56)	6/6 (50)	X^2 (2) =1.17	.55
Handedness R/L (% right)	22/1 (96)	15/3 (83)	11/1 (92)	X^2 (2) =0.15	.92
White-British, n (%)	22 (91.66)	17 (100)	12 (100)	X^2 (2) =2.51	.28
Educational qualifications					
No qualifications (%)	3 (12.5)	1 (5.88)	0 (0)	X^2 (2) =1.89	.38
16/18 years old qualifications (GSCSEs , A-level or equiv.) (%)	10 (41.66)	6 (35.29)	6 (50)	X^2 (2) =0.62	.73
Degrees, and advanced vocational qualifications (%)	11 (48.83)	10 (58.82)	6 (50)	X^2 (2) =0.67	.71
Cognition and IQ					
MMSE- Total , mean (SD)	30.30 (3.08)	28.94 (1.59)	28.41 (1.37)	F(2, 49) =3.07	.07
IQ, mean (SD)	102.04 (11.92)	89.64 (14.59)	86.54 (13.45)	F(2, 49) =7.1	.002
			Controls vs all PD p<.01; PD vs PD psychosis p>.1		
Parkinson severity					
Hoehn and Yahr stage, % 1/2/3/4/5	-	59/35/0/0/6	60/10/30/0/0	X^2 (1) =0.5	.72
Medications					
Levodopa Equivalent Dosage, mean (SD)	-	614.81 (503.42)	714.03 (531.42)	T(22)=0.5	.61
Levodopa therapy, n (% yes)		13 (76.47)	9 (75)	X^2 (1) =.00	.9
DA agonist, n (% yes)		8 (47.05)	10 (83.33)	X^2 (1) =2.54	.11
Antidepressants, n (% yes)	-	7 (41.17)	1 (8.33)	X^2 (1) =2.33	.12
Anxiolitics, n (% yes)	-	3 (17.64)	3 (25)	X^2 (1) =.009	.92
Psychopathology					
BDI Total score, mean (SD)	3.08 (3.69)	8.88 (4.94)	12.66 (7.83)	F(2,50) =14.73	<.001
			Controls vs all PD p<.003; PD vs PD psychosis p>.1		
CAARMS score equal or over 3, global rating scales					
Unusual Thought Content	-	-	-	-	-
Non Bizarre Ideas , n (%)	-	-	3 (25)	-	-
Perceptual Abnorm. , n (%)	-	-	9 (75)	-	-
Disorganized Speech , n (%)	-	-	2 (16.66)	-	-
Aggression/Dangerous beh.	-	-	-	-	-
Suicidality and Self-harm	-	-	-	-	-

Table 4.1 – Descriptive statistics
Characteristics of the whole sample separated into the three groups: Controls, Parkinson Disease and Parkinson Disease with psychosis. The inclusion criterion for the psychosis group is Lifetime CAARMS (Comprehensive Assessment of At Risk Mental States scale, which measures psychotic symptoms) scoring equal or over 3 in global and frequency scales. BDI, Beck Depression Inventory.

either losing or winning £1; during Neutral Trials each choice was associated with a 80%/20% chance of receiving two kinds of neutral feedbacks (Fig. 4.1). Each trial type was repeated 30 times in a random sequence, for a total of 90 trials. The order and position of the pictures presented were counterbalanced across trials of the same kind. To win money, the participants had to learn, by trial and error, to select the stimulus that was more likely to produce a reward. All participants were informed that the total amount won during the task would be paid to them at the end of the experimental session.

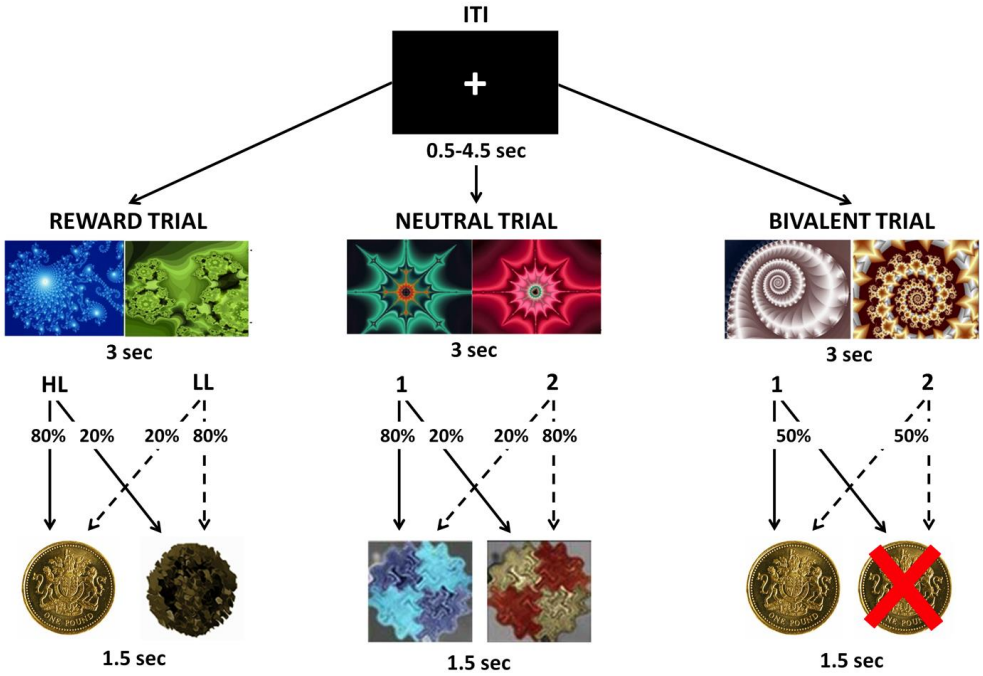

Figure 4.1 – Visual representation of the instrumental conditioning task
Participants were presented with two fractal images and required to choose one of them in order to maximise their pay-offs. Each choice was followed by a visual feedback indicating the associated outcome. During Reward Trials one choice was associated with an 80% chance to win 1£ (High-Likelihood - HL) and the other choice was associated with a 20% chance to win 1£ (Low-Likelihood - LL). During Bivalent Trials each choice was associated with a 50% chance of either losing or winning £1. During Neutral Trials each choice was associated with an 80%/20% chance of receiving two kinds of neutral feedbacks.

Rating scales

Before scanning, participants underwent a general interview and clinical assessment using the Comprehensive Assessment of At-Risk Mental States (CAARMS) (Yung et al., 2005). The Beck Depression Inventory (BDI) (Beck and Steer, 1996) was used to assess depressive symptoms during last 2 weeks. IQ was estimated using the Culture Fair test (Cattell et al., 1973). Hoehn and Yahr scale (Hoehn and Yahr, 1967) was used to assess the stage of Parkinson's Disease in patients.

The Levodopa Equivalent Dosage (LED) was calculated according to Tomlinson et al. (2010).

Computational model

A standard reinforcement learning algorithm (Q-learning) was used to fit each subject's sequence of choices (Pessiglione et al., 2006; Seymour et al., 2007; Murray et al., 2008). For each pair of stimuli A and B, the model estimates the expected values of choosing A (Qa) and B (Qb), on the basis of individual sequences of choices and outcomes. This value (Q) reflects the expected reward obtained by taking that particular action. Q values were set at zero before learning, and after every trial (t>0) the value of the chosen stimulus was updated according to the following rule. For example, for choosing A:

$$Qa(t+1)=Qa(t) + \alpha*\delta(t)$$

The prediction error was:

$$\delta(t)=R(t)-Qa(t)$$

where R(t) was defined as the reinforcement obtained as an outcome of choosing A at trial t. In other words, the prediction error δ(t) was the difference between the expected outcome (Q(t)) and the actual outcome (R(t)). The reinforcement magnitude (R) was +1 for reward feedback, 0 for neutral feedback and -1 for loss feedback. Given the Q values, the associated probability of selecting each action was estimated by implementing the Softmax Rule. For example, for choosing A:

$$Pa(t)= (expQa(t)/\beta)/(expQa(t)/\beta+expQb(t)/\beta)$$

The Softmax Rule is a standard stochastic decision rule that calculates the probability of taking one of a set of actions according to their associated values. The constants α (learning rate) and β (temperature, i.e. the balance between exploration and exploitation) were adjusted to maximize the likelihood of the actual choices under the model. The learning model was used to obtain a statistical regressor corresponding to the modelled outcome prediction error, to be used in the analysis of the imaging data. The constants α and β were individually estimated. The accuracy of fit was assessed by using negative log-likelihood, which can be summed across trials, sessions and subjects.

fMRI data acquisition and analysis

A Siemens Magnetom Trio Tim operating at 3 T was used to collect imaging data. Gradient-echo echo planar T2*-weighted images depicting BOLD contrast were acquired from 35 non-contiguous oblique axial planes to minimize signal drop-out in ventral regions.

TR was 1620ms; echo time was 30ms; flip angle was 65; in-plane resolution was 3.0x3.0, matrix size was 64x64; field of view was 192x192 mm, bandwidth 2442 HZ/Px. A total of 550 volumes per subject were acquired (27 slices each of 2 mm thickness, interslice gap 1mm). The first 5 volumes were discarded to allow for T1 equilibration effects.

The data were analysed in SPM 12 software (Wellcome Department of Cognitive Neurology, London, UK). Images were realigned, spatially normalized to a standard template and spatially smoothed with a Gaussian kernel (8mm at full-width half-maximum). A high pass filter was applied (128 Hz).

Explanatory variables (EVs) used were as follows: onset time of the bivalent cues; onset time of the neutral cues; onset time of the reward cues; onset time of loss feedback during Bivalent Trials; onset time of win feedback during Bivalent Trials; onset time of neutral feedback 1 during neutral trial; onset time of neutral feedback 2 during Neutral Trials; onset time of neutral feedback during Reward Trials; onset time of win feedback during Reward Trials. Parametric modulators of expected incentive value of the chosen cue, as estimated by the Q learning computational model, were used for all onset times of cue. Parametric modulators of prediction error (estimated by the Q learning computational model) were used for all onset times of feedback. The model constituted of 18 EVs in total. All regressors were modelled as 0s events and convolved with a canonical hemodynamic response function. Three further motion parameters were used as additional regressors to control for any movement artefacts along the three axes. All regressors of interest were convolved with a canonical hemodynamic response function with a temporal derivative.

The contrasts of interest were computed at the individual subject level and then taken to a group level for statistical analysis. There were two main contrast of interest: Reward Feedback and Prediction Error. For the Reward Feedback, a contrast between the presentation of win feedbacks (during Reward Trials) minus the presentation of neutral feedbacks (during Neutral Trials) was performed. The rationale here is that the two conditions differed in the actual monetary value (respectively, £1 and £0). To estimate Prediction Error, a contrast between the presentation of win feedbacks during bivalent minus Reward Trials was performed. In this case, the two conditions shared the same value (£1), but differed in the expected value (£1 during Reward Trials and 0£ during Bivalent Trials). To measure Reward Anticipation, the parametric modulator of the expected value of the chosen cue (estimated by the Q learning model) during Reward Trials was evaluated. The contrast of interest was PD versus PD psychosis. Brain activation in healthy Controls and Controls versus all PD patients are also reported. All analysis consisted of whole-brain t-contrasts, initial cluster-forming threshold at p=.01, Family-Wise Error (FWE) corrected at p=.05.

Results

Movement artifacts

To ensure that the three groups did not differ in the amount of movements during the fMRI scanning session, the averaged time series of translations and rotations along the three axes (x, y, z), as estimated during the realignment phase of pre-processing, were compared across groups.

Two separate mixed-effects models were used, with Group (Controls/PD/PD psychosis) as in depended variable and translations (mm) or rotations (degrees) ad dependent variable.

No group differences were found both in the amount of translations ($p = .35$; part. $\mu^2 = .04$) or rotations ($p = .32$; part. $\mu^2 = .04$).

Behavioural results

To ensure successful learning of the contingencies across the three groups, participant's performance during the task was analysed.

Three separate 3x2 mixed-effects models were performed for each trial condition (Reward/Neutral/Bivalent), using the total number of responses as the dependent variable (Fig. 4.2-A). Group (Controls/PD/PD psychosis) and Choice (High/Low-Likelihood or 1/2) were used as the independent variables. Results from Reward Trials showed a significant main effect of Choice ($F(1, 50) = 83.22$; two-tailed $p < .001$; part. $\mu^2 = .62$), evidencing a higher number of High-Likelihood over Low-Likelihood choices, and a Group X Choice interaction $F(2, 50) = 4,42$; two-tailed $p = .02$; part. $\mu^2 = .15$). Bonferroni-corrected post-hoc tests revealed a significant difference only between Controls and PD psychosis for both High-Likelihood ($p < .01$) and Low-Likelihood ($p < .01$) choices. On average, High-Likelihood choices were performed more by Controls (mean = 25.62, sd = 5.21) than PD psychosis (mean =19.25, sd =8.31); whereas Low-Likelihood choices were performed less by Controls (mean = 4.12, sd = 5.31) than PD psychosis (mean =10.58, sd =8.36).

Results from Neutral and Bivalent Trials showed no statistically significant effect (ps > .1).

A further 3x3 mixed-effects model was performed using the reaction times as the dependent variable (Fig. 4.2-C). Group (Controls/PD/PD psychosis) and Trial type (Reward/Neutral/Bivalent) were used as the independent variables. Results showed a significant main effect of Trial type ($F(1, 87) = 5.09$; two-tailed $p = .007$; part. $\mu^2 = .03$). Bonferroni-corrected post-hoc tests revealed a significant difference only between Reward and Neutral Trials (p=.006). All other effects were not significant (ps >.1).

Overall, these results indicate that all participants, irrespective of the group to which they belonged, showed contingencies learning during Reward Trials, expressed both as a preference for the High-Likelihood (HL) choice (i.e., more frequently associated with a reward – 80%) over the Low-Likelihood (LL) choice (i.e., less frequently associated with a reward – 20%), and as faster reaction times when choosing during such trials, as compared with Bivalent and Neutral Trials. Nevertheless, Controls presented a significantly higher ability to distinguish between choices with higher and lower chance of rewards, as compared with PD patients with psychosis. Such a difference was subsequently directly tested by creating an index of the ability to learn from rewards (reward learning), obtained by subtracting for each subject the number of LL responses from the number of HL responses (Fig. 4.2-B). A mixed-effects model was performed using the reward learning index as the dependent variable and Group (Controls/PD/PD psychosis) as the independent variable. Results showed a significant main effect of Group $F(2, 50) = 4.37$; two-tailed $p = .02$; part. $\mu 2 = .15$). This further analysis confirmed what observed with the previous analysis: Control subjects are better able to differentiate between choices with

higher and lower chance of reward and to use this information to adapt their behaviour. In this regard, it is important to notice that PD patients with psychosis still showed a significant difference between High-Likelihood and Low-Likelihood choices (HL > LL), so they did learn the contingencies during Reward Trials. Critically, no significant difference between PD patients with and without psychosis was found.

Neither preferences for a choice, nor differences in reaction times, were evidenced in Neutral and Bivalent Trials.

Furthermore, the number of actual rewards received during Reward Trials was checked. A mixed-effects model was performed using the number of rewards received as the dependent variable and Group (Controls/PD/PD psychosis) as the independent variable. Results showed a significant main effect of Group $F(2, 50) = 5.11$; two-tailed $p = .01$; part. $\mu 2 = .17$). Bonferroni-corrected post-hoc tests revealed a significant difference only between Controls and PD psychosis ($p < .02$). On average, more rewards were received by Controls (mean = 21.16, sd = .22) relative to PD psychosis (mean =17.25, sd =4.39). Critically, no significant difference between PD patients with and without psychosis (mean =18.17, sd =4.39) was found ($p > .9$).

Computational Model results

The parameters α, β and negative log-likelihood (estimated with the Q-learning model), were statistically compared to check for possible differences between groups and conditions.

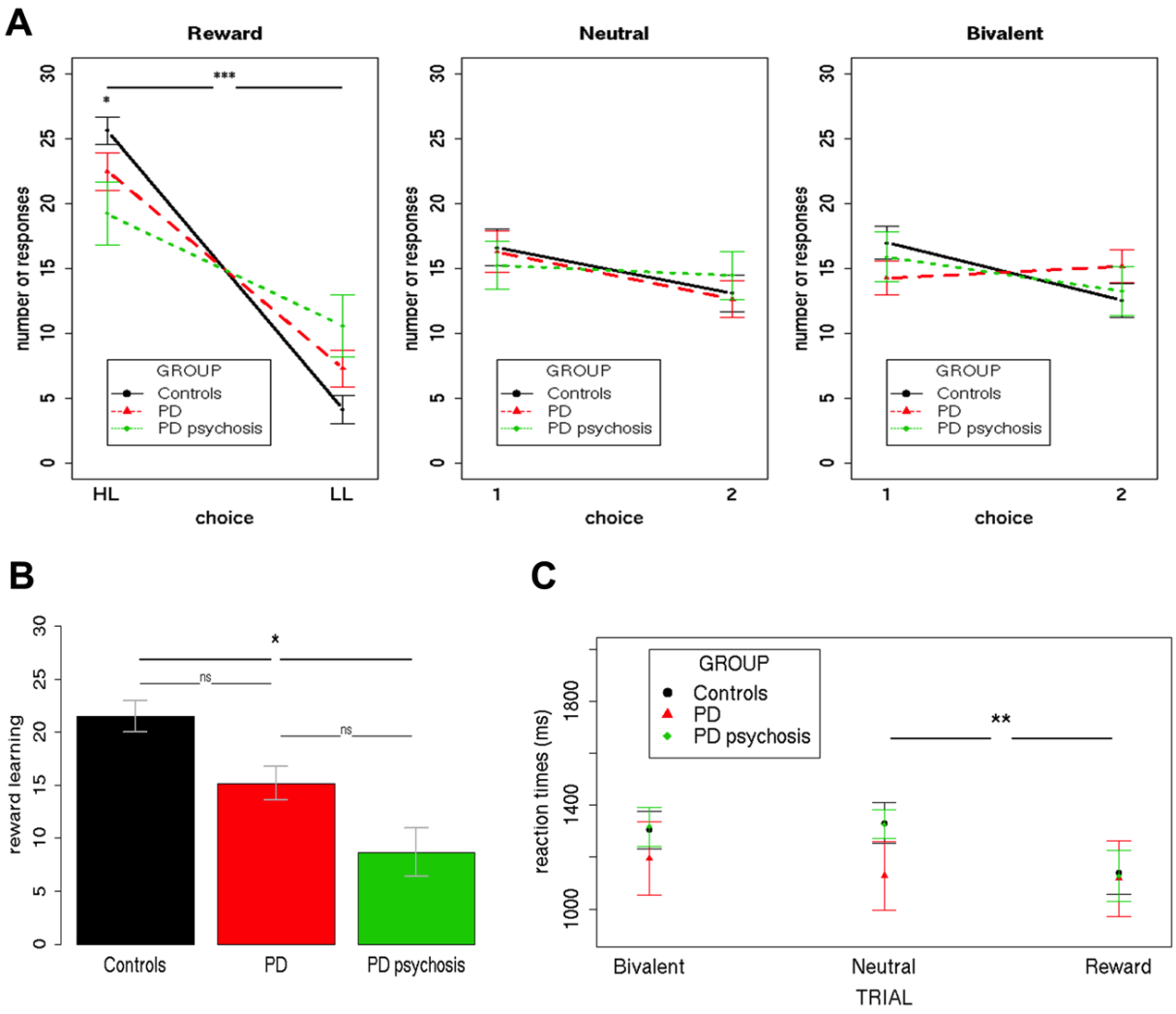

Figure 4.2 – Behavioural performance
Panel A shows the total number of responses during the three trial types (Reward/Neutral/Bivalent) for the three groups (Controls/PD/PD psychosis). Panel B shows the reward learning index (High-Likelihood - Low-Likelihood) for the three groups (Controls/Parkinson Disease/Parkinson Disease with psychosis). Panel C shows mean reaction times during the three trial types (Reward/Neutral/Bivalent) for the three groups (Controls/Parkinson Disease/Parkinson Disease with psychosis).
HL = High-Likelihood; LL = Low-Likelihood; 1 = choice one; 2 = choice two; PD = Parkinson's Disease. Bars indicate standard error. ***p<.0001; **p<.001.; *p<.05.

Three separate 3x3 mixed-effects models were performed using each parameter as the dependent variable and Group (Controls/PD/PD psychosis) and Trial type (Reward/Neutral/Bivalent) as the independent variables. For the parameter α, no significant differences were found (ps>.10).

For the parameter β, a significant effect of condition was found $F(1, 95) = 7.87$; $p = .001$; $\mu^2 = .10$. Bonferroni-corrected post-hoc tests revealed a significant ($p = .0004$) difference between Reward Trials (mean =0.84; sd =1.54) and Neutral Trials (mean =2.42; sd =2.17). All other effects were not significant (ps >.08). This result is interpretable as a stronger expression of explorative behaviour during Neutral Trials (thus, when no response is preferred over the other) as compared with Reward Trials (where the High-Likelihood choice is preferred).

In the analysis of the negative log-likelihood a significant effect of condition was found $F(1, 95) = 27.56$; $p < .001$; $\mu^2 = .24$. Bonferroni-corrected post-hoc tests revealed a significant ($p < .0001$) difference between Reward Trials (mean =-10.43; sd =8.23) and both Neutral Trials (mean =-17.28; sd =5.52) and Bivalent Trials (mean =-18.38; sd =2.70). All other effects were not significant (ps >.1). This result is interpretable as a better fit of the model in which performances is expressed as an actual preference for a specific option, as compared to a more random choice (i.e., when there is no preference).

Overall these results showed that no group differences were present on the parameters α, β and negative log-likelihood as estimated by the Q-learning model.

Imaging results

Reward Anticipation

In Controls, the Reward Anticipation was associated with activation in the occipital cortex (peak voxel MNI coordinates: x, y, z = 16, -70, 22; peak voxel Z = 3.53, FWE cluster corrected p = 0.02, 915 voxels). In the whole PD patient group, Reward Anticipation did not show any significant activation.

In all PD patients, as compared with Controls, there was reduced activation in the occipital cortex (peak voxel MNI coordinates: x, y, z = 16, -64, -10; peak voxel Z = 3.90; FWE cluster corrected $p < 0.001$; 5519 voxels).

In PD psychosis patients, as compared with PD patients, there was reduced activation in two clusters, one localized in the medial prefrontal cortex (peak voxel MNI coordinates: x, y, z = 6, 42, 4; peak voxel Z = 3.21; FWE cluster corrected p = 0.008; 1202 voxels) and in a cluster spanning from the right medial temporal lobe to the middle temporal gyrus laterally (peak voxel MNI coordinates: x, y, z = 48, -18, -6; peak voxel Z = 3.34; FWE cluster corrected p = 0.03; 908 voxels).Standard parametrical maps of the Reward Feedback contrasts is presented in Figure 4.3.

Reward Feedback

In Controls, the Reward Feedback was associated with activation in a cluster spanning from the occipital cortex to the ventral striatum bilaterally (peak voxel MNI coordinates: x, y, z = 8, -50, 6; peak voxel Z = 4.89, FWE cluster corrected $p < 0.001$, 1079 voxels) with a further cluster in the medial prefrontal cortex (peak voxel MNI coordinates: x, y, z = -10, 28, 0; peak voxel Z =

4.87; FWE cluster corrected p =.01; 1180 voxels). In the whole PD patient group, Reward Feedback did not show any significant activation.

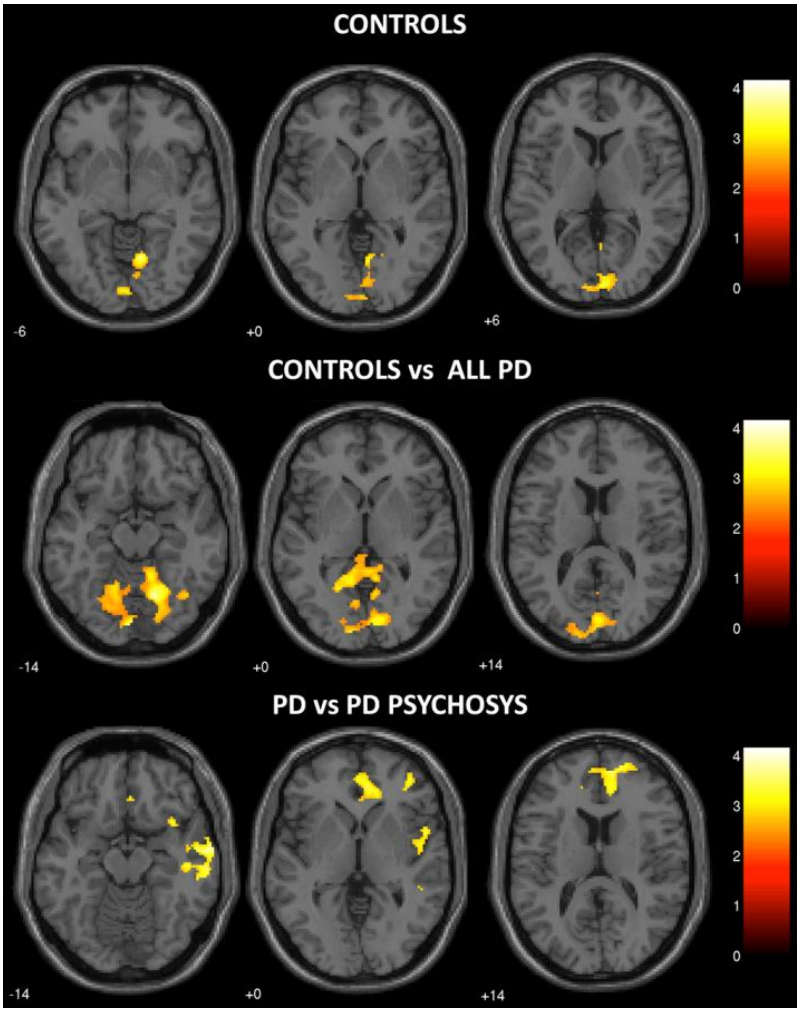

Figure 4.3 – Reward Anticipation
On top -Statistical parametric maps of the Reward Anticipation paramentric modulator. Effects significant at $p < 0.05$ FWE cluster corrected for multiple comparisons are shown in yellow and orange. Left hemisphere is plotted on the left and right hemisphere is plotted on the right. PD = Parkinson's Disease.

In all PD patients, as compared with Controls, there was reduced activation bilaterally in the ventral striatum (peak voxel MNI coordinates: x, y, z = -18, 10, -8; peak voxel Z = 4.97; FWE cluster corrected p =.02; 1206 voxels) and in the occipital cortex (peak voxel MNI coordinates: x, y, z =0, -82, 16; peak voxel Z = 4.94; FWE cluster corrected p < 0.001; 8483 voxels).

In PD psychosis patients, as compared with PD patients, there was reduced activation in a cluster spanning from the medial prefrontal cortex to the ventral striatum bilaterally (peak voxel MNI coordinates: x, y, z = -10, 34, -4; peak voxel Z = 4.16; FWE cluster corrected p = 0.002; 1715 voxels) and a further cluster beginning in the left lateral temporal lobe and peaking in the left orbitofrontal cortex (peak voxel MNI coordinates: x, y, z = -26, 16, -18; peak voxel Z = 4.36; FWE cluster corrected p = 0.03; 1038 voxels).

Standard parametrical maps of the Reward Feedback contrasts is presented in Figure 4.4.

Prediction Error

In Controls, the *Prediction Error* was associated with activation in the left parietal cortex (peak voxel MNI coordinates: x, y, z = -46, -56, 54; peak voxel Z = 4.02; FWE cluster corrected p = 0.001; 2175 voxels). In the whole PD patient group, *Prediction Error* did not show any significant activation.

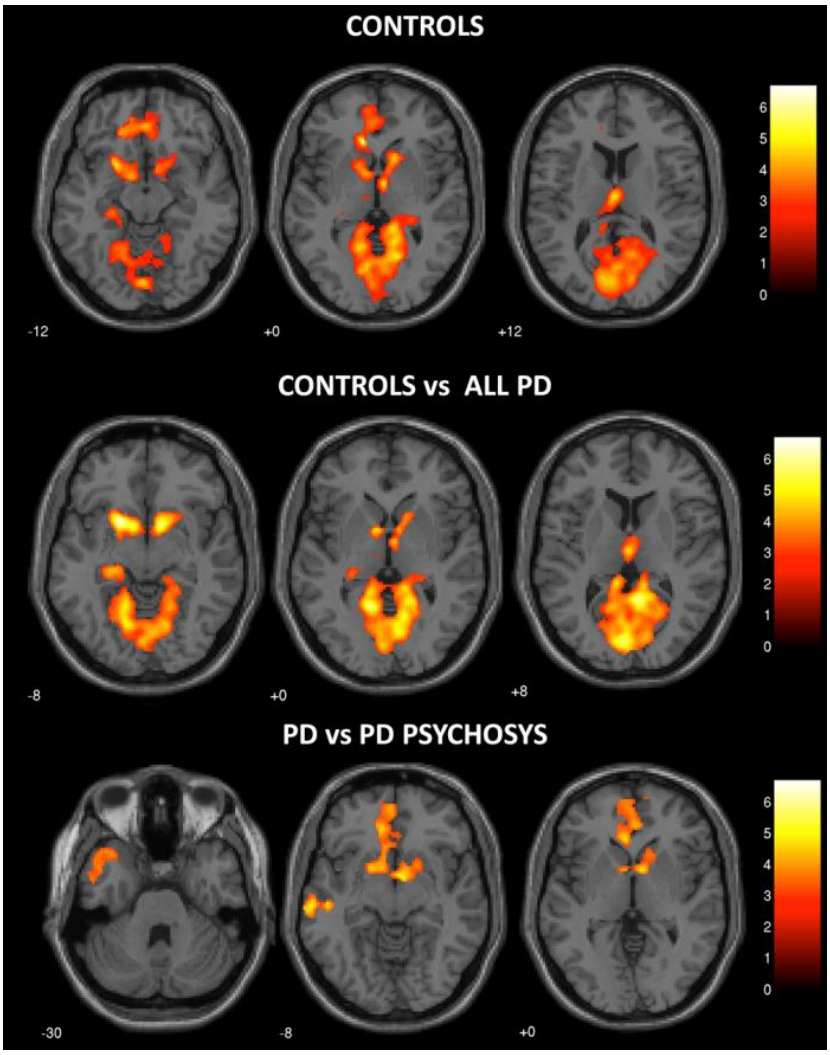

Figure 4.4 – Reward Feedback
On top - Statistical parametric maps of the Reward Feedback contrast. Effects significant at $p < 0.05$ FWE cluster corrected for multiple comparisons are shown in yellow and orange. Left hemisphere is plotted on the left and right hemisphere is plotted on the right. PD = Parkinson's Disease.

In all PD patients, as compared with Controls, there was reduced activation in the left parietal cortex (peak voxel MNI coordinates: x, y, z = -34, -60, 44; peak voxel Z = 4.11; FWE cluster corrected $p < 0.001$; 2254 voxels)

When comparing PD patients without and with psychosis, there was no significant associated activation.Standard parametrical maps of the Reward Feedback contrasts is presented in Figure 4.5.

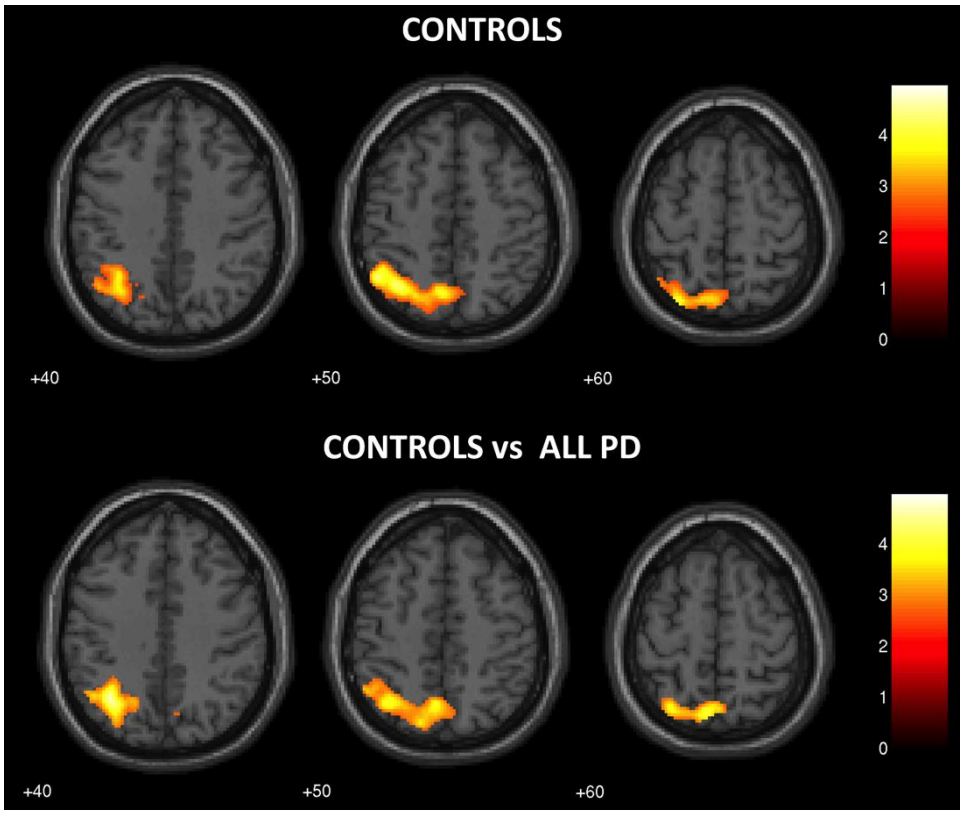

Figure 4.5 – Prediction Error
On top - Statistical parametric maps of the Prediction Error contrast. Effects significant at $p < 0.05$ FWE cluster corrected for multiple comparisons are shown in yellow and orange. Left hemisphere is plotted on the left and right hemisphere is plotted on the right. PD = Parkinson's Disease.

Conclusion

The present study aimed at investigating the neural correlates of reward processing in three groups of subjects: patients with Parkinson's Disease without any history of psychosis (PD) (N=17); patients with Parkinson's Disease associated Psychosis (PD psychosis) (N=12) and a matched group of healthy Controls (N=24).

All participants underwent fMRI scanning while performing an instrumental learning task during which they were presented with visual stimuli and required to choose one of them to maximize their pay-offs. During Reward Trials one choice was associated with an 80% chance to win 1£ (High-Likelihood - HL) and the other with a 20% chance to win 1£ (Low-Likelihood - LL). During Neutral Trials each choice was associated with 80%/20% chance of receiving two kinds of neutral feedbacks. During Bivalent Trials each choice was associated with 50% chance of either losing or winning £1 (Fig.4.1).

Behavioural performance

Analysis of the behavioural performance during the task, suggested that all participants acquired a preference for the highly rewarded choice during Reward Trials (High-Likelihood > Low-Likelihood), thus showing learning of the task contingencies. No preference for a specific choice was present during Neutral and Bivalent Trials. Furthermore, all participants expressed faster reaction times when choosing during Reward Trials, as compared with Neutral and Bivalent Trials. Nevertheless, when considering the number of choices, PD patients with psychosis displayed a significantly lower preference for the rewarded choice, relative to Controls. It is

important to notice that these patients still showed a significant difference between High-Likelihood and Low-Likelihood choices (HL > LL), thus expressing learning. Nevertheless their ability to learn reward-related associations and to use such learning to optimize their behaviour was significantly lower as compared with Controls. Critically, no differences between patients with and without psychosis were found.

fMRI

Reward anticipation and Reward Feedback

The fMRI contrast, looking at Reward Anticipation, showed hypoactivation of the medial prefrontal cortex and the right temporal lobe in PD patients with psychosis, as compared to PD patients without psychosis.

The fMRI contrast looking at Reward Feedback, showed hypoactivation of medial prefrontal cortex, ventral striatum, left temporal lobe and left orbitofrontal cortex in PD patients with psychosis, as compared to PD patients without psychosis.

The finding of abnormal prefrontal activity linked to psychotic symptoms is in line with a significant body of literature showing - across a wide range of procedures and manipulations - altered fronto-striatal connectivity and dysfunctional neural processing of reward in such individuals (Murray et al., 2008a; Fusar-Poli and Meyer-Lindenberg, 2012; Juckel et al., 2006a; de Leeuw et al., 2015; Juckel et al., 2012; Schlagenhauf et al., 2014; 2009). Interestingly, the hypoactivation observed in the medial prefrontal cortex partially overlaps with the one found in a previous study (Bernacer et al., 2013), in which a correlation between psychotic symptoms and

impaired activation for reward anticipation was found in the ventromedial prefrontal cortex, by using a methamphetamine manipulation.

Prediction Error

No relevant differences between PD patients with and without psychosis were found regarding Prediction Error. This lack of a difference might be explained in the light of recent evidences reporting incomplete prediction error signals in older adults (Chowdhury et al., 2013). Senescence can indeed affect the ability to use information about rewards, which can result in little to null reward prediction error signals, probably determining the impossibility to capture such signals from the sample of interest of the present study. In healthy Controls, Reward Anticipation correlated with activation in the occipital cortex only, and this was also significantly underactivated in all PD patients compared with Controls.

Prediction Error correlated in healthy Controls with activation in the right posterior parietal lobe, and this area was hypoactivated in all PD patients as compared with Controls. Even though it did not survive FWE correction, activity in frontal areas was visible when using less stringent cluster-level thresholds ($p = .05$).

However, previous studies suggested a role of the parietal cortex in signalling surprising outcomes (Nieuwenhuis et al, 2005; Glascher et al, 2010), with some evidences also pointing to a dopaminergic mediation of reward processing in this region (Medic et al, 2014).

PART III - EXPERIMENTAL STUDIES ON PAVLOVIAN-TO-INSTRUMENTAL TRANSFER

Study 5 - Individual differences in the influence of task-irrelevant Pavlovian cues on human decision

Abstract

Pavlovian-to-instrumental transfer (PIT) refers to the process of a Pavlovian reward-paired cue acquiring incentive motivational proprieties that drive choices. It represents a crucial phenomenon for understanding cue-controlled behaviour, and it has both adaptive and maladaptive implications (i.e., drug-taking). In animals, individual differences in the degree to which such cues bias performance have been identified in two types of individuals that exhibit distinct Conditioned Responses during Pavlovian conditioning: Sign-Trackers (ST) and Goal-Trackers (GT). Using an appetitive PIT procedure with a monetary reward, the present study investigated, for the first time, the extent to which such individual differences might affect the influence of reward-paired cues in humans. In a first task, participants learned an instrumental response leading to reward; then, in a second task, a visual Pavlovian cue was associated with the same reward; finally, in a third task, PIT was tested by measuring the preference for the reward-paired instrumental response when the task-irrelevant reward-paired cue was presented, in the absence of the reward itself. In ST individuals, but not in GT individuals, reward-related cues biased behaviour, resulting in an increased likelihood to perform the instrumental response independently paired with the same reward when presented with the task-irrelevant reward-

paired cue, even if the reward itself was no longer available (i.e., stronger PIT effect). This finding has important implications for developing individualized treatment for maladaptive behaviours, such as addiction.

Experimental Design

The aim of the present study was to investigate individual differences in human PIT. Specifically, the present study explored, for the first time in humans, whether individual differences in the propensity to approach and engage a Sign (cue-predicting reward) or a Goal (reward) are predictive of cue-controlled behaviour. To this end, a typical PIT experimental design was used, comprising three tasks. In the first phase, participants performed an Instrumental Conditioning task, in which they were presented with two possible choices, one paired with an actual monetary win (Rewarded Choice) and the other paired with a neutral outcome (Unrewarded Choice). In a subsequent session, participants performed a Pavlovian Conditioning task, during which they learned to associate a specific visual cue with an actual monetary win (CS+), and another visual cue with a neutral outcome (CS-). During this phase, eye-movements were recorded and subsequently analysed in order to identify the expressed CR and characterize participants as ST or GT. Mirroring previous studies conducted in animals (Boakes, 1977; Flagel et al., 2011, 2008; Flagel, Watson, Robinson, and Akil, 2007; Robinson, Yager, Cogan, and Saunders, 2013), in which the CR is identified based on the amount of approaching behaviour expressed during CS presentation, in the present study ST and GT participants were distinguished based on a learned oculomotor CR. Specifically, it was measured the tendency to direct contiguous eye-gazes toward the location where the visual CS (Sign) or

the reward (Goal) would be presented. Finally, PIT was tested in an extinction phase (without any rewards), during which participants had to choose between the same two options given during instrumental conditioning, while presented with the task-irrelevant CS. In this final phase, PIT would be observed if presentation of the CS+, compared to the CS-, enhanced instrumental responses to the choice rewarded during instrumental conditioning (Congruent Choice), relative to the previously unrewarded choice (Incongruent Choice). If consistent with animal literature, this effect should be stronger in ST individuals than in GT individuals, possibly indicating a stronger biasing effect of Pavlovian cues over behaviour in the first group relative to the second.

Materials and Methods

Participants

Forty-five volunteers (27 female; 2 left-handed; mean age = 24.87 sd = 2.5; mean education = 17.53, sd = 1.5) with no history of neurological diseases were recruited from the student population at the University of Bologna. All participants gave written informed consent to take part in the experiment and received payment corresponding to the amount earned during the tasks. The study was conducted in accordance with institutional guidelines and the 1964 Declaration of Helsinki. It was approved by the Ethics Committee for Psychological Research at the University of Bologna.

Stimuli and procedure

The whole experiment consisted of three tasks. The same visual background was used in all three tasks. Four black squares (4 cm^2) were displayed on a 17-inch colour monitor with a black background. The squares were highlighted by a white frame and positioned as follows: top

center, bottom center, right center, left center. Two black-and-white fractal images (balanced for luminance, complexity and colour saturation) were used as Pavlovian cues (CS) and presented within the top center square. An image of a 10 euro cent coin was used as the reward, and a light-yellow circle (equally sized) was used as the neutral outcome (no-reward). Both these visual cues appeared within the bottom center square (Fig. 5.1). A computer running Presentation software (Neurobehavioural Systems, Albany, CA) controlled stimulus presentation. On arrival, participants were comfortably seated in a silent room and their position was centered relative to the screen, at a viewing distance of 60 cm from the eye-tracker and 75 cm from the screen. The eye-tracker was positioned under the screen, and was centered relative to both the screen and the participant. Eye-movements and behavioural responses were collected throughout the experiment and stored for offline analysis. Participants were asked to remain as still as possible to avoid confounding effects on eye-movements. The whole experiment was conducted in a dark room to facilitate eye-movement recording. The experimental session began with calibration of the eye-tracker device, during which the participant fixated 9 specific points on the computer screen The experimental session followed the standard paradigm for testing Pavlovian-to-Instrumental Transfer It was composed of three tasks administered in succession: an Instrumental Conditioning task, in which participants learned a response-contingent reward; a Pavlovian Conditioning task, in which participants learned a cue-contingent reward; and a Pavlovian-to-Instrumental Transfer task, during which the influence of irrelevant Pavlovian cues on instrumental responding was tested. In each task, participants were required to pay attention to the screen and follow the instructions reported at the beginning of the task. A few example trials were always performed and, if necessary, further clarifications were given before beginning each

task. At the end of the experimental session, participants completed the Barratt Impulsiveness Scale (BIS-11) (Patton, Stanford, and Barratt, 1995). Previous studies on animals reported an association between Sign-Tracking behaviour and reduced impulse control (Flagel et al., 2011). Thus, this measure allowed further investigation into the differences between ST and GT individuals.

Instrumental Conditioning task. Participants were instructed to choose between two squares to gain a reward. One square was paired with an actual monetary win (Rewarded Choice), while the other was paired with a neutral outcome (Unrewarded Choice). The right and left squares were presented in white and indicated as possible choices to be selected by a mouse click. The mouse pointer was centrally positioned before each choice, in order to not encourage a specific choice. Only one square was associated with a reward following a partial reinforcement schedule, so that between one reward and the next a variable interval between 4 and 12 seconds was always associated with no-reward. After each choice, a corresponding neutral image (light-yellow circle) or reward image (10 euro cents coin) appeared for 1 second in the bottom square (Fig. 5.1-A). Participants were aware that they would receive an actual payment corresponding to the amount of coins collected during the task. The association between square and outcome was counterbalanced across subjects. The rationale of this task was to make participants learn an association between a specific response (left or right square) and the reward; thus, participants would get a higher frequency of Rewarded Choices if they learned the correct association. The task lasted about 6 minutes, during which subjects were free to perform as many choices as they wished, with no time pressure.

Pavlovian Conditioning task. In each trial, one of two possible visual cues (fractal images) appeared for 5 seconds within the top square, followed by a white patch within the bottom square. Upon presentation of the patch, participants were instructed to press the left-Ctrl button on the keyboard as quickly as possible to remove the patch and discover the outcome hidden below. To perform this button press, participants did not need to remove their gaze from the screen. The outcome was then presented for 1 second. One fractal was associated with a reward (10 euro cent coin) on 80% of trials (CS+), while the other fractal was associated with no-reward (light-yellow circle) on all trials (CS-) (Fig. 5.1-B). The task consisted of 40 trials (20 per condition) with a variable inter-trial-interval between 0.5 and 4 seconds. Participants were aware that they would receive an actual payment corresponding to the amount of coins collected during the task. The association between visual cue and outcome was counterbalanced across subjects. The whole task lasted around 6 minutes.

The Pavlovian speeded reaction time response described above ("press the button upon patch presentation") has been successfully used in previous studies (Talmi et al., 2008) and was introduced to obtain a behavioural measure of Pavlovian conditioning. The main reason for using a speeded response was to mirror PIT studies on animals, in which Pavlovian conditioning is measured by a behaviour performed to gain the reward (e.g., latency of the first nose-poke or frequency of nose-pokes) (Corbit and Balleine, 2005; Dickinson, Smith, and Mirenowicz, 2000; Holland, 2004). The rationale here is to observe a faster reaction times when a reward was predicted (CS+ condition) than when a neutral outcome was predicted (CS- condition). To avoid a possible instrumental influence on the task, participants were explicitly told that, in this task, the reward was not contingent on their response. It was demonstrated that, if no answer was

given, the patch would disappear anyway after 1.5 seconds, revealing the outcome. Importantly, this speeded reaction time response allowed us to obtain a measure of the learning rate that is independent from ST/GT behaviour.

To identify ST and GT Conditioned Responses, eye-movements were recorded in order to evaluate contiguous eye-gazes directed toward the "Sign" (top center square) and the "Goal" (bottom center square). Mirroring animal studies, these two Conditioned Responses were subsequently used to distinguish participants as ST or GT, depending on the tendency to direct eye-gaze toward the Sign or the Goal during the 5 seconds of CS presentation (Flagel et al., 2011).

Pavlovian-to-Instrumental Transfer Task. Participants received exactly the same instructions as in the Instrumental Conditioning phase requiring them to choose between the right and left white squares. The task was identical to the Instrumental Conditioning task, except in two aspects: first, the task-irrelevant Pavlovian CS were presented sequentially within the top square, changing every 30 seconds; second, the task was completely performed in extinction, so all choices always lead to no-reward. (Fig. 5.1-C). Extinction is a standard procedure for assessing PIT, both in human and animal research, since it allows one to test the influence of Pavlovian cues on instrumental responding without the confounding effects of the reward (Bray et al., 2008; Corbit, Muir, and Balleine, 2001; Rescorla, 1994a, 1994b; Talmi et al., 2008). Indeed, the rationale here is to test the ability of a task-irrelevant Pavlovian cue to drive choices (presumably, towards the response previously associated with a reward) even if the reward is not

available anymore. The PIT task lasted about 6 minutes, during which subjects were free to perform as many choices as they wished, with no time pressure.

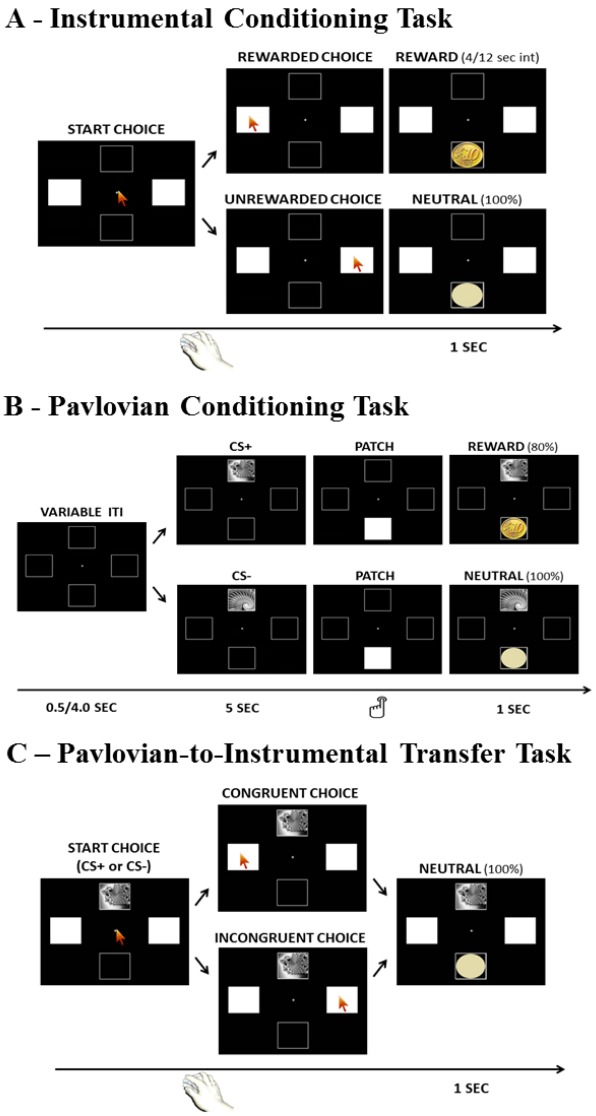

Figure 5.1
Graphical illustration of the three tasks: Instrumental Conditioning Task (Panel A), Pavlovian Conditioning Task (Panel B); Pavlovian-to-Instrumental Transfer Task (Panel C).

Eye tracking

Eye movements were recorded in a dimly lit room using a Pan/Tilt optic eye-tracker (Eye-Track ASL-6000) which registers real-time gaze at 50Hz. Data acquired during the Pavlovian Conditioning task were analysed offline using EyeNal Analysis Software (ASL). Dwell time during the 5 seconds of CS presentation was then measured for two specific areas of interest (AOI): "Sign", corresponding to the 4 cm square at the top center, plus a 1 cm margin; "Goal", corresponding to the 4 cm square at the bottom center, plus a 1 cm margin. Dwell time was defined as the amount of time during which a series of contiguous fixations remained within the same AOI.

Sign-Tracker and Goal-Tracker categorization

Participants were categorized as ST or GT based on the oculomotor CR expressed during the Pavlovian Conditioning task. Previous studies used approaching and engaging behaviours during Pavlovian Conditioning to identify ST and GT. In these studies, the numbers of contacts with the Sign (i.e., lever) and the Goal (i.e., food tray) were compared to obtain an index of behaviour, and divide the subjects into ST (i.e., high probability to engage the lever) and GT (i.e., high probability to engage the food-tray) individuals (Flagel et al., 2011, 2008, 2007; Robinson and Flagel, 2009; Robinson et al., 2014; Saunders and Robinson, 2013). This method was adapted in the present experiment by calculating contiguous eye-gazes (Dwell Time) toward the cue (Sign) and the reward (Goal) Areas of Interest (AOI), during CS presentation (see above). ST behaviour has been defined as a Conditioned Response to approach and engage “the cue or sign that indicates impending reward delivery"; while GT behaviour has been defined as a tendency to

"engage the location of unconditioned cue delivery, even though it is not available until conditioned cue termination" (Flagel et al., 2011). Thus, a learned oculomotor CR towards the location of the Sign or the Goal is a practical method for distinguishing between ST and GT individuals. On this basis, an eye-gaze index was created based on the Dwell Time spent on the Sign and Goal locations. An individual dwell is defined as the time period during which a fixation or series of temporally contiguous fixations remain within an Area of Interest (AOI). That is, an individual dwell is defined as the sum of the durations across all fixations within the current AOI, from entry to exit. To compute fixations, EyeNal Analysis Software was used, which defines a fixation if the observer' s gaze position remains within a diameter of 0,5 degree of visual angle for at least 120 ms (6 consecutive samples, at 50 Hz sampling rate; Eye-Analysis software Manual, v. 1.41, Applied Science Laboratories, 2007). The Dwell Time spent on the Sign and Goal locations was calculated for each trial and then averaged for each participant. The eye-gaze index was calculated as the difference between the Dwell Time on Sign minus the Dwell Time on Goal over the total Dwell Time (Sign - Goal / (Sign + Goal)), so that a higher value corresponded to a higher Dwell Time toward the Sign (Sign-Tracking behaviour) and a lower value corresponded to a higher Dwell Time toward the Goal (Goal-Tracking behaviour). Since the interest here was to disentangle two reward-specific Conditioned Responses, only CS+ trials in the second half of the task were considered, when contingency learning was more established. Based on this index, the top and bottom 50% of the total sample were categorized as ST (eye-gaze index between 0.38 and 1.00) and GT (eye-gaze index between -1.00 and 0.27), respectively.

Results

Sign-Tracker and Goal-Tracker Conditioned Responses

To ensure that the oculomotor responses used to categorize ST and GT individuals were learned CRs, eye-gaze indices were separately analysed for CS+ and CS- trials in the first and second halves of the Pavlovian Conditioning task. Two separate mixed-effects models with Group (ST/GT) and Hemiblock (1/2) as independent variables were performed for CS+ and CS- conditions. The eye-gaze index described above was the dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results from CS+ trials showed a significant interaction effect ($F(1, 42) = 14.75$; two-tailed $p = .0004$; part. $\eta^2 = .26$). Bonferroni-corrected post-hoc tests revealed a significant difference ($p=.003$) between ST (mean = .35; sd = .77) and GT (mean = -.06; sd = .79) in the second Hemiblock (Fig. 5.2-A). No other post-hoc comparisons were significant (ps > .15). Results from CS- trials did not show any significant effects (ps > .05) (Fig. 5.2-B). Overall, these results indicate two important points: first, a bias toward either the Sign or the Goal is a learned Conditioned Response, since it is not present at the beginning of the task but emerges later in time, when contingencies have been learned (Fig. 5.2-A); moreover, this looking bias is specific to the reward-paired cue (CS+), as no differences were observed for the unpaired cue (CS-) (Fig. 5.2-B).In Fig. 5.2-A it is evident how, at the beginning of the Pavlovian task, during CS+ presentation, no tendency seems evident, while, towards the end ST show higher Dwell Time towards the Sign (eye-gaze index increases) while GT show higher Dwell Time towards

the Goal (eye-gaze index decreases). Fig. 5.2-B, on the other hand, shows that the same pattern is not observable during the presentation of the neutral stimulus (CS-).

To further test that this behaviour is a reward-specific CR, the eye gaze index was also directly compared between CS+ and CS- trials from the second hemiblock (when contingencies had been learned) within each group. Two separate paired t-tests were performed for the ST and GT groups, using Condition (CS+/CS-) as the independent variable and the eye-gaze index as the dependent variable. In both groups a significant difference between the two conditions was found. The ST group showed a significantly higher eye-gaze index in the CS+ condition than in the CS- condition ($t(21) = 1.69$; one-tailed $p = .03$; Cohen's $d = .19$), indicating a greater tendency to direct contiguous eye-gazes towards the Sign during CS+ trials than during CS- trials (Fig. 5.2-C). The GT group showed a significantly lower eye-gaze index in the CS+ condition than in the CS- condition ($t(21) = 2.21$; one-tailed $p = .01$; Cohen's $d = .24$), indicating a greater tendency to direct contiguous eye-gazes towards the Goal during CS+ trials than during CS- trials (Fig. 5.2-D).

Given the specific spatial locations of the Sign and the Goal in the present paradigm, visual exploratory behaviour was also considered by analyzing the total dwell time spent on the top and the bottom portions of the screen, in order to exclude the presence of a spatial bias that could account for ST and GT behaviour. A mixed-effects model was used, with Group (ST/GT) and AOI (Top/Bottom) as independent variables and Total Dwell Time as dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a marginal main effect of AOI ($F(1, 42) =$

4.01; two-tailed p = .05; part. η^2 = .09), with more Dwell Time spent on the Top of the screen (mean = .76; sd = .91) than on the Bottom (mean = .41; sd = .64) in both groups (Fig. 5.2-E). Neither group differences, nor interaction effects emerged (ps > .87). These results strengthen the evidence that the behavioural differences observed between ST and GT cannot be ascribed to a mere spatial bias towards the upper or the lower part of the screen. The general difference in time spent looking at the Top and the Bottom of the screen is compatible with the fact that dwell time was calculated during the 5 seconds of CS presentation. These results thus indicate that both groups spent more time visually exploring the region of the screen where a stimulus was being presented (Top), rather than where there was no stimulus (Bottom). No difference in this spatial bias was found between the two groups (Fig. 5.2-E).

Taken together, the last two analyses demonstrated that group differences in the tendency to direct contiguous eye-gazes to the location of the Sign or the Goal cannot be ascribed to a mere spatial bias, but rather reflect a learned reward-related conditioned response.

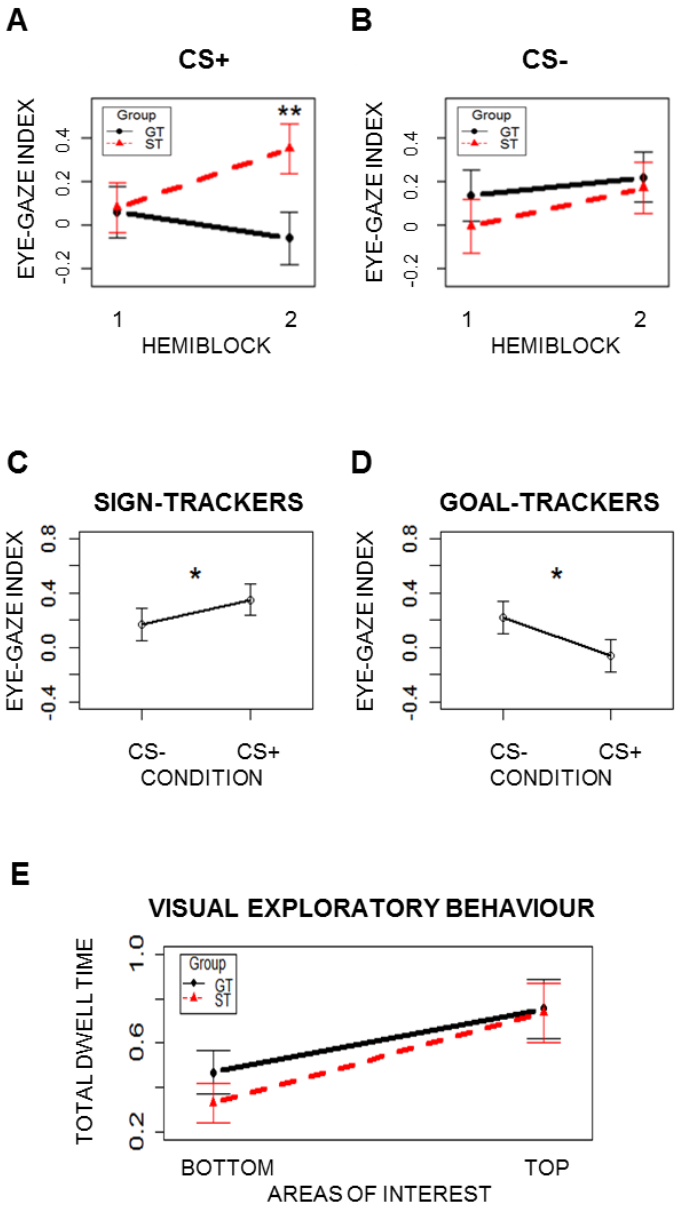

Figure 5.2
Oculomotor response. Panels A and B show the eye-gaze index in the two groups (ST = Sign-Trackers; GT = Goal-Trackers) and the two task hemiblocks. Panel A represents CS+ trails and panel B represents CS- trials. Panels C and D show the eye-gaze index in the two conditions (CS+ = reward-associated cue; CS- = neutral cue) in Sign-Trackers and Goal-Trackers, respectively. Panel E shows visual exploratory behaviour in the two groups (ST = Sign-Trackers; GT = Goal-Trackers) throughout the task. Bars indicate standard error of the mean. * p <.05; **p <.01; *** p <.001.

Instrumental Conditioning

To ensure that instrumental conditioning was successful in both the ST and the GT groups, so that all participants learned which response leads to a reward, the number of choices (mouse clicks) made on the two white squares were compared. Choosing the square associated with reward was considered a Rewarded Choice, and choosing the square associated with no-reward was considered an Unrewarded Choice. A mixed-effects model was used, with Choice (Rewarded/Unrewarded) and Group (ST/GT) as independent variables and the number of choices as the dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a main effect of Choice ($F(1, 42) = 20.88$; two-tailed $p < .0001$; part. $\eta^2 = .33$), with Rewarded Choices (mean = 32.80; sd = 9.38) occurring more frequently than Unrewarded Choices (mean = 22.09; sd = 9.10) (Fig. 5.3-A). Neither group differences, nor interaction effects emerged (ps > .55). These results indicate that the ST and GT groups learned to discriminate between the rewarding and non-rewarding choices equally well.

Pavlovian Conditioning

To ensure that Pavlovian learning occurred in both ST and GT groups, reaction times to patch presentation were analysed. If participants correctly learned to discriminate between the two Pavlovian cues, faster reaction times should be observed for CS+ trials relative to CS- trials. A mixed-effects model was used, with Condition (CS+/CS-) and Group (ST/GT) as independent variables, and reaction times as the dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were

verified. Results showed a significant main effect of Condition ($F(1, 842) = 110.24$; two-tailed $p = .0001$; part. $\eta^2 = .72$), with faster reaction times for CS+ trials (mean = 306.33; sd = 44.41) relative to CS- trials (mean = 351.21; sd = 50.05) (Fig. 5.3-B). Neither group differences, nor interaction effects emerged (ps > .29). These results indicate that participants generally reacted more quickly to the patch on trials with the reward-paired cue (CS+) than on trials with the unpaired cue (CS-). This reward-specific response facilitation indicates successful Pavlovian conditioning in both ST and GT.

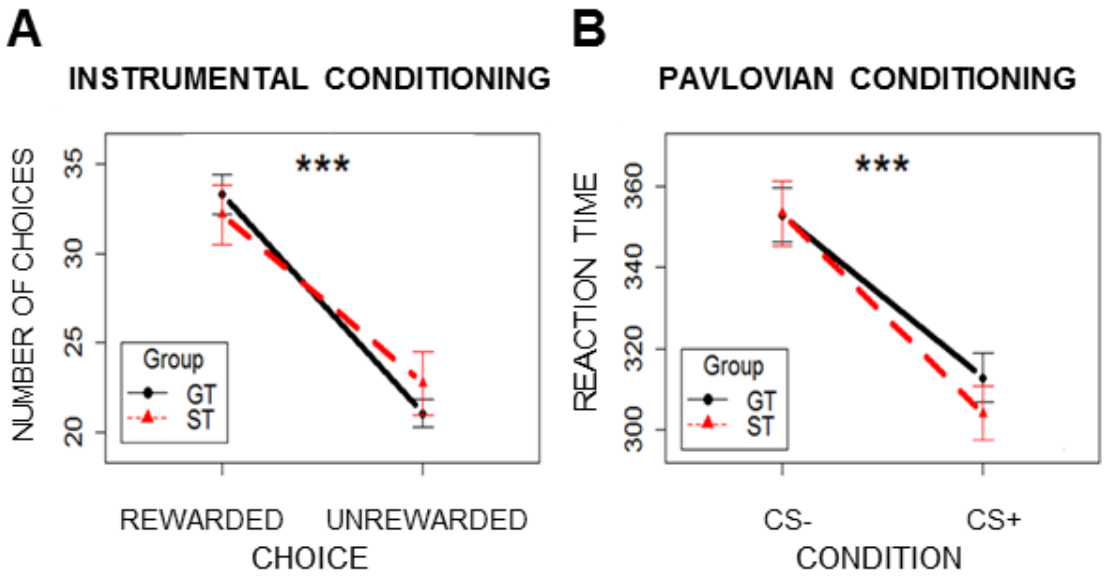

Figure 5.3
Learning rates in the two groups (ST = Sign-Trackers; GT = Goal-Trackers) during Instrumental Conditioning (panel A) and Pavlovian Conditioning (panel B). Bars indicate standard error of the mean. * p <.05; **p <.01; *** p <.001.

Pavlovian-to Instrumental Transfer

To test for PIT, the numbers of Congruent choices (associated with the reward during Instrumental Conditioning) and Incongruent choices (associated with no-reward during Instrumental Conditioning) during CS+ and CS- presentation were compared. A response index was calculated as the probability of selecting the Congruent choice minus the probability of selecting the Incongruent choice (number of congruent – incongruent choices / total number of choices). Higher values correspond to a higher probability of making the Congruent choice, while lower values correspond to a higher probability of making the Incongruent choice. A mixed-effects model was used, with Condition (CS+/CS-) and Group (ST/GT) as independent variables and the response index, described above, as the dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a significant Condition X Group interaction ($F(1, 42) = 8.22$; two-tailed $p = 0.006$; part. $\eta^2 = .16$). Bonferroni-corrected post-hoc comparisons revealed a significant difference ($p = .001$) between CS+ (mean = .18; sd = .12) and CS- (mean = .04; sd = .13) only in ST group, and a significant difference ($p = .04$) between ST (mean = .18; sd = .12) and GT (mean = .08; sd = .12) during CS+ (Fig. 5.4-A). No other comparisons were significant ($ps > .13$). These results indicate that the ST group was more likely to choose the congruent option when they saw the task-irrelevant CS+ than when they saw the CS-. thus revealing a PIT effect. Critically, this bias was stronger in ST than in GT individuals.

While the first analysis on PIT focused on the overall effect, a second analysis divided the task into three equal blocks of 2 minutes (4 trials) to check for differences in task performance over time. A mixed-effects model was used, with Condition (CS+/CS-), Group (ST/GT) and Block

(1/2/3) as independent variables, and the response index as the dependent variable. Subjects were modeled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a significant main effect of Condition ($F(1, 42) = 6.39$; two-tailed $p = 0.02$; part. $\eta^2 = .13$), a significant Condition X Group interaction ($F(1, 42) = 7.69$; two-tailed $p = 0.008$; part. $\eta^2 = .15$), and a significant Block X Group interaction ($F(1.27, 53.32) = 50.61$; two-tailed $p < 0.001$; part. $\eta^2 = .5$) (Fig. 5.4-B/C). Bonferroni-corrected post-hoc tests on the Condition X Group interaction revealed a significant difference ($p = .003$) between CS+ and CS- in ST group but not the GT group, and a significant difference ($p = .02$) between ST and GT groups in CS+ trials (Fig. 5.4-B/C). Bonferroni-corrected post-hoc tests on the Block X Group interaction revealed a significant difference ($p < .0001$) between ST and GT groups in the third block, but not in the first and second blocks (Fig. 5.4-B/C). Fig. 5.4-D shows the number of responses.

In line with the results of the first analysis, these results showed that, unlike GT, ST group was more likely to choose the congruent option when they saw the task-irrelevant CS+ than when they saw the CS-, throughout the entire PIT task. The only effect of time revealed by this analysis was in the last block, where a group difference in responses emerged. Since this difference was unrelated to the displayed stimulus (CS+/CS-), it does not constitute a difference in PIT. This result instead indicates that the ST and GT groups differed in the proportion of congruent choice made towards the end of the task, thereby revealing a possible faster extinction of the preference for the previously rewarded choice in the GT group than in the GT group.

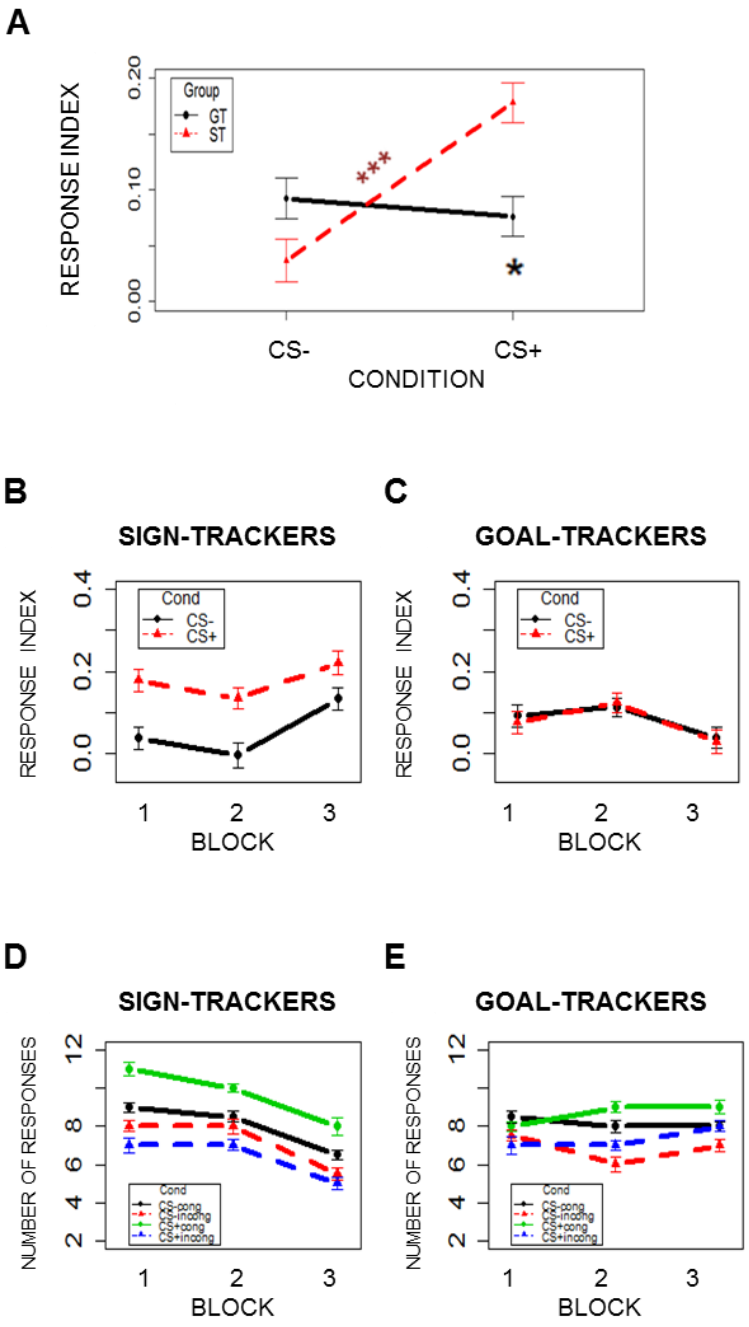

Figure 5.4

Pavlovian-to-Instrumental Transfer. Panel A shows the response index (Congruent-Incongruent/Total) in the two groups (ST = Sign-Trackers; GT = Goal-Trackers) during CS- and CS+ trials. Panels B and C show the response index over time by dividing the task into three blocks of two trials. Panel D shows the number of responses. Bars indicate standard error of the mean. * p <.05; **p <.01; *** p <.001.

Impulsiveness

To further investigate differences between ST and GT individuals, self-reported impulsiveness, as rated by the BIS-11 questionnaire (Patton et al., 1995), was compared between the two groups. A two-sample t-test was performed using Group (ST/GT) as the independent variable and BIS-11 scores as the dependent variable. Results revealed a significant difference between the two groups (t(28.75) =2.06; two-sided p = 0.04, with the ST group (mean = 61.0; sd =9.91) showing higher impulsiveness than the GT group (mean = 54.09; sd = 8.86) (Fig. 5.5). This finding is consistent with previous studies showing significantly higher levels of impulsiveness as compared to GT (Flagel et al., 2009; Tomie, Aguado, Pohorecky, and Benjamin, 2000).

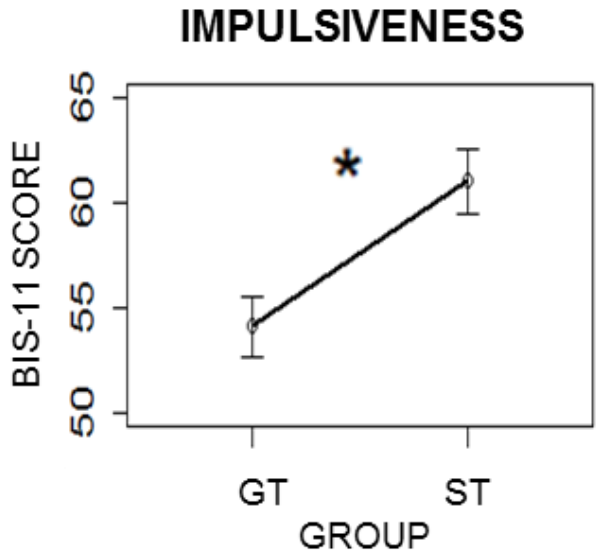

Figure 5.5
Impulsiveness levels in the two groups (ST = Sign-Trackers; GT = Goal-Trackers) as measured by the Barratt Impulsiveness Scale (BIS-11). Bars indicate standard error of the mean. * p <.05; **p <.01; *** p <.001.

Conclusion

In the present study, the Pavlovian-to-Instrumental Transfer (PIT) paradigm was used to examine individual differences in the excitatory influence that signals associated with reward can exert on human choices. PIT is a well-known procedure for testing the ability of a Pavlovian reward-paired cue to acquire incentive motivational properties and influence instrumental performance (de Wit and Dickinson, 2009; Estes, 1943, 1948; Holmes et al., 2010; Rescorla and Solomon, 1967). Here, participants performed a standard PIT paradigm composed of 3 tasks: an Instrumental Conditioning task, during which response-outcome associations were learned; a Pavlovian Conditioning task, during which stimulus-outcome associations were leaned; and a PIT task, in which the ability of a Pavlovian cue to drive instrumental responses was tested. Individual differences were characterized by two distinct oculomotor Conditioned Responses (CR) exhibited during Pavlovian Conditioning, corresponding to two different learning styles previously identified and described in animal literature: Sign-Tracking (ST) and Goal-Tracking (GT) (Boakes, 1977; Estes, 1943, 1948; Flagel et al., 2011). In the present study, ST behaviour consisted of a tendency to direct contiguous eye-gazes towards the cue (CS) that indicated impending reward delivery (Sign); in contrast, GT behaviour was characterized by a tendency to direct contiguous eye-gazes towards the location of reward (US) delivery (Goal), even if not available until CS termination. An eye-gaze index was based on the emergence of these two behavioural patterns during presentation of the reward-paired stimulus (CS+) in the second half of the task (when contingencies had been learned), and a median split was used to categorize participants as ST or GT. Importantly, the present results demonstrate that this oculomotor CR was (i) acquired over time (i.e., learned), since a specific CR towards the Sign or the Goal only

emerged towards the end of the task, when stimulus-reward associations had been acquired selectively during the presentation of reward-paired cues (CS+) (Fig. 5.2-A/B); and (ii) reward specific, since the CR was only evident when participants saw the reward-related cue (CS+) and not when they saw the neutral cue (CS-) (Fig. 5.2-C/D). Coherently with what expected, the task-irrelevant CS had a much stronger influence on the ST group than on the GT group during the PIT task.

Group differences in the PIT effect are not attributable to differences in the strength of Instrumental or Pavlovian learning between the groups, which could have potentially induced a bias towards the rewarded choice in the Instrumental Conditioning task, or a stronger influence of the reward-paired cue in the second Pavlovian Conditioning task. Analyses of both the number of rewarded choices during Instrumental Conditioning, and reaction times during Pavlovian Conditioning, exclude such a possibility by revealing that both the ST and GT groups learned the response-outcome and stimulus-outcome contingencies equally well (Fig. 5.3). Consequently, differences in the PIT effect cannot be explained by group differences in the ability to learn either the instrumental or the Pavlovian contingencies. In line with the animal literature (Robinson and Flagel, 2009), the Pavlovian cue (CS+) was clearly predictive of reward, since it elicited faster reaction times during Pavlovian conditioning than the neutral stimulus (CS-) did in both groups, along with a CR corresponding to the behavioural profile of each group (ST/GT).

Since the “Sign” and the “Goal” had specific spatial locations (the top and bottom portions of the screen, respectively), it is important to rule out the possibility that spatial biases in gaze direction

might account for the difference in the PIT effect between groups. A bias towards looking at the top of the screen might cause result in a stronger influence of the Sign on the ST group just because they spent more time looking at it. Analysis of visual exploratory behaviour during Pavlovian Conditioning, however, revealed that the ST and GT groups did not differ in the total amount of time spent looking at the top and bottom of the screen (Fig. 5.2-E). Critically, behavioural differences only emerged during CS+ trials towards the end of the task, once the association between the cue and the reward had been learned. Consequently, it is conclude that there was no a-priori bias in gaze direction; rather, such a bias emerged during the Pavlovian Conditioning task as a learned reward-specific CR.

Moreover, a recent study (Trick, Hogarth, and Duka, 2011) directly investigated the relation between fixation times during Pavlovian learning and the PIT effect. The authors found that fixation times during Pavlovian learning increased with uncertainty (that is, more attention was paid to stimuli with uncertain outcome probabilities, e.g., 50%, than to stimuli with more certain outcome probabilities, e.g., 90%). In contrast, the PIT effect increased with the probability of reward (that is, it was stronger for stimuli associated with a high probability of reward, e.g., 90%, than for stimuli associated with uncertain outcomes, e.g., 50%, or a low probability of reward, e.g. 10%). Thus, Trick et al. (2011) concluded that the behavioural influence exerted by conditioned stimuli (i.e., the PIT effect) is dissociated from attention to Pavlovian stimuli in humans, (see Kaye and Pearce, 1984, for similar findings in animals). Instead, PIT is linked to the predictive value acquired by stimuli during learning.

ST behaviour has been explained as a consequence of attributing incentive salience to reward-paired cues (Pavlovian CS), arising from the interaction between previous experience (reinforcement learning processes) and individual propensities (Berridge and Robinson, 2003; Berridge, 2001; Flagel et al., 2011). This incentive salience motivates reward-related action (Flagel et al., 2008; Robinson and Flagel, 2009; Tomie et al., 2000). In the present study, ST and GT groups differed in the extent to which Pavlovian reward-paired cues biased their behaviour: relative to the GT group, the ST group showed an increased likelihood of performing the instrumental response independently paired with the same reward when presented with the task-irrelevant reward-paired cue, even if the reward itself was no longer available (i.e., a stronger PIT effect) (Fig. 5.4-A). Therefore, reward-paired cues exerted a stronger source of influence on the behaviour of ST individuals, as predicted. Importantly, time course analysis revealed that this effect occurred early and remained stable throughout the entire PIT test session (Fig. 5.4-B/C), thereby suggesting that the group difference in the PIT effect most likely reflects greater incentive salience to reward cues in ST than in GT individuals.

A group difference in the overall amount of congruent responses (during both CS+ and CS- presentation, thus not reflecting PIT) emerged towards the end of the task (Fig. 5.4-B/C). This difference might reflect a tendency for GT individuals to extinguish reward-associated instrumental responses more quickly than ST individuals. However, this interpretation should be considered with caution, given the short duration of the task (6 minutes), which might not have been sufficient to properly explore group differences in extinction.

Previous studies have found an association between ST behaviour and other traits, such as higher levels of behavioural impulsivity and a greater propensity to develop addiction (Flagel et al., 2008; Robinson and Flagel, 2009; Tomie et al., 1998). In line with these studies, the present study found reduced self-reported impulse control in the ST group than in the GT group (Fig. 5.5). These findings seem to corroborate the idea that ST and GT behaviours are just one expression of a broader profile of individual differences, which might be clinically relevant. Many studies have reported that ST individuals are more impulsive and prone to develop potentially maladaptive behaviours, such as addiction (Flagel et al., 2011; Robinson and Flagel, 2009; Tomie et al., 1998). For example, the propensity to sign-track is associated with a stronger effect of psychomotor sensitization, a higher susceptibility to a form of cocaine-induced plasticity that may contribute to the development of addiction (Flagel et al., 2008). Furthermore, ST behaviour in relation to a specific Pavlovian cue (i.e., a cue predicting monetary reward) is also predictive of the propensity to attribute incentive salience to other reward-paired cues, such as food-related or drug-related cues (e.g., cocaine and alcohol) (Clark et al., 2013; Cunningham and Patel, 2007; Flagel et al., 2008; Uslaner, Acerbo, Jones, and Robinson, 2006). The extent to which such individual differences might play a role in the development of addiction and in the propensity to relapse is not yet clear, but their implications for developing individually targeted treatment programs are promising.

Study 6 - Triggering avoidance: the power of aversive conditioned stimuli on human behaviour

Abstract

In the present experiment the influence of Pavlovian aversive stimuli on avoidance instrumental responses has been tested by using a Pavlovian-to-Instrumental Transfer (PIT) task. Transfer effect reflects the learned motivational influence of the CS over the instrumental response (Blundell, Hall, and Killcross, 2001; Corbit, Muir, and Balleine, 2001; Dickinson, Smith, and Mirenowicz, 2000; Hall, Parkinson, Connor, Dickinson, and Everitt, 2001; Holland and Gallagher, 2003). Two forms of transfer have been object of investigation: outcome-specific transfer -which is tough to be mediated by the sensory-specific properties of the reinforcer, deriving from a S-O-R association - and general transfer – which is thought to be mediated by a direct S–R association, reflecting a more general motivational response. Additionally, in order to clarify the nature of the transfer effect (goal-directed/habit), the impact of reinforcer devaluation after different amounts of instrumental training has been investigated. A double dissociation between specific and general aversive transfer effects was reported for the first time in humans. Specific transfer was expressed in the percentage of responses but not in the vigour of such responses; whereas the opposite pattern was observed for general transfer, where the effect was captured by the vigour of the responses but not percentage of responses. Moreover, specific transfer was enhanced by instrumental overtraining, but not by reinforcer devaluation. General transfer was not affected by neither instrumental overtraining nor reinforcer devaluation.

Experimental Design

Given the lack of literature human PIT, especially in aversive contexts, a first aim of the present study is to investigate the ability of aversive Pavlovian stimuli to increase the number and vigour of instrumental responses independently paired with the same (outcome-specific transfer) and different (general transfer) punishments. Furthermore, in order to evaluate the nature (goal-directed/habits) of transfer effect, a second aim of the present study is to test the effect of outcome devaluation on specific and general forms of PIT. Finally, to investigate the mechanisms behind transfer, a third aim of this study is to test the effect of instrumental overtraining on outcome-specific and general transfer.

With these three aims in mind, an aversive experimental design was used, comprised of three main tasks (Instrumental Conditioning, Pavlovian Conditioning, PIT), and presented to the participants as a space-war game. Crucially, two measures of transfer were acquired: the percentage of instrumental responses performed in each task and the vigour (hand-grip) of such responses.

Materials and Methods

Participants

Thirty-eight volunteers (18 female; 4 left-handed; mean age = 25.18, sd = 5.69 years; mean education = 16.5, sd = 2.42 years) with no history of neurological diseases were recruited from the student population of the University of Cambridge (UK). All participants gave written

informed consent to take part in the experiment and received payment corresponding to the amount of time needed to complete the tasks. The study was conducted in accordance with institutional guidelines and the 1964 Declaration of Helsinki. It was approved by the Department of Psychology Ethics Committee of the University of Cambridge.

Skin Conductance Response (SCR) recording and analysis

Ambu WS electrodes connected to a DC amplifier (Biopac Systems - MP150 - GSR100) were used for recording galvanic skin response. These were attached to subjects' volar surface of the index and middle fingertip in their left hand (which did not require any motor movement during the task). A gain factor of 5 μS/V and low-pass filter set at 10 Hz were used for recording the analog signal, which was then passed through the digital converter at a 200Hz rate. The signal was then fed into AcqKnowledge 3.9 (Biopac Systems) and transformed into microsiemens for offline analysis. Skin Conductance Response (SCR) was extracted from the continuous signal and calculated for each trial as the peak-to-peak amplitude of the largest deflection during the 0.5-4.5 sec time window following stimulus onset (Schiller et al. 2008). The minimal response criterion was 0.02 μS and smaller responses were encoded as zero. Raw SCR scores were square root transformed to normalize the distributions and scaled to each subject's maximal US response to account for interindividual variability (Schller et al., 2008).

Data were analysed offline using custom-made MATLAB scripts (The MathWorks, Inc., Natick, Massachusetts, United States) and all statistical analyses were performed with RStudio v0.98.1062 (Boston, MA).

This signal was recorded to assess implicit Pavlovian learning. During Pavlovian Conditioning task all trials were recorded, however analyses only included trials in which no aversive noise was delivered (40% of all CS+ and all CS- trials), in order to exclude artifacts.

Hand-grip and response recording and analysis

An isometric hand dynamometer was used (Biopac Systems - MP150 - TSD121C - DA100C) to record hand gripping (compression), by simply squeezing the handle of the transducer. To ensure correct recording of the hand grip, the hand dynamometer was calibrated and each participant was given the chance to familiarize with the maximum and minimum strength possible to be recorded. The hand grip was recorded in kilograms and extracted from the continuous signal by calculating the maximum peak amplitude for each trial.

To allow for multiple responses, the hand dynamometer was attached to the base of a joystick and used as handle. In this way, participants could squeeze the handle bar while moving it towards the left or the right. Both the side and the hand-grip were simultaneously recorded.

These measures were collected to get a measure of the vigour of all responses performed.

Stimuli

Five different custom made images depicting space scenarios were presented in the background of a computer screen during all tasks. Centrally, in the lower part of the screen, a smaller image of a monitor was used to display visual feedbacks. Visual feedbacks consisted in: a green circle

with the inscription “missed”; a red triangle with the inscription “hit”; the inscription “defend yourself”. Images were presented on a 17 inches computer screen, at a viewing distance of 80 cm.

The “hit” feedback was always paired with one of three different aversive noises, consisting in 100db sounds played for 1 second. The three noises had been rated as equally aversive and clearly distinguishable by and independent group of subjects prior to the experiment.

A computer running Presentation software (Neurobehavioural Systems, Albany, CA) controlled stimulus presentation.

Procedure

On arrival, participants were comfortably seated in a silent room and their position was centred relative to the screen. They were required to wear a headset used to deliver aversive sounds during the task.

Galvanic skin response, hand grip and behavioural responses were collected throughout the experiment and stored for offline analysis.

The experiment consisted of three main tasks, with a repetition of some of those. The tasks were presented in the following order: Instrumental Conditioning, Pavlovian Conditioning, PIT, PIT under devaluation (PIT-dev), Instrumental Overtraining, PIT after overtraining (PITo), PIT after overtraining under devaluation (PITo-dev). In each task, participants were required to pay

attention to the screen and follow the instructions. A few example trials were always performed at the beginning and, if necessary, further clarifications were given before each task.

At the end of the experimental session, participants completed the Behavioural Inhibition/Activation System (BIS/BAS) inventory (Carver and White, 1994).

Instrumental Conditioning task. Participants were engaged in a space-war kind of game. In the initial instructions, participants were thanked for taking part to the space mission and warned that during this first part of the task there were two possible kinds of attacks (bombs and missiles), each one corresponding to a different Unconditioned Stimulus (US), consisting in an aversive noise. The two USs were played at this stage to allow participant's familiarization. A single space scenario was presented on the background for the whole duration of the task and, after a random inter trial interval (1.5-2 seconds), a "defend yourself" message was prompted into the small monitor (2 seconds). For the following 30 seconds, one of the two possible USs (corresponding to either bombs or missiles) was randomly delivered according to a random interval (1.5-3 seconds). To avoid attacks, participants were provided with a joystick and required to move it towards left or right, while squeezing, in order to learn the correct response to avoid each specific US. Each side allowed to avoid only one US. If the US was correctly avoided a "missed" visual feedback was displayed in the small monitor; if not a "hit" visual feedback was presented (1 second) (see Stimuli section for a description) (Fig. 6.1-A). The association between response (left/right) and attack (bombs/missiles) was counterbalanced across subjects. The rationale of this task was to learn the association between a specific US and the

correct response (left or right) required to avoid it. Participants performed 4 trials for each kind of attack, for a total of 8 trials and a duration of about 5 minutes.

At the end of this task, explicit learning was assessed by asking participants to pair each US with the corresponding correct avoidance response.

Pavlovian Conditioning task. Participants were presented with new instructions informing that in this stage they would be travelling through different galaxies (corresponding to the variable space scenario backgrounds used as Conditioned Stimuli) and that more attacks could be delivered. They were also informed that they would not be able to use the joystick to avoid those attacks and were required to pay attention to the contingencies.

In each trial, after a variable inter trial interval (7-9 seconds), one of four possible space scenarios (conditioned stimuli - CSs) was presented in background (4.5 seconds) and was followed by a "missed" or "hit visual feedback (1 second). Two scenarios (CS+1/CS+2) were paired with the same two USs previously used during Instrumental Conditioning; a third scenario (CS+3) was paired with a new US; a fourth scenario (CS-) was associated with no sound. All CS+ followed a 60-40 partial reinforcement schedule (Fig. 6.1-B). The association between CS and US was counterbalanced across subjects. The rationale of this task was to learn the association between the different space scenarios (CSs) and each US. Participants performed 20 trials for each CS condition, for a total of 80 trials and a duration of about 15 minutes.

At the end of this task, explicit learning was assessed by asking participants to pair each US with the corresponding CS.

Pavlovian-to-Instrumental Transfer (PIT) task. Participants were instructed that at this stage they could use again the joystick to avoid attacks (as during instrumental conditioning task), but that a malfunction occurred to the monitor and no visual feedback was going to be displayed on the monitor. The task was identical to the Instrumental Conditioning task, except in two aspects: first, the task-irrelevant space scenarios used during Pavlovian Conditioning as CSs were randomly presented in background, one for each trial; secondly, the task was completely performed in extinction, so neither visual feedbacks nor aversive noises ever occurred (Fig. 6.1-C). Extinction is a standard procedure for assessing PIT, both in human and animal research, since it allows one to test the influence of Pavlovian cues on instrumental responding without the confounding effects of the unconditioned stimulus (Bray et al., 2008; Corbit, Muir, and Balleine, 2001; Rescorla, 1994a, 1994b). The rationale of this phase is to test the ability of a task-irrelevant Pavlovian cue to trigger avoidance responses (presumably, towards the one previously associated with the same or a similar punishment) even if no aversive stimuli are delivered. Participants performed 4 trials for each CS condition, for a total of 16 trials and a duration of about 8 minutes.

PIT with devaluation. In this phase, the PIT task was repeated exactly as the first time, but the headset was removed from the participant, so that no USs could be delivered. A similar procedure was successfully used in previous study to assess devaluation (Gillan et al., 2014).

Instrumental Overtraining. During this task, participants repeated the Instrumental Conditioning task for a total of 24 trials and a duration of about 14 minutes, thus three times more than the previous Instrumental Conditioning task.

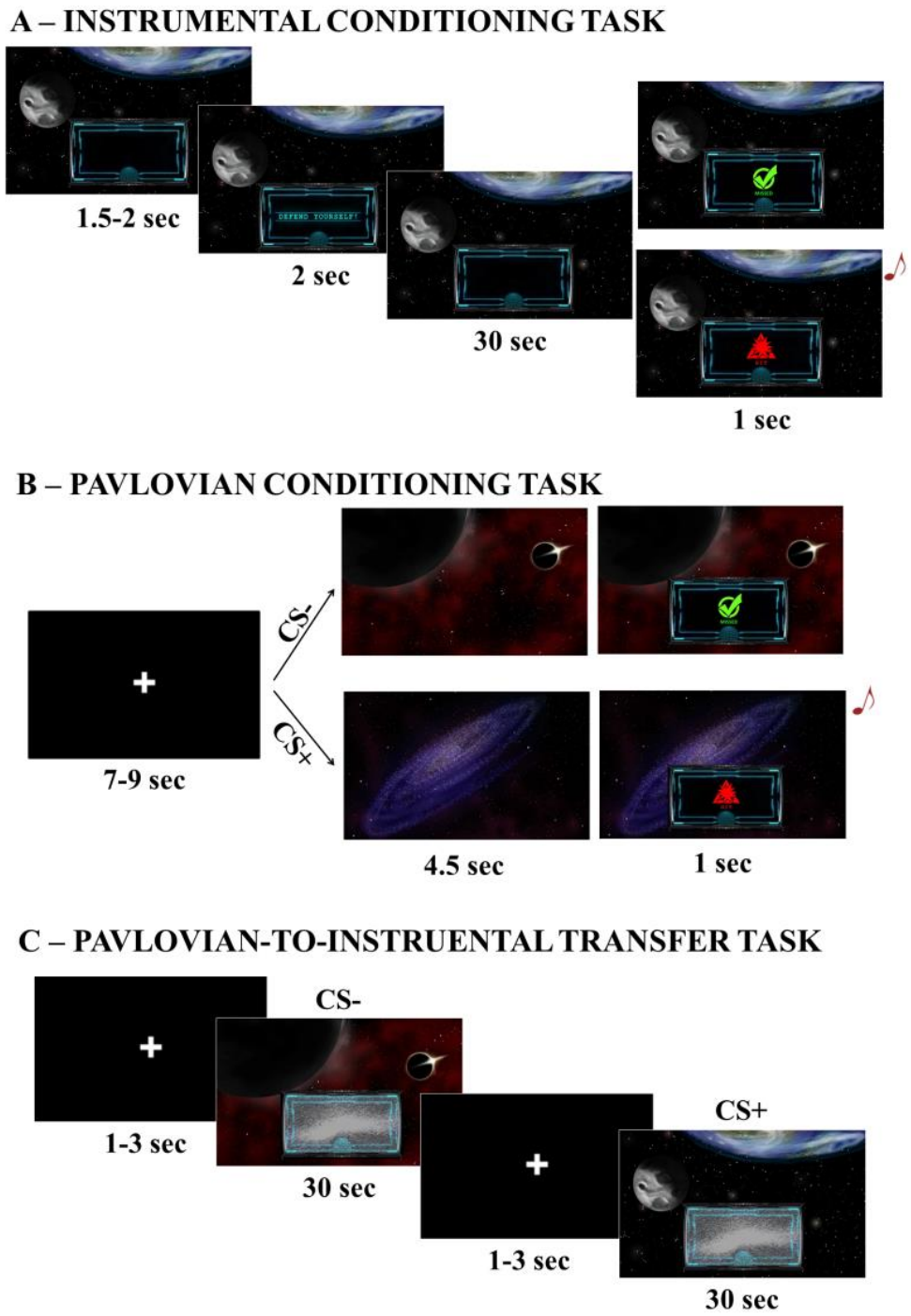

Figure 6.1
Graphical illustration of the three MAIN tasks: Instrumental Conditioning Task (Panel A), Pavlovian Conditioning Task (Panel B); Pavlovian-to-Instrumental Transfer Task (Panel C).

Results

Instrumental Conditioning

To assess implicit learning of the Instrumental Conditioning task, the number of responses performed with the joystick were analysed. For each trial, a response (left/right) was categorized as correct or wrong according to its ability to avoid the current attack (bombs/missiles). Each side, indeed, allowed to avoid only one specific US, uniquely associated with a particular attack. A mixed-effects model was used, with Response (correct/wrong) and US (US1/US2) as independent variables, and the total number of responses as dependent variable. Subjects were modelled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a main effect of Response ($F(1, 37) = 194.4$; two-tailed $p < .0001$; part. $\eta^2 = .84$), with correct responses (mean =49.17; sd =7.41) being performed more than wrong responses (mean =17.82; sd =8.49) (Fig. 6.2-A). All other effects were not significant (ps > .58).

Moreover, 89% of participants gave a correct response when explicitly asked to indicate the avoidance response associated with each US.

These results indicate that participants learned, both implicitly and explicitly, to discriminate between the two USs and the corresponding avoidance response.

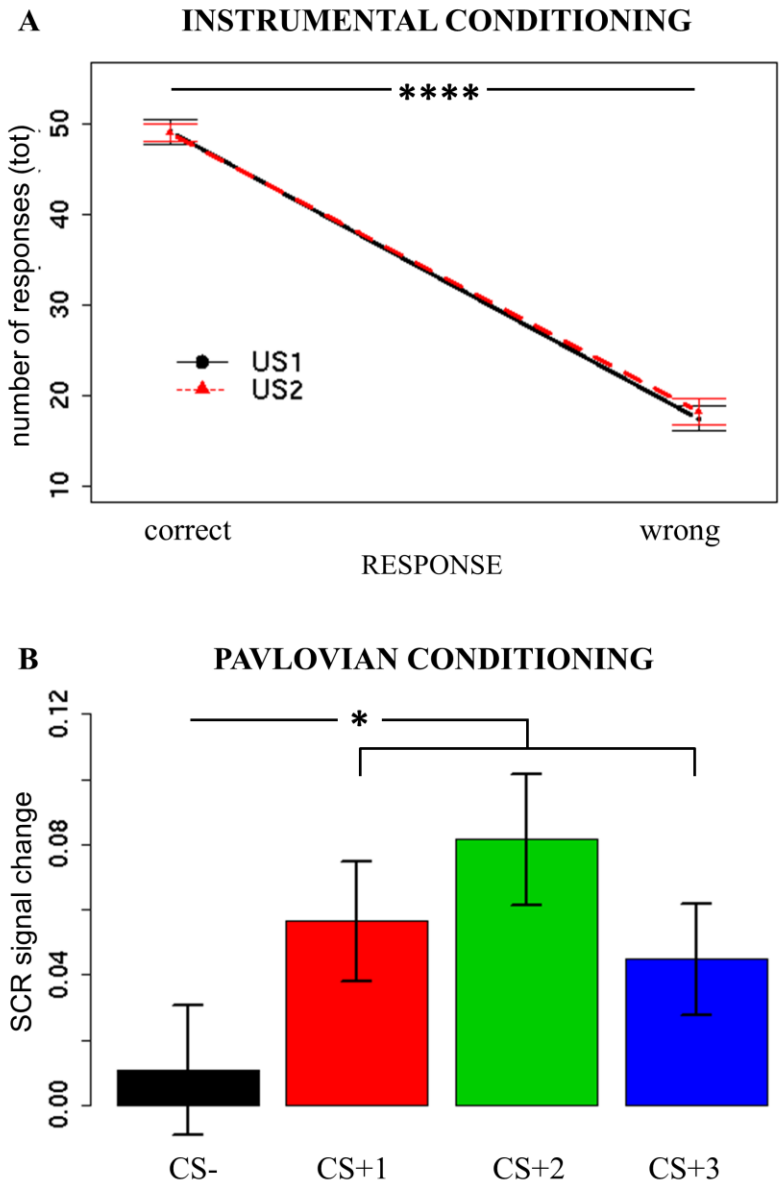

Figure 6.2
Instrumental Conditioning and Pavlovian Conditioning results.
Panel A reports the number of correct and wrong responses performed when presented with the two different attacks/noises (Unconditioned Stimuli – US) during Instrumental Conditioning.
Panel B reports the Skin Conductance Response (SCR) signal change (second hemiblock – first hemiblock) when presented with all possible CSs during Pavlovian Conditioning. Bars indicate standard error of the mean. * p <.05.

Pavlovian Conditioning

To assess implicit learning during the Pavlovian Conditioning task, an SCR signal change index was calculated, to get a measure of changes in the arousal level as learning occurred. To detect variations in time, the difference between SCR during the second and first hemiblocks of the task was calculated (SCR signal change) for each CS. If participants correctly learned to discriminate between aversive and neutral Pavlovian cues, a higher signal change should be observed for all CS+ trials relative to CS- trials. A mixed-effects model was used, with CS (CS+1/CS+2/CS+3/CS-) as independent variables, and SCR signal change as dependent variable. Subjects were modelled as random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified. Results showed a significant main effect of CS ($F(1, 111) = 3.17$; two-tailed $p = .03$; part. $\eta^2 = .08$). Bonferroni-corrected post-hoc analysis revealed a significant difference between CS- (mean =0.011; sd =0.12) and all CS+ (CS+1 mean =0.056, sd =0.11; CS+2 mean =0.081, sd =0.12; CS+1 mean =0.044, sd =0.10;) conditions (respectively, p=.03; p=.01; p=.04), but not between the CS+ (ps > .2) (Fig. 6.2-B).

These results show that, as learning occurred over time, participants' arousal significantly increased when presented with aversive stimuli (all CS+) as compared to a neutral stimulus (CS-), thus indicating successful Pavlovian conditioning.

Pavlovian-to Instrumental Transfer

To asses transfer effect, two dependent variables were considered, namely the percentage of responses and the hand-grip.

Specific transfer was tested considering only CS+1 and CS+2 trials, as these were paired with the same USs used during Instrumental Conditioning. The rationale of specific transfer is, indeed, to test if CSs are able to elicit a response that has been independently associated with the same reinforcer. To this aim, all responses were categorized as congruent (e.g., choosing R1 when presented with CS+1 or choosing R2 when presented with CS+2) or incongruent (e.g., choosing R2 when presented with CS+1 or choosing R1 when presented with CS+2) and compared. Two separate mixed-effects models were used, with Congruency (congruent/incongruent) as independent variable and percentage of responses or hand-grip as dependent variables. Subjects were modelled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified.

Results on percentage of responses showed a significant difference ($F(1, 37) = 2.56$; two-tailed $p = .05$; part. $\eta 2 = .12$) between congruent and incongruent responses, the first being higher than the second (means are reported in Fig. 6.3-A).

Results on hand-grip showed no difference ($p = .7$) between congruent and incongruent responses (means are reported in Fig. 6.3-A).

In sum, specific transfer effect was observed when considering the response frequency (percentage of responses), but not when considering the vigour (hand-grip) of such responses.

General transfer, on the other hand, was tested considering only CS+3 and CS- trials, as these were respectively paired with a US not used during Instrumental Conditioning (unpaired condition) and with no US (neutral condition). The rationale of general transfer is, indeed, to test

if CSs are able to elicit a response that has been independently associated with a similar reinforcer (unpaired), relative to a neutral CS.

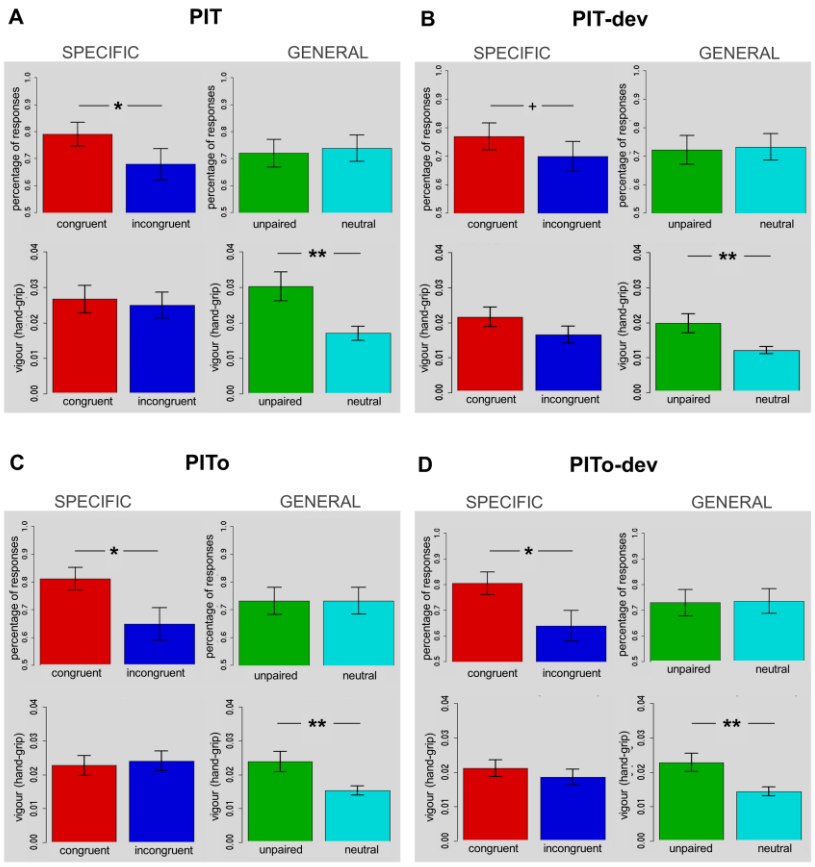

Figure 6.3
Pavlovian-to-Instrumental Transfer (PIT) task results.
Panel A shows the first PIT task performed (PIT); panel B shows PIT under devaluation (PIT-dev); panel C shows PIT after overtraining (PITo); panel C shows PIT after overtraining under devaluation (PITo-dev). All panels show results for Specific and general PIT on both percentage of responses and hand-grip. Bars indicate standard error of the mean. * p <.05; **p <.01.

Two separate mixed-effects models were used, with CS (unpaired/neutral) as independent variable and percentage of responses or hand-grip as dependent variables. Subjects were modelled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified.

Results on percentage of responses showed no difference (p = .5) between congruent and incongruent responses (means are reported in Fig. 6.3-A).

Results on hand grip showed a significant difference (F(1, 37) = 37.18; two-tailed p < .001; part. η2 = .51) between unpaired and neutral responses, the first being higher than the second (means are reported in Fig. 6.3-A).

		Percentage of responses			Hand-grip		
		F (1, 37)	p	part-η2	F (1, 37)	p	part-η2
PIT	Specific	2.56	.06	.12	.61	.44	.02
	General	.2	.6	.01	26.19	<.001	.42
PIT-dev	Specific	5.67	.02	.13	.13	.72	0
	General	.43	.52	.01	37.18	<.001	.51
PITo	Specific	14.06	<.001	.28	.86	.36	.02
	General	0	.9	0	45.03	<.001	.55
PITo-dev	Specific	13.85	<.001	.27	.18	.67	.01
	General	.04	.83	0	65.87	<.001	.64

Table 6.1
Specific and General PIT results in all tasks. PIT = Pavlovian-to-Instrumental Transfer; PIT-dev = PIT under devaluation; PITo = PIT after overtraining; PITo-dev = PIT after overtraining under devaluation.

In sum, general transfer effect was observed when considering the vigour (hand-grip) of responses, but not when considering the frequency of such responses (percentage of responses).

Crucially, this result indicates the presence of a double dissociation between specific and general aversive transfer effects.

All the following transfer tasks (PIT-dev, PITo, PITo-dev) were analysed following the same criteria and showed a similar trend. All results are described in detail in Table 6.1 and means are reported in Fig. 6.3-B-C-D.

PIT effect across tasks

To directly test how transfer effect on both percentage of responses and hand-grip was modulated by the experimental manipulations of devaluation and instrumental overtraining, an index of transfer effect was computed.

For specific transfer, the index was calculated on the percentage of responses (i.e., using the dependent variable that expressed a specific transfer effect) by subtracting incongruent from congruent responses.

For general transfer, the index was calculated on the hand-grip (i.e., using the dependent variable that expressed an effect) by subtracting responses during neutral trials (neutral) from responses during aversive trials (unpaired).

Two separate mixed-effects models were used, with Task (PIT/PIT-dev/PITo/PITo-dev) as independent variable and specific or general transfer index as dependent variables. Subjects were modelled as a random effect. Assumptions of normal distribution, independence of residuals and sphericity were verified.

Results on specific transfer showed a significant main effect of Task ($F(2.3, 85.19) = 3.53$; two-tailed $p = .03$; part. $\eta 2 = .09$) (Fig. 6.4-A). Bonferroni-corrected post-hoc analysis revealed a

significant difference (p = .04) between PIT (mean =0.07; sd =0.26) and PITo (mean =0.16; sd =0.26) and between PIT and PITo-dev (mean =0.17; sd =0.27) (p > .03).

Results on general transfer index reported no significant differences between the tasks (p = .2) (Fig. 6.4-B).

These results indicates that specific transfer was enhanced by instrumental overtraining, but not by reinforcer devaluation, and, conversely, general transfer was not affected by neither instrumental overtraining nor reinforcer devaluation.

Sensitivity to punishments and rewards: correlation with BIS/BAS inventory

To further investigate PIT effect, a correlation with sensitivity to punishments and rewards, as captured from the Behavioural Inhibition/Activation System (BIS/BAS) inventory (Carverand White, 1994) was tested.

Specific and General PIT indexes (obtained during the first time PIT task was performed) were correlated with both BIS and BAS subscales separately.

A significant correlation between General PIT and BIS was found (r =.34; p =.04). All other correlations were not significant (ps >.1).

This result indicates that stronger motivation to avoid potential punishments is linked to a stronger sensitivity to such punishments.

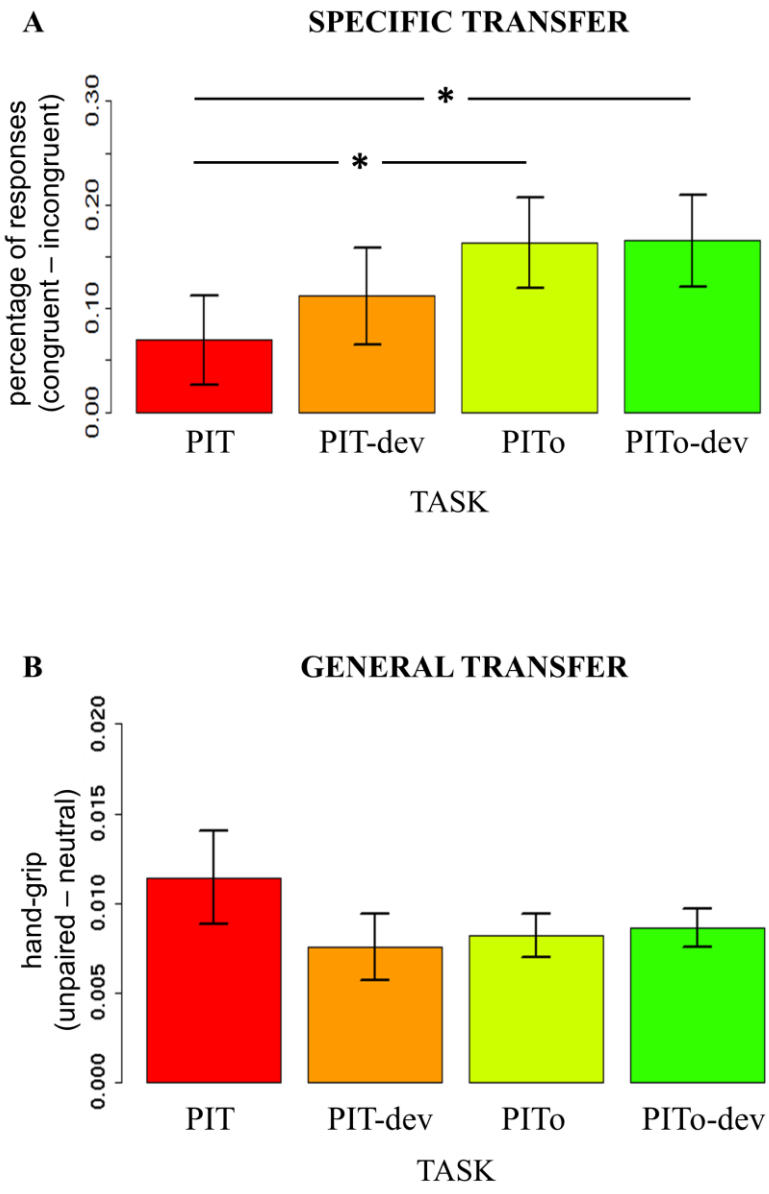

Figure 6.4
Pavlovian-to-Instrumental Transfer effect across tasks.
Panel A shows an index of the Specific PIT effect calculated on percentage of responses (Congruent-Incongruent). Panel B shows an index of the General PIT effect calculated on hand-grip (Congruent-Incongruent). PIT = Pavlovian-to-Instrumental Transfer; PIT-dev = PIT under devaluation; PITo = PIT after overtraining; PITo-dev = PIT after overtraining under devaluation. Bars indicate standard error of the mean. * p <.05.

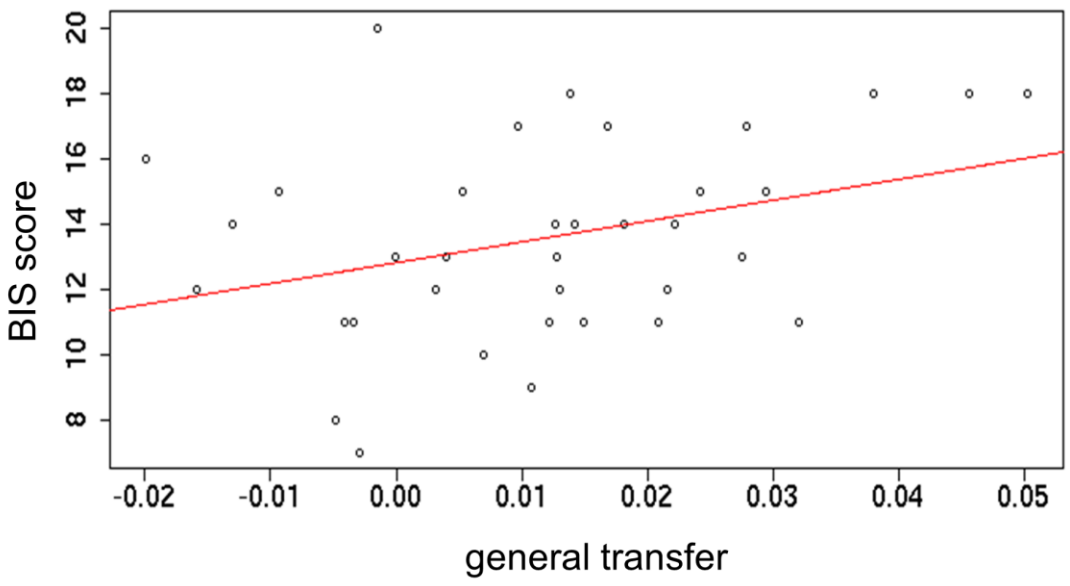

Figure 6.5
Correlation between the index of General PIT effect and the Behavioural Inibition System (BIS) scale.

Conclusion

A double dissociation between specific and general aversive transfer effects was reported for the first time in humans. Specific transfer was expressed in the percentage of responses but not in the vigour of such responses; whereas the opposite pattern was observed for general transfer, where the effect was captured by the vigour of the responses but not percentage of responses. Moreover, specific transfer was enhanced by instrumental overtraining, but not by reinforcer devaluation. General transfer was not affected by neither instrumental overtraining nor reinforcer devaluation.

A positive correlation between general transfer and sensitivity to punishments (as measured by the BIS scale) was found.

DISCUSSION

The role of mediofrontal negativity in prediction error signalling

Unexpected omission of aversive events

RL-ERN theory (Holroyd ad Coles, 2002) holds that mediofrontal negativities (such as, ERN and mediofrontal negativity) signal worse than expected outcomes, while the PRO-model (Alexander and Brown, 2011) proposes that mediofrontal activity can be explained as unexpected non-occurrence signals, consequent to the unexpected omission of both positive and negative outcomes. Many studies reported the presence of mediofrontal negativities after unexpected negative outcomes (Crowley et al., 2009; Chase et al., 2010; Holroyd et al., 2003; 2007), but only few studies explored the presence of such ERPs after unexpected positive outcomes (Oliveira et al 2007; Talmi et al., 2013; Ferdinand et al., 2012). However, the unexpected non-occurrence hypothesis postulated by the PRO-model has never been directly tested. The aim of this study was to investigate mediofrontal ERPs associated with the unexpected omission of an aversive outcome.

In the first study here presented, the unexpected omission of a predicted shock generated a stronger frontocentral negativity, signalling unexpected omission.

Although the current this result supports PRO-model, it also challenges the idea that unexpected omission is specifically related to action-outcomes. Since a Pavlovian aversive conditioning paradigm was used, no action was required during the task. Thus, no action-outcome association

was possible and the acquired expectancy about shock-delivery was only attributable to a stimulus-outcome association. The ERP results reported here clearly indicate that mediofrontal activity can be also triggered by stimulus-outcome expectancy violations. Hence, as also reported by other studies (Talmi et al., 2013; Yeung et al., 2004; Donkers et al., 2005), mPFC seems to code for unexpected outcomes even when there is no need for action.

An important contribution of this study to the literature is the direct demonstration that the observed neural signal to unexpected omission is conveyed by the absence of the expected event, rather than by the elaboration of a different stimulus. In the present study the shock absence itself acts as the feedback, while previous studies (Oliveira et al 2007; Talmi et al., 2013; Ferdinand et al., 2012) operationalized the unexpectedness by presenting a stimulus different than the one expected. The presence of a neural response to the absence of a stimulus leads unambiguously to the conclusion that an event was expected and that its absence produced a prediction error signal. Thus, the present results strongly support the PRO-model conception of mPFC coding for the unexpected omission of a predicted event.

The present study tested a central claim of the PRO-model, namely that mPFC is specifically involved in coding unexpected omission events. Still, it remains to be clarified whether the role of mPFC is selectively related to the unexpected omission of predicted outcomes, as demonstrated here, or whether it also includes the unexpected occurrence of unpredicted outcomes, as some studies seem to indicate (Oliveira et al 2007; Talmi et al., 2013; Ferdinand et al., 2012; Jessup et al., 2010). Although it would be functionally inefficient to have separate processors for detecting two kinds of unexpectedness, it could make sense considering the

different strategic processing adjustments that these two events could drive[42]. For example, if the prediction error signal concerns an unexpectedly omitted punishment, the behaviourally relevant information to learn could be to stop avoiding the associated cue or choice; while, if it signals an unexpectedly delivered punishment, it would be relevant to learn to avoid the associated cue or choice. The inverse relation should be valid if the unexpected outcome is a reward omission or delivery. PRO-model's assumptions about the role of mPFC in the detection of unexpected events could be strengthened by addressing this issue.

Timing of salient outcomes

Learning about the time of events in the environment is a fundamental adaptive behaviour (Buhusi and Meck, 2005; Cohen, 2011). Like other types of associative learning, it is deemed to depend critically on detecting mismatches (i.e., Prediction Errors) between expected and actual experience (Rescorla and Wagner, 1972; Mackintosh, 1975; Pearce and Hall, 1980; Sutton and Barto, 1998; Holroyd and Coles, 2002).

Overall, the findings presented in the second study provide novel evidence about neural activity in ACC/mPFC as it is reflected in the mediofrontal negativity. They reveal that this component tracks the timing of salient events, and reports an error signal when an aversive outcome delivery is shifted away from an expected time.

Although some experimental work has attempted to address whether the mediofrontal negativity is modulated by the timing of outcome, no prior studies, to our knowledge, have examined whether this ERP component is sensitive to temporal aspect of outcome

prediction by showing prediction error responses with surprising changes (i.e., unexpectedness) in outcome timing, as proposed by the PRO model.

In previous studies, temporal expectations or predictions were not varied: either external cues signalled in advance whether timing preceding outcome delivery would be short or long (Wang et al., 2014; Peterburs et al., 2015), or short and long outcome timings occurred in 50% of cases, and thus were equally expected (Weinberg, et al., 2012). By contrast, the current study directly manipulated the probability of occurrence of an outcome at a given time (i.e., temporal expectation), thus making it possible to contrast expected vs. unexpected timing information.

Our observation that the mediofrontal negativity is highly sensitive to the unexpected variation in the temporal interval preceding an outcome, is in agreement with electrophysiological recordings of midbrain dopamine neurons responding to errors in the temporal prediction of outcome (Hollerman and Schultz, 1998; Fiorillo et al., 2003, 2008; Bromberg-Martin et al., 2010) and targeting prefrontal cortical regions (Williams and Goldman-Rakic, 1993). As such, the present findings provide a strong functional link between mediofrontal negative ERP component and the neural data about midbrain dopaminergic signals carrying timing information of salient events (Holroyd and Coles, 2002). Furthermore, our findings are in line with fMRI studies reporting activity in frontal regions and VTA (a crucial source of the dopamine innervation of the prefrontal cortex) modulated by predictions of the timing of outcome delivery (Forster and Brown, 2011; Klein-Flügge et al., 2011).

Crucially, the observed pattern of mediofrontal negativity for outcomes occurring at unexpected times is consistent with the predictions of the PRO model that ACC/mPFC signal violations of predicted outcome timing (Alexander and Brown, 2011; 2014). Like many previous models (Kennerley et al., 2006; Rudebeck et al., 2006; Matsumoto et al., 2007), the PRO model posits that mPFC can learn to predict outcomes by using a reinforcement learning algorithm based on prediction errors (surprising outcomes). However, in a contrast to previous accounts, the PRO model also maintains that the mPFC encodes multiple independent predictions in parallel, which not only concern the likelihood of occurrence of outcomes but also their timing. Consistent with the PRO model predictions, we found that the scalp potentials overlying the ACC/mPFC activity were significantly more negative when outcomes occurred at unexpected times, even if the outcomes themselves were expected, suggesting that ACC/mPFC reports errors in the prediction of the timing of outcomes independently of their probability of occurrence.

Given that mPFC is thought to play a critical role in action selection (Posner and Petersen, 1990; Holroyd and Coles, 2002), earlier formulation of the PRO model focussed on the role of mPFC in predicting the outcome of an action (Alexander and Brown, 2011). However, to account for several findings implicating mPFC in paradigms not explicitly related to response generation, more recent simulation works (Alexander and Brown, 2014; 2015) have proposed an extended PRO model of mPFC, according to which this region may encode predictions and prediction errors even when the predicted outcomes are not contingent on prior actions. In close agreement with the extended implementation of the PRO model, the present findings reveal that the

mediofrontal negativity encodes the unexpected timing of outcomes during a task with no action requirement (i.e., pavlovian aversive conditioning). These results appear consistent with a variety of recent empirical evidences suggesting that ACC/mPFC is critically implicated in learning the casual structure of the environment (Dayan and Niv, 2008), allowing an individual to predict what follows and when.

Notably, the PRO model maintains that mPFC learns to predict and signal discrepancies between expected and actual events of any stimulus and outcome a subject may encounter during the course of an experimental task. In striking contrast with this prediction, however, the current results reveal that ACC/mPFC activation is sensitive to the salience of the outcome, as a higher amplitude of mediofrontal negativity was selectively observed for aversive outcomes occurring at unexpected timings, as compared to neutral outcomes similarly delivered at unexpected timings. These findings are highly consistent with earlier EEG works (Talmi et al., 2013; Maier et al., 2012; 2013; Lou et al., 2015; Ganushchak et al., 2008; Hajcak et al., 2005), suggesting that the mediofrontal negativity expresses a prediction error related to the motivational salience of outcomes.

It is worth noting that the way salience is defined may vary between studies. While in some studies the term salience defines the magnitude of the outcome and provides a measure of the degree of motivation or drive with which the outcome is approached or avoided (Talmi et al., 2013; Roesch and Olson, 2004; Matsumoto and Hikosaka, 2009: Kahnt et al., 2014), in others, salience refers to a large category of arousing and alerting events that signal the need for cognitive and behavioural changes (Redgrave et al., 1999; Horvitz, 2000; Zink et al, 2003),

including and extending beyond reinforcing events. Unfortunately, the present study fails in separating magnitude and arousal-based concepts of salience, since aversive stimuli are both arousing and reinforcing events. Regardless of the particular definition, however, the salience of an event determines the amount of attentional resources and control processes that are engaged in order to mobilize the most adaptive responses (i.e., behavioural, autonomic, cognitive; see Seeley et al., 2007; Shackman et al., 2011; Ham et al., 2013). In this context, aversive events (as electrical shocks used in the present experiment) can be regarded as salient, since the defensive preparatory adjustments elicited can help an organism to minimize the impact of negative consequences (Ploghaus et al., 2003). Therefore, learning the precise timing of delivery of a salient outcome constitutes a critical feature of an accurate prediction, whereas the timing of occurrence of a neutral outcome does not have the same relevance.

To conclude, many current theories view mPFC as crucially implicated in processing outcomes, and signalling discrepancies between actual and expected outcomes (Holroyd and Coles, 2002; Alexander and Brown, 2011). The present findings on the mediofrontal negativity, an ERP component generated by activity in anterior cingulate cortex and adjacent mPFC, provide further evidence for this view. We found that the mediofrontal negativity is sensitive to the timing of salient outcomes by reporting prediction error responses with surprising changes in outcome timing. These signals may be used to learn about temporal contingencies in the environment, and make preparatory adjustments in advance of the occurrence of motivationally significant events (Shackman et al., 2011). These findings represent a critical step towards uncovering the link between midbrain dopaminergic activity following violations of the expected timing of outcome and mediofrontal ERP signals of prediction error.

Reward processing in Parkinson's Disease associated psychosis

The altered activation found in the temporal lobe during both Reward Anticipation and Reward Feedback (right and left lobes, respectively) adds to the idea that psychotic symptoms might be secondary to dysfunction in cortical regions that regulate the dopamine system and display progressive alterations in psychosis, such as the temporal lobe itself (Lodge and Grace, 2011). Alterations of the temporal lobe have been reported in patients affected by psychosis both when performing tasks and from postmortem studies (Tamminga et al., 2010; Weiss et al., 2003; Heckers, et al. 1998). Furthermore, abnormal activity in the temporal lobe has been directly linked with clinical measures of psychosis (Schobel et al., 2009; Molina et al., 2003). This theoretical framework probably deserves more attention in future studies, as it sheds new light in the understanding of the illness and in the development of better pharmacological treatments for psychosis. For instance, of particular interest would be the direct investigation of the lateralization found in the present study (hypoactivation in the right temporal lobe for Reward Anticipation and left temporal lobe for Reward Feedback).

Moreover, of interest would be exploring the hypothesis that psychotic aspects of PD (especially hallucinations and paranoia) are a possible precursor of the dementia, which predominantly involves posterior visual association cortices.

This study provided new evidences about abnormal reward processing in patients with PD associated psychosis, as compared with patients presenting PD only and healthy Controls. In PD patients with psychosis, medial-prefontal, striatal and temporal areas have been found to be

hypoactive in two critical stages: assessment of the expected value of a choice and evaluation of the associated reward. Furthermore, impairments in the ability to efficiently use reward occurrence to guide their behaviour, relative to Controls, was also evidenced in the behavioural performance during the task.

These results support the idea that psychotic symptoms might take their origin from abnormalities in the hierarchical Bayesian processing of reality, as hypothesized by Fletcher and Frith (2009).

An impaired attribution of salience to rewards (higher-level system) can provide altered prior beliefs based on which inferences and beliefs are updated (lower-level system), consequently favouring an abnormal loop of interpretations and expectations about the world.

The influence of task-irrelevant Pavlovian cues on human decision

Individual differences in appetitive contexts

Motivated behaviour is characterized by a wide span of inter-individual differences in both human and non-human animals. Some recent studies highlighted a complex scenario relating ST and GT behaviours to addiction. While ST individuals are more susceptible to the influence of discrete cues, GT individuals are more influenced by contextual cues, which can motivate drug-seeking behaviour (Robinson et al., 2014). Consequently, these learning styles seem to reflect

differences in the kinds of triggers to which the individual is susceptible (e.g., discrete/contextual), rather than a propensity to addiction per se. This finding emphasizes that there are diverse pathways to addiction, and has remarkable implications for the development of personalized treatments in the future.

But what exactly is the mechanism underlying the attribution of incentive salience to discrete stimuli, such as Pavlovian cues? A large amount of evidence points to the role of dopaminergic transmission within circuits known to be involved in addiction. The core of the nucleus accumbens, for example, was reported to be involved in ST behaviour, and mediates the reinstatement of drug-seeking and drug-taking behaviour (Clark et al., 2013; Flagel et al., 2011, 2008, 2007). Furthermore, various studies have supported the involvement of the mesolimbic dopamine system in the emergence of ST behaviour. ST individuals are characterized by stronger dopaminergic gene expression and increased levels of dopamine in the nucleus accumbens (correlated with the vigour with which the CR is performed) (Flagel et al., 2008, 2007). Even if differences in basic dopaminergic levels cannot fully account for differences in dopamine responsiveness, it has been argued that higher reward-related dopamine release before conditioning might increase attribution of incentive salience to reward-related cues (Wyvell and Berridge, 2000, 2001). Additionally, Flagel and colleagues (2011) directly demonstrated that dopaminergic transmission is not involved in all forms of learning, but it is necessary for the acquisition of a sign-tracking CR, playing a crucial role in the assignment of incentive salience to reward-related cues. The same study also showed that dopaminergic prediction-error signals, coded by activity in the nucleus accumbens, are present in ST individuals, but not in GT individuals. In the present study, a similar mechanism might occur: high levels of dopamine

release might boost attribution of incentive salience to reward-related cues, increasing their ability to motivate and drive behaviour.

Future studies might further investigate individual differences in the influence of Pavlovian cues on behaviour by taking additional measures into account, such as phasic dopamine levels, psychophysiological indices (e.g., galvanic skin response and heart rate) and as neuroimaging measurements. These methods would allow better comparisons between human and non-human animal research on individual differences in ST/GT behaviour and learning styles. A general limitation in the standard PIT paradigm is that the "Sign" and the "Goal" are presented in distinct spatial locations. Thus, unrelated spatial biases in gaze direction might obscure the effect of interest. Although the analysis conducted in this study already confirmed that the present findings cannot be accounted for by any a-priori difference in spatial bias between groups, another way to control for this possibility would be to replicate the experiment with the spatial positions of the "Sign" and the "Goal" inverted in the three tasks.

In conclusion, the individual differences demonstrated here offer a promising direction for further investigating the degree to which incentive salience is attributed to environmental stimuli associated with rewards, as well as the link between this process and maladaptive behaviours, ranging from over-eating to pathological gambling and addiction (Saunders and Robinson, 2013). Further, the present findings have important implications for the treatment of impulse-control disorders. Overall, these individual differences in PIT offer new insights into the mechanisms underlying the transition from intentional to habitual/compulsive behaviour.

Double dissociation

The double dissociation found between specific transfer, captured by the percentage of responses, and general transfer, captured by the vigour of responses, might be informative about the different mechanisms behind the two effects.

Previous studies reporting outcome-specific and general transfer as measured by percentage of responses and vigour of responding, respectively, did not demonstrate a double dissociation between these two measures (e.g., Watson et al., 2016). Furthermore, the vigour of response was calculated as the average total number of key presses, whereas in the present study a more direct measure of the power exerted for each response (hand-grip) has been considered.

The presence of a double dissociation finds support in the presence of dissociable brain areas found in human and animal studies investigating outcome-specific and general transfer.

Unfortunately, only few evidences directly investigate humans.

Outcome-specific transfer correlates with activity in the dorsal striatum (putamen) (Bray et al., 2008) – believed to be implicated in instrumental learning - and ventral amygdala (Prèvost et al., 2012) – believed to be implicated in processing reinforcer value and S-O associations.

General transfer correlates with activity in the ventral striatum (nucleus accumbens) (Talmi et al., 2008) – believed to be implicated in Pavlovian learning - and dorsal amygdala (Prèvost et al., 2012; Talmi et al., 2008) – believed to be implicated in assigning incentive motivation to stimuli.

From non-human studies comes a larger amount of evidences, but results are not entirely straightforward.

Outcome-specific transfer has been reported to be impaired by lesions of the basolateral amygdala (BLA) (Corbit and Balleine, 2005; Hall et al.,2001; Holland and Gallagher, 2003), nucleus accumbens shell (NAC-shell) (Corbit et al., 2001; Corbit and Balleine, 2011) and ventral tegmental area (VTA) (Corbit et al., 2007). Contrasting evidences have been reported regarding the role of the prefrontal cortex in outcome-specific transfer (see Ostlund and Balleine, 2007; Cardinal et al., 2003; Homayounand Moghaddam, 2009).

General transfer has been reported to be impaired by lesions of the central nucleus of the amygdala (CeN) (Corbit and Balleine, 2005; Hall et al.,2001; Holland and Gallagher, 2003), nucleus accumbens core (NAC-core) (Hall et al., 2001; Corbit and Balleine, 2011), substantia nigra (SN) (El-Amamy and Holland, 2007) and ventral tegmental area (VTA) (Corbit et al., 2007). General transfer has also been reported to correlate with activity in the prefrontal cortex (Ostlund and Balleine, 2007; Corbit and Balleine, 2003; Homayounand Moghaddam, 2009) and dorsal striatum (Corbit and Janak, 2007).

It is interesting to notice that some overlapping regions have been identified in human and non-human animals, which suggest a relation across species and supports the importance of translational studies. However, more studies are needed to clarify such relationships as well as the mechanism underlying transfer in each of the species.

Devaluation

In line with the idea that transfer does not reflect an intentional goal-directed behaviour, the present experiment reported no effect of devaluation on both specific and general transfer in humans, regardless of the amount of training (thus, before and after instrumental overtraining).

Previous studies which obtained similar results with animals (Colwill and Rescorla, 1990; Rescorla, 1994; Holland, 2004; Corbit, Janak and Balleine 2007; Hogarth, Dickinson and Duka 2010; Hogarth and Chase 2011; Hogarth 2012), reported different possible interpretations.

For some authors, devaluation acts on the sensory specific properties of the reinforcer, though not reducing the motivational influence that a Pavlovian CS (associated with an reinforcer before its devaluation) can exert on instrumental responding, like in the case of transfer effect (Colwill and Motzkin, 1994; Holland, 1990; Holland, 2004).

Conversely, other authors pointed more to the idea that the decreased motivational value consequent to devaluation procedure does not impact transfer effect, as it reflects a signalling role of the CS to the appropriate instrumental responding, activated by the representation of sensory-specific (rather than motivational) properties of the reinforcer (Rescorla, 1994).

Much of the confusion might be explained by taking into account the different devaluation procedures used in literature. From a theoretical point of view, it is hard to state if devaluation selectively acts on sensory-specific or motivational properties of the reinforcer. If a food pellet is paired with a toxin, only the motivation towards that specific reward is reduced, but the general motivation to eat might not be affected. On the other hand, if devaluation is operationalized as satiation, then the overall general motivation to eat might be compromised. In the first case there

is a modulation of liking; in the second, a modulation of wanting. Thus, devaluation might selectively impact sensory-specific, motivational or even both representations of the reinforcer, depending on the procedure used. Consequently, devaluation per se cannot definitely disentangle the kind of information (sensory-specific or motivational) implicated in transfer or any other effects.

For example, previous studies that investigated the impact of devaluation on transfer effect in humans reported contrasting results. Watson and colleagues (2014) reported that food satiation did not influence both specific and general transfer effects. Allman and colleagues (2010), on the other hand, found reduced outcome-specific transfer for stimuli associated with a devalued currency. Such different results might be explained by the use of very different devaluation procedures, respectively satiety and currency deflation. As argued, the first one might have acted more on the general wanting of food (satiation for popcorn might reduce not only craving for popcorns, but also general hunger), thus impairing motivation. Whereas, the second one might have acted on the selective "liking" of the deflated currency, as compared to the more valuable one (devaluing a currency does not impair the desire to win more money, especially if there is a more valuable currency that would allow to do so), thus being more sensory-specific.

These different procedures complicates the interpretation of the results. If motivation to win money was not reduced by Allman's procedure, is it possible that the effect observed was just driven by a vivid representation of the sensory-specific features of the two reinforcers, that is known to be implicated in outcome-specific transfer?

In the present experiment devaluation was operationalized as a complete impossibility for the learned aversive reinforcers to be delivered, which should have an impact on both general fear (motivation) and aversion for those exact aversive stimuli (sensory-specific).

What can be undoubtedly stated about such a devaluation procedure is that the current value of the reinforcer is altered, thus inevitably reducing potential goal-directed responses. Consequently, actions that are not altered by that are more habit-like actions, as they are not driven by the value of the potential reinforcer.

Consequently, the absence of an effect of devaluation on both outcome-specific and general transfer effect here reported, reflects the independence of transfer from the current value of an reinforcer, demonstrating the absence goal-directed behaviour. Moreover, outcome-specific and general transfer are always observed in extinction, thus in the absence of the goal itself.

This proposed framework adds on the idea that the power exerted by Pavlovian CSs on instrumental responses does not determine a goal-directed behaviour (see also Hogarth et al., 2013).

Overtraining

A possible different mechanism underlying specific and general transfer effects might still reflected by the different sensitivity to instrumental overtraining (which increased specific transfer but did not alter general transfer) and by the presence of a double dissociation in the

measure that allowed the observation of each effect (namely, specific transfer was observed on percentage of responses only and general transfer was observed on vigour of response only).

Instrumental overtraining favours the formation of habits at the expense of goal-directed behaviours, determining a reduction of devaluation effect on the instrumental response (Adams, 1982; Adams and Dickinson, 1981; Dickinson, Balleine, Watt, Gonzalez, and Boakes, 1995; Yin and Knowlton, 2002; Tricomi, Balleine, and O'Doherty, 2009; Dayan and Berridge, 2014).

Overtraining is often interpreted as a progressive shift in control of instrumental responding from a direct R-O association to a more indirect S–R associations, which overcomes and weakens the reinforcer representation (Holland et al., 2004).

In line with this interpretation, overtraining increased outcome-specific transfer, driven by a S-O-R association, but did not alter general transfer effect, driven by an S-R association and thus unrelated to the specific reinforcer.

In a study on rats, Holland (2004) found that moderate and extensive instrumental training (7 or 20 sessions, respectively), but not minimal training (2 sessions), could elicit transfer. This result was explained as an effect of practice on the response. More specifically, the idea proposed is that early in training instrumental responses are more governed by detailed R-O association, but as training proceeds a S–R takes over, which is more susceptible to the influence of conditioned incentives (Holland, 2004).

This evidence is coherent with the idea and results here presented that transfer effect does not reflect a goal-directed influence on instrumental performance, as it is not influenced by the

current value of the reinforcer, especially after that instrumental overtraining weakens even more the outcome representation which still plays a role in outcome-specific transfer.

In conclusion, dissociable aspects of associative representations seem to underlie transfer effect, overtraining, and devaluation effects.

The importance of investigating these issues is given by the chance to achieve new insights about the mechanisms underlying the transition from intentional goal-directed actions to habitual and compulsive behaviours, which could have significant implications in the treatment of impulse-control disorders.

CONCLUDING REMARKS

Rewards and punishments play a key role in shaping human and animal learning. Thanks to their intrinsic value, by interacting with the environment we learn how to get what we need (not only in terms of food, but also socially and emotionally), how to avoid potential dangers and, most importantly, how to adapt to a changing environment (Sutton and Barto, 1998; Doya, 2008; Cohen et al., 2011; Behrens et al, 2008; Brown and Brüne, 2012).

Reinforcement learning has the potential to function as a bridge in neuroscientific research across different species and to offer new theoretical and methodological frameworks able to get new insights into many clinical disorders (Cohen et al., 2011), ranging from addiction (Baker, et al. 2011; Crean et al. 2011; Park. et al., 2010) to psychiatric diseases (Everitt and Robbins, 2005; Gold, et al. 2008; Fletcher and Frith, 2009; Berridge et al., 1998).

The present work contributed to the current scientific knowledge around this topic by exploring some of the adaptive and maladaptive implications of reinforcement learning processes, analysing its behavioural and neural correlates.

Brain signals of prediction error have been investigated to achieve new awareness of the strengths and weaknesses of the current theoretical models and hypothesis about the normal and pathological brain. Cue-driven behaviours have been analysed in bot appetitive and aversive contexts to further understand how external and internal factors can impact on human behaviour and drive choices.

Significant gaps still remain in understanding how brain networks interact between them to support reinforcement learning, which suggests the need further investigation. Critical to advancing our knowledge will be the achievement of new neurobiological models looking at our brain, mind, body and environment from a more integrated perspective.

BIBLIOGRAPHY

Adams DF. 1982. How safe is the coronary angiogram? Cardiovasc Intervent Radiol. 5:168–173.

Adams SS, Humphries RG, Mason CG. 1981. The relationship between development of ultraviolet erythema and release of prostaglandins in guinea pig skin. Agents Actions. 11:473–476.

Alexander WH, Brown JW. 2014. A general role for medial prefrontal cortex in event prediction. Front Comput Neurosci. 8:69.

Alexander WH, Brown JW. 2015. Hierarchical Error Representation: A Computational Model of Anterior Cingulate and Dorsolateral Prefrontal Cortex. Neural Comput. 27:2354–2410.

Alexander WH, Brown JW. 2011. Medial prefrontal cortex as an action-outcome predictor. Nat Neurosci. 14:1338–1344.

Allman MJ, DeLeon IG, Cataldo MF, Holland PC, Johnson AW. 2010. Learning processes affecting human decision making: An assessment of reinforcer-selective Pavlovian-to-instrumental transfer following reinforcer devaluation. J Exp Psychol Anim Behav Process. 36:402–408.

Balleine B. 1994. Asymmetrical interactions between thirst and hunger in Pavlovian-instrumental transfer. Q J Exp Psychol B. 47:211–231.

Balleine BW, Dickinson A. 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology. 37:407–419.

Bell AJ, Sejnowski TJ. 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7:1129–1159.

Bernacer J, Ph D, Corlett PR, Ramachandra P, Turner DC, Psych MRC. 2070. Methamphetamine-Induced Disruption of Frontostriatal Reward Learning Signals: Relation to Psychotic Symptoms, (November), . http://doi.org/10.1176/appi.ajp..1978.

Berridge KC, Robinson TE. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev. 28:309–369.

Berridge KC. 2001. Reward learning: Reinforcement, incentives, and expectations. In: The Psychology of Learning and Motivation. p. 223–278.

Berridge KC, Robinson TE. 2003. Parsing reward. Trends Neurosci.

Bettuzzi S. 2009. Introduction. Adv Cancer Res.

Blundell P, Hall G, Killcross S. 2001. Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. J Neurosci. 21:9018–9026.

Boakes RA, In HMB, Davis H, Hillsdale NJ, Bray S, Rangel A, Shimojo S, Balleine B, Doherty JP. 2008. Performance on learning to associate a stimulus with positive reinforcement. H eds Operant Interact Erlbaum pp OThe neural Mech underlying Influ Pavlov cues Hum Decis Mak J Neurosci Off J Soc Neurosci 58616 doi101523JNEUROSCI089708. 28:67–97.

Boksem M a S, Tops M, Kostermans E, De Cremer D. 2008. Sensitivity to punishment and reward omission: evidence from error-related ERP components. Biol Psychol. 79:185–192.

Boksem M a S, Tops M, Wester AE, Meijman TF, Lorist MM. 2006. Error-related ERP components and individual differences in punishment and reward sensitivity. Brain Res. 1101:92–101.

Boureau Y-L, Dayan P. 2011. Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology. 36:74–97.

Bray S, Rangel A, Shimojo S, Balleine B, O'Doherty JP. 2008. The neural mechanisms underlying the influence of pavlovian cues on human decision making. J Neurosci. 28:5861–5866.

Cannon RL. 2012. Low Resolution Brain Electromagnetic Tomography (LORETA). Basic concepts and clinical applications. BMED Press.

Cardinal RN, Parkinson JA, Marbini HD, Toner AJ, Bussey TJ, Robbins TW, Everitt BJ. 2003. Role of the anterior cingulate cortex in the control over behavior by Pavlovian conditioned stimuli in rats. Behav Neurosci. 117:566–587.

Carver CS, White TL. 1994. Behavioural inhibition, behavioural activation, and affective responses to impending ieward and punishment: the BIS/BAS scales. J Pers Scocial Psychol. 67:319–333.

Carver CS, White TL, J ,. 1994. Behavioral inhibition, behavioral activation, and affective responses to impending ieward and punishment: the BIS/BAS scales. J Personal Scocial Psychol. 67:319–333.

Chase HW, Swainson R, Durham L, Benham L. 2010. Feedback-related negativity codes prediction error but not behavioural adjustment during probabilistic reversal learning. J Cogn Neurosci. 23:936–946.

Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ. 2013. Dopamine restores reward prediction errors in old age. Nat Neurosci. 16:648–653.

Clark JJ, Collins AL, Sanford CA, Phillips PEM. 2013. Dopamine encoding of Pavlovian incentive stimuli diminishes with extended training. J Neurosci. 33:3526–3532.

Colwill RM, Motzkin DK. 1994. Encoding of the unconditioned stimulus in Pavlovian conditioning. Anim Learn Behav. 22:384–394.

Colwill RM, Rescorla RA. 1990. Effect of reinforcer devaluation on discriminative control of instrumental behavior. J Exp Psychol Anim Behav Process. 16:40–47.

Colwill RM, Rescorla RA. 1988. The role of response-reinforcer associations increases throughout extended instrumental training. Anim Learn Behav. 16:105–111.

Corbit LH, Balleine BW. 2003. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. J Exp Psychol Anim Behav Process. 29:99–106.

Corbit LH, Balleine BW. 2003. The role of prelimbic cortex in instrumental conditioning. Behav Brain Res. 146:145–157.

Corbit LH, Balleine BW. 2005. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. J Neurosci. 25:962–970.

Corbit LH, Balleine BW. 2011. The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. J Neurosci. 31:11786–11794.

Corbit LH, Janak PH. 2007. Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. J Neurosci. 27:13977–13981.

Corbit LH, Janak PH, Balleine BW. 2007. General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. Eur J Neurosci. 26:3141–3149.

Corbit LH, Muir JL, Balleine BW. 2001. The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. J Neurosci. 21:3251–3260.

Corlett PR, Fletcher PC. 2012. The neurobiology of schizotypy: fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. Neuropsychologia. 50:3612–3620.

Crowley MJ, Wu J, Bailey CA, Mayes LC. 2009. Bringing in the negative reinforcements: the avoidance feedback-related negativity. Neuroreport. 20:1513–1517.

Cunningham CL, Patel P. 2007. Rapid induction of Pavlovian approach to an ethanol-paired visual cue in mice. Psychopharmacology (Berl). 192:231–241.

Dawson ME, Schell AM, Filion DL, Cacioppo LG, Tassinary GG. 1990. The electrodermal system. In: Cacioppo JT,, Tassinary LG,, Berntson GG, editors. J.T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), Handbook of psychophysiology. Cambridge: Cambridge University Press. Cambridge: Cambridge University Press. p. 295–334.

Dayan P, Berridge KC. 2014. Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. Cogn Affect Behav Neurosci. 14:473–492.

Dayan P, Huys QJM. 2008. Serotonin, Inhibition, and Negative Mood. PLoS Comput Biol. 4:e4.

Dayan P, Niv Y, Seymour B, Daw ND. 2006. The misbehavior of value and the discipline of the will. Neural Netw. 19:1153–1160.

de Leeuw M, Kahn RS, Vink M. 2015. Fronto-striatal Dysfunction During Reward Processing in Unaffected Siblings of Schizophrenia Patients. Schizophr Bull. 41:94–103.

de Wit S, Dickinson A. 2009. Associative theories of goal-directed behaviour: a case for animal–human translational models. Psychol Res Psychol Forsch. 73:463–476.

de Wit S, Dickinson A. 2009. Associative theories of goal-directed behaviour: a case for animal-human translational models. Psychol Res. 73:463–476.

de Wit S, Watson P, Harsay HA, Cohen MX, van de Vijver I, Ridderinkhof KR. 2012. Corticostriatal Connectivity Underlies Individual Differences in the Balance between Habitual and Goal-Directed Action Control. J Neurosci. 32:12066–12075.

Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK. 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J Neurosci. 25:11730–11737.

Delamater AR. 1990. Effects of several extinction treatments upon the integrity of Pavlovian stimulus-outcome associations. Anim Learn Behav doi103758BF0315. 24:437–449.

Delamater AR. 1980. Outcome-selective effects of intertrial reinforcement in a Pavlovian appetitive conditioning paradigm with rats. Anim Learn Behav doi103758BF0313. 23:31–39.

Delamater AR, Holland PC. 2008. The influence of CS-US interval on several different indices of learning in appetitive conditioning. J Exp Psychol Anim Behav Process. 34:202–222.

Delorme A, Makeig S. 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134:9–21.

Delorme A, Sejnowski T, Makeig S. 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage. 34:1443–1449.

Dickinson A, Balleine B, In H. 2001. The role of learning in the operation of motivational systems.

Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA. 1995. Motivational control after extended instrumental training. Anim Learn Behav. 23:197–206.

Dickinson AD, Balleine BW. 1994. Motivational control of goal-directed action. Anim Learn Behav. 22:1–18.

Dickinson AD, Smith J, Mirenowicz J. 2000. Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav Neurosci. 114:468–483.

Dolan RJ, Dayan P. 2013. Goals and habits in the brain. Neuron. 80:312–325.

Donkers FCL, Nieuwenhuis S, van Boxtel GJM. 2005. Mediofrontal negativities in the absence of responding. Brain Res Cogn Brain Res. 25:777–787.

Doya K. 2008. Modulators of decision making. Nat Neurosci. 11:410–416.

Dunsmoor JE, Labar KS. 2012. Brain activity associated with omission of an aversive event reveals the effects of fear learning and generalization. Neurobiol Learn Mem. 97:301–312.

Egner T. 2011. Surprise! A unifying model of dorsal anterior cingulate function? Nat Neurosci. 14:1219–1220.

El-Amamy H, Holland PC. 2007. Dissociable effects of disconnecting amygdala central nucleus from the ventral tegmental area or substantia nigra on learned orienting and incentive motivation. Eur J Neurosci. 25:1557–1567.

Esslinger C, Englisch S, Inta D, Rausch F, Schirmbeck F, Mier D, Kirsch P, Meyer-Lindenberg A, Zink M. 2012. Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. Schizophr Res. 140:114–121.

Estes WK. 1943. Discriminative conditioning. I A Discrim Prop Cond Anticip J Exp Psychol doi101037h0058316. 32:150–155.

Estes WK. 1948. Discriminative conditioning; effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. J Exp Psychol. 38:173–177.

Everitt BJ, Robbins TW. 2005. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci. 8:1481–1489.

Falkenstein M, Hohnsbein J, Hoormann J, Blanke L. 1990. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. Psychophysiological brain Res. 1:192–195.

Ferdinand NK, Mecklinger A, Kray J, Gehring WJ. 2012. The processing of unexpected positive response outcomes in the mediofrontal cortex. J Neurosci. 32:12087–12092.

Fiorillo CD, Newsome WT, Schultz W. 2008. The temporal precision of reward prediction in dopamine neurons. Nat Neurosci. 11:966–973.

Fiorillo CD, Tobler PN, Schultz W. 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science. 299:1898–1902.

Flagel SB, Akil H, Robinson TE. 2009. Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. Neuropharmacology. 56 Suppl 1:139–148.

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers C a, Clinton SM, Phillips PEM, Akil H. 2011. A selective role for dopamine in stimulus-reward learning. Nature. 469:53–57.

Flagel SB, Watson SJ, Akil H, Robinson TE. 2008. Individual differences in the attribution of incentive salience to a reward-related cue: influence on cocaine sensitization. Behav Brain Res. 186:48–56.

Flagel SB, Watson SJ, Robinson TE, Akil H. 2007. Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. Psychopharmacology (Berl). 191:599–607.

Fletcher PC, Frith CD. 2009. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia., Nature Reviews Neuroscience.

Forster SE, Brown JW. 2011. Medial prefrontal cortex predicts and evaluates the timing of action outcomes. Neuroimage. 55:253–265.

Fusar-Poli P, Meyer-Lindenberg A. 2013. Striatal presynaptic dopamine in schizophrenia, part II: meta-analysis of [(18)F/(11)C]-DOPA PET studies. Schizophr Bull. 39:33–42.

Gallagher DA, Schrag A. 2012. Psychosis, apathy, depression and anxiety in Parkinson's disease. Neurobiol Dis. 46:581–589.

Ganushchak LY, Schiller NO. 2008. Motivation and semantic context affect brain error-monitoring activity: an event-related brain potentials study. Neuroimage. 39:395–405.

Garofalo S, di Pellegrino G. 2015. Individual differences in the influence of task-irrelevant Pavlovian cues on human behavior. Front Behav Neurosci. 9:163.

Garofalo S, Maier ME, di Pellegrino G. 2014. Mediofrontal negativity signals unexpected omission of aversive events. Sci Rep. 4:4816.

Gehring WJ, Goss B, Coles MG, Meyer DE, Donchin E. 1993. A neural system for error detection and compensation. Psychol Sci. 4:385–390.

Gehring WJ, Willoughby AR. 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. Science. 295:2279–2282.

Gibb WR, Lees AJ. 1988. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry. 51:745–752.

Gillan CM, Morein-Zamir S, Urcelay GP, Sule A, Voon V, Apergis-Schoute AM, Fineberg N a, Sahakian BJ, Robbins TW. 2014. Enhanced avoidance habits in obsessive-compulsive disorder. Biol Psychiatry. 75:631–638.

Gläscher J, Daw N, Dayan P, O'Doherty JP, Doherty JPO, Glascher J, Daw N, Dayan P, Doherty JPO, Gläscher J, Daw N, Dayan P, Doherty JPO. 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 66:585–595.

Gratton G, Coles MG, Donchin E. 1983. A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol. 55:468–484.

Gray JX, Rolfe BG. 1990. Exopolysaccharide production in Rhizobium and its role in invasion. Mol Microbiol. 4:1425–1431.

Griffiths O, Langdon R, Pelley ME, Coltheart M, Le Pelley ME, Coltheart M. 2014. Le & Delusions and prediction error: re-examining the behavioural evidence for disrupted error signalling in delusion formation., Cognitive neuropsychiatry.

Grimm O, Heinz A, Walter H, Kirsch P, Erk S, Haddad L, Plichta MM, Romanczuk-Seiferth N, Pöhland L, Mohnke S, Mühleisen TW, Mattheisen M, Witt SH, Schäfer A, Cichon S, Nöthen M, Rietschel M, Tost H, Meyer-Lindenberg A. 2014. Striatal response to reward anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. JAMA psychiatry. 71:531–539.

Hajcak G. 2012. What we've learned from mistakes: insights from error-related brain activity. Assoc Psychol Sci. 21:101–106.

Hajcak G, McDonald N, Simons RF. 2004. Error-related psychophysiology and negative affect. Brain Cogn. 56:189–197.

Hajcak G, Moser JS, Yeung N, Simons RF. 2005. On the ERN and the significance of errors. Psychophysiology. 42:151–160.

Hajcak G, Moser JS, Holroyd CB, Simons RF. 2006. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol Psychol. 71:148–154.

Hall J, Parkinson JA, Connor TM, Dickinson A, Everitt BJ. 2001. Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating pavlovian influences on instrumental behaviour. Eur J Neurosci. 13:1984–1992.

Hauser TU, Iannaccone R, Stämpfli P, Drechsler R, Brandeis D, Walitza S, Brem S. 2014. The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. Neuroimage. 84:159–168.

Hebb DO. 1949. The organization of behaviour. New York: Wiley and Sons.

Heckers S, Rauch SL, Goff D, Savage CR, Schacter DL, Fischman AJ, Alpert NM. 1998. Impaired recruitment of the hippocampus during conscious recollection in schizophrenia. Nat Neurosci. 1:318–323.

Heinz A. 2002. Dopaminergic dysfunction in alcoholism and schizophrenia--psychopathological and behavioral correlates. Eur Psychiatry. 17:9–16.

Helden J Van Der, Boksem MAS. 2010. The importance of failure: feedback-related negativity predicts motor learning efficiency. Cereb Cortex. 20:1596–15603.

Hoehn MM, Yahr MD. 1967. Parkinsonism: onset, progression and mortality. Neurology. 17:427–442.

Hogarth L, Balleine BW, Corbit LH, Killcross S. 2013. Associative learning mechanisms underpinning the transition from recreational drug use to addiction. Ann N Y Acad Sci. 1282:12–24.

Hogarth L, Chase HW. 2011. Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. J Exp Psychol Anim Behav Process. 37:261–276.

Hogarth L, Dickinson A, Duka T. 2010. The associative basis of cue-elicited drug taking in humans. Psychopharmacology (Berl). 208:337–351.

Hogarth L, Dickinson A, Wright A, Kouvaraki M, Duka T, Hogarth, Lee; Dickinson, Anthony; Wright, Alexander; Kouvaraki, Mariangela; Duka T, Hogarth L, Dickinson A, Wright A, Kouvaraki M, Duka T. 2007. The role of drug expectancy in the control of human drug seeking. J Exp Psychol Anim Behav Process. 33:484–496.

Hogarth L, Field M, Rose AK. 2013. Phasic transition from goal-directed to habitual control over drug-seeking produced by conflicting reinforcer expectancy. Addict Biol. 18:88–97.

Holland PC. 1990. Event representation in Pavlovian conditioning: image and action. Cognition. 37:105–131.

Holland PC. 2004. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. J Exp Psychol Anim Behav Process. 30:104–117.

Holland PC, Gallagher M. 2003. Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer. Eur J Neurosci. 17:1680–1694.

Holland PC, Petrovich GD, Gallagher M. 2002. The effects of amygdala lesions on conditioned stimulus-potentiated eating in rats. Physiol Behav. 76:117–129.

Hollerman JR, Schultz W. 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci. 1:304–309.

Holmes NM, Marchand AR, Coutureau E. 2010. Pavlovian to instrumental transfer: a neurobehavioural perspective. Neurosci Biobehav Rev. 34:1277–1295.

Holroyd CB, Coles MGH. 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev. 109:679–709.

Holroyd CB, Krigolson OE. 2007. Reward prediction error signals associated with a modified time estimation task. Psychophysiology. 44:913–917.

Holroyd CB, Krigolson OE, Baker R, Lee S, Gibson J. 2009. When is an error not a prediction error? An electrophysiological investigation. Cogn Affect Behav Neurosci. 9:59–70.

Holroyd CB, Larsen JT, Cohen JD. 2004. Context dependence of the event-related brain potential associated with reward and punishment. Psychophysiology. 41:245–253.

Holroyd CB, Nieuwenhuis S, Yeung N, Cohen JD. 2003. Errors in reward prediction are reflected in the event-related brain potential. Neuroreport. 14:2481–2484.

Homayoun H, Moghaddam B. 2009. Differential representation of Pavlovian-instrumental transfer by prefrontal cortex subregions and striatum. Eur J Neurosci. 29:1461–1476.

Horvitz JC. 2000. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience.

Jessup RK, Busemeyer JR, Brown JW. 2010. Error Effects in anterior cingulate cortex reverse when error likelihood is high. J Neurosci. 30:3467–3472.

Juckel G, Schlagenhauf F, Koslowski M, Wüstenberg T, Villringer A, Knutson B, Wrase J, Heinz A. 2006. Dysfunction of ventral striatal reward prediction in schizophrenia. Neuroimage. 29:409–416.

Jutten C, Herault J. 1991. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. Signal Processing. 24:1–10.

Kapur S. 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry. 160:13–23.

Kaye H, Pearce JM. 1984. The strength of the orienting response during Pavlovian conditioning. J Exp Psychol Anim Behav Process. 10:90–109.

Kehagia AA, Murray GK, Robbins TW. 2010. Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr Opin Neurobiol. 20:199–204.

Klein-Flügge MCC, Hunt LTT, Bach DRR, Dolan RJJ, Behrens TEJEJ. 2011. Dissociable reward and timing signals in human midbrain and ventral striatum. Neuron. 72:654–664.

Kobayashi S, Schultz W. 2008. Influence of reward delays on responses of dopamine neurons. J Neurosci. 28:7837–7846.

Konorski J. 1987. Integrative activity of the brain: an interdisciplinary approach.

Kreussel L, Hewig J, Kretschmer N, Hecht H, Coles MGH, Miltner WHR. 2012. The influence of the magnitude, probability, and valence of potential wins and losses on the amplitude of the feedback negativity. Psychophysiology. 49:207–219.

Leone N, Pellicano R, Brunello F, Rizzetto M, Ponzetto A. 2002. Elevated serum chromogranin A in patients with hepatocellular carcinoma. Clin Exp Med. 2:119–123.

Lodge DJ, Grace AA. 2011. Hippocampal dysregulation of dopamine system function and the pathophysiology of schizophrenia. Trends Pharmacol Sci. 32:507–513.

Lou B, Hsu W-YW-Y, Sajda P. 2015. Perceptual Salience and Reward Both Influence Feedback-Related Neural Activity Arising from Choice. J Neurosci. 35:13064–13075.

Lovibond PF. 1981. Appetitive Pavlovian-instrumental interactions: effects of inter-stimulus interval and baseline reinforcement conditions. Q J Exp Psychol B. 33:257–269.

Lovibond PF, Colagiuri B. 2013. Facilitation of voluntary goal-directed action by reward cues. Psychol Sci. 24:2030–2037.

Magosso E, Forcelli V, Garofalo S, di Pellegrino G, Ursino M. 2015. Event-related brain potential signaling unexpected timing of feedback: A source localization analysis. 2015 37th Annu Int Conf IEEE Eng Med Biol Soc. 618–621.

Maier ME, Steinhauser M. 2013. Updating Expected Action Outcome in the Medial Frontal Cortex Involves an Evaluation of Error Type. J Neurosci. 33:15705–15709.

Maier ME, Pellegrino G, Steinhauser M. 2012. Enhanced error-related negativity on flanker errors: Error expectancy or error significance? Psychophysiology. 49:899–908.

Maier ME, Yeung N, Steinhauser M. 2011. Error-related brain activity and adjustments of selective attention following errors. Neuroimage. 56:2339–2347.

Makeig S, J. Bell. A, Jung T-P, Sejnowski TJ. 1996. Independent Component Analysis of Electroencephalographic Data. Adv Neural Inf Process Syst. 8:145–151.

Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B, Holmes C, Collins L, Thompson P, MacDonald D, Iacoboni M, Schormann T, Amunts K, Palomero-Gallagher N, Geyer S, Parsons L, Narr K, Kabani N, Le Goualher G,

Boomsma D, Cannon T, Kawashima R, Mazoyer B. 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci. 356:1293–1322.

Medic N, Ziauddeen H, Vestergaard MD, Henning E, Schultz W, Farooqi IS, Fletcher PC. 2014. Dopamine modulates the neural representation of subjective value of food in hungry subjects. J Neurosci. 34:16856–16864.

Miller R. 1976. Schizophrenic psychology, associative learning and the role of forebrain dopamine. Med Hypotheses. 2:203–211.

Molina V, Reig S, Pascau J, Sanz J, Sarramea F, Gispert JD, Luque R, Benito C, Palomo T, Desco M. 2003. Anatomical and functional cerebral variables associated with basal symptoms but not risperidone response in minimally treated schizophrenia. Psychiatry Res. 124:163–175.

Morris RW, Vercammen A, Lenroot R, Moore L, Langton JM, Short B, Kulkarni J, Curtis J, O'Donnell M, Weickert CS, Weickert TW. 2012. Disambiguating ventral striatum fMRI-related BOLD signal during reward prediction in schizophrenia. Mol Psychiatry. 17:235,280–289.

Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell a D, Honey G, Jones PB, Bullmore ET, Robbins TW, Fletcher PC. 2008. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry. 13:267–276.

Murray GK, Clark L, Corlett PR, Blackwell AD, Cools R, Jones PB, Poustka L, BMC , Robbins TW, Poustka L. 2008. Incentive motivation in first-episode psychosis: a behavioural study., BMC psychiatry.

Nadler N, Delgado MR, Delamater AR. 2011. Pavlovian to Instrumental Transfer of Control in a Human Learning Task. Emotion. 11:1112–1123.

Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp. 15:1–25.

Nieuwenhuis S, Heslenfeld DJ, von Geusau NJ, Mars RB, Holroyd CB, Yeung N. n.d. Alting Activity in human reward-sensitive brain areas is strongly context dependent. Neuroimage. 25:1302–1309.

Nieuwenhuis S, Holroyd CB, Mol N, Coles MGH. 2004. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci Biobehav Rev. 28:441–448.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. Science (80-). 304:452–454.

Oliveira FTP, Mcdonald JJ, Goodman D. 2007. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. J Cogn Neurosci. 19:1994–2004.

Ostlund SB, Balleine BW. 2007. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J Neurosci. 27:4819–4825.

Paredes-Olay C, Abad MJF, Gámez M, Rosas JM. 2002. Transfer of control between causal predictive judgments and instrumental responding. Anim Learn Behav. 30:239–248.

Park A, Stacy M. 2009. Non-motor symptoms in Parkinson's disease.

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

Patton JH, Stanford MS, Barratt ES. 1995. Factor structure of the Barratt Impulsiveness Scale. J Clin Psychol. 51:768–774.

Perks SM, Clifton PG. 1997. Reinforcer revaluation and conditioned place preference. Physiol Behav. 61:1–5.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature. 442:1042–1045.

Peterburs J, Kobza S, Bellebaum C. 2015. Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). Psychophysiology. 1–6.

Prévost C, Liljeholm M, Tyszka JM, O'Doherty JP. 2012. Neural correlates of specific and general Pavlovian-to-Instrumental Transfer within human amygdalar subregions: a high-resolution fMRI study. J Neurosci. 32:8383–8390.

Rescorla RA. 2000. Associative changes with a random CS-US relationship. Q J Exp Psychol B. 53:325–340.

Rescorla RA. 1997. Response inhibition in Extinction. Q J Exp Psychol B. 50:238–252.

Rescorla RA. 1994. Transfer of instrumental control mediated by a devalued outcome. Anim Learn Behav.

Rescorla RA. 1994. Control of instrumental performance by Pavlovian and instrumental stimuli. J Exp Psychol Anim Behav Process. 20:44–50.

Rescorla RA, Solomon RL. 1967. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. Psychol Rev. 74:151–182.

Robinson TE, Yager LM, Cogan ES, Saunders BT. 2014. On the motivational properties of reward cues: Individual differences. Neuropharmacology. 76 Pt B:450–459.

Robinson TE, Flagel SB. 2009. Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. Biol Psychiatry. 65:869–873.

Rolls ET. 2000. Précis of the brain and emotion. Behav Brain Sci. 23:177–233.

Rushworth MFS, Walton ME, Kennerley SW, Bannerman DM. 2004. Action sets and decisions in the medial frontal cortex. Trends Cogn Sci. 8:410–417.

Rushworth MFS, Behrens TEJ. 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci. 11:389–397.

Sambrook TD, Goslin J. 2016. Principal Components Analysis of Reward Prediction Errors in a Reinforcement Learning Task. Neuroimage. 124:276–286.

San Martín R, Manes F, Hurtado E, Isla P, Ibañez A. 2010. Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. Neuroimage. 51:1194–1204.

Saunders BT, Robinson TE. 2013. Individual variation in resisting temptation: implications for addiction. Neurosci Biobehav Rev. 37:1955–1975.

Schiller D, Cain CK, Curley NG, Schwartz JS, Stern S a, Ledoux JE, Phelps E a. 2008. Evidence for recovery of fear following immediate extinction in rats and humans. Learn Mem. 15:394–402.

Schiller D, Levy I, Niv Y, Ledoux JE, Phelps EA. 2008. From fear to safety and back: reversal of fear in the human brain. J Neurosci. 28:11517–11525.

Schlagenhauf F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze H-JJ, Dolan R, Heinz A. 2014. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. Neuroimage. 89:171–180.

Schlagenhauf F, Sterzer P, Schmack K, Ballmaier M, Rapp M, Wrase J, Juckel G, Gallinat J, Heinz A. 2009. Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. Biol Psychiatry. 65:1032–1039.

Schmidt TT, Ostwald D, Blankenburg F. 2014. Imaging tactile imagery: changes in brain connectivity support perceptual grounding of mental images in primary sensory cortices. Neuroimage. 98:216–224.

Schobel SA, Lewandowski NM, Corcoran CM, Moore H, Brown T, Malaspina D, Small SA. 2009. Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. Arch Gen Psychiatry. 66:938–946.

Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK, Yuji K. 2009. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat Rev Neurosci. 10:885–892.

Schultz W. 2007. Multiple dopamine functions at different time courses. Annu Rev Neurosci. 30:259–288.

Schultz W, Dayan P, Montague PR. 1997. A Neural Substrate of Prediction and Reward. Science (80-). 275:1593–1599.

Schultz W, Dickinson A. 2000. Neuronal coding of prediction errors. Annu Rev Neuoscience. 23:473–500.

Seymour B, Daw N, Dayan P, Singer T, Dolan R. 2007. Differential encoding of losses and gains in the human striatum. J Neurosci. 27:4826–4831.

Smeets PM, Barnes-Holmes D. 2003. Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. J Econ Psychol. 24:603–618.

Spoormaker VI, Blechert J, Goya-maldonado R, Sämann PG, Wilhelm FH, Czisch M. 2012. Additional support for the existence of skin conductance responses at unconditioned stimulus omission. Neuroimage. 63:1404–1407.

Talmi D, Atkinson R, El-Deredy W. 2013. The feedback-related negativity signals salience prediction errors, not reward prediction errors. J Neurosci. 33:8264–8269.

Talmi D, Seymour B, Dayan P, Dolan RJ. 2008. Human pavlovian-instrumental transfer. J Neurosci. 28:360–368.

Tamminga CA, Stan AD, Wagner AD. 2010. The hippocampal formation in schizophrenia. Am J Psychiatry. 167:1178–1193.

Tomie A, Aguado AS, Pohorecky LA, Benjamin D. 2000. Individual differences in pavlovian autoshaping of lever pressing in rats predict stress-induced corticosterone release and mesolimbic levels of monoamines. Pharmacol Biochem Behav. 65:509–517.

Tomie A, Aguado AS, Pohorecky LA, Benjamin D. 1998. Ethanol induces impulsive-like responding in a delay-of-reward operant choice procedure: impulsivity predicts autoshaping. Psychopharmacology (Berl). 139:376–382.

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. 2010. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 25:2649–2653.

Trick L, Hogarth L, Duka T. 2011. Prediction and uncertainty in human Pavlovian to instrumental transfer. J Exp Psychol Learn Mem Cogn. 37:757–765.

Tricomi E, Balleine BW, O'Doherty JP. 2009. A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci. 29:2225–2232.

Ungless MA. 2004. Dopamine: the salient issue. Trends Neurosci. 27:702–706.

Ungless MA, Magill PJ, Bolam JP. 2004. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science (80-). 303:2040–2042.

Ursino M, Cuppini C, Magosso E. 2014. Neurocomputational approaches to modelling multisensory integration in the brain: A review. Neural Networks. 60:141–165.

Ursino M, Zavaglia M, Astolfi L, Babiloni F. 2007. Use of a neural mass model for the analysis of effective connectivity among cortical regions based on high resolution EEG recordings. Biol Cybern. 96:351–365.

Uslaner JM, Acerbo MJ, Jones SA, Robinson TE. 2006. The attribution of incentive salience to a stimulus that signals an intravenous injection of cocaine. Behav Brain Res. 169:320–324.

Viola FC, Thorne J, Edmonds B, Schneider T, Eichele T, Debener S. 2009. Semi-automatic identification of independent components representing EEG artifact. Clin Neurophysiol. 120:868–877.

Volkow ND, Wang G-J, Fowler JS, Telang F. 2008. Overlapping neuronal circuits in addiction and obesity: evidence of systems pathology. Philos Trans R Soc Lond B Biol Sci. 363:3191–3200.

Walsh MM, Anderson JR. 2012. Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. Neurosci Biobehav Rev. 36:1870–1884.

Waltz JA, Schweitzer JB, Ross TJ, Kurup PK, Salmeron BJ, Rose EJ, Gold JM, Stein EA. 2010. Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. Neuropsychopharmacology. 35:2427–2439.

Wang J, Chen J, Lei Y, Li P. 2014. P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. Neuroreport. 25:1044–1048.

Watson P, de Wit S, Hommel B, Wiers RW. 2012. Motivational Mechanisms and Outcome Expectancies Underlying the Approach Bias toward Addictive Substances. Front Psychol. 3:250–262.

Watson P, Wiers RW, Hommel B, De Wit S. 2014. Working for food you don't desire. Cues interfere with goal-directed food-seeking. Appetite. 79:139–148.

Weinberg A, Luhmann CC, Bress JN, Hajcak G. 2012. Better late than never? The effect of feedback delay on ERP indices of reward processing. Cogn Affect Behav Neurosci. 12:671–677.

Weiss AP, Schacter DL, Goff DC, Rauch SL, Alpert NM, Fischman AJ, Heckers S. 2003. Impaired hippocampal recruitment during normal modulation of memory performance in schizophrenia. Biol Psychiatry. 53:48–55.

Williams SM, Goldman-Rakic PS. 1993. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. Cereb Cortex. 3:199–222.

Wyvell CL, Berridge KC. 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement.

Wyvell CL, Berridge KC. 2001. Incentive sensitization by previous amphetamine exposure: increased cue-triggered "wanting" for sucrose reward. J Neurosci. 21:7831–7840.

Yeung N, Botvinick MM, Cohen JD. 2004. The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol Rev. 111:931–959.

Yin HH, Knowlton BJ. 2006. The role of the basal ganglia in habit formation. Nat Rev Neurosci. 7:464–476.

Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell'Olio M, Francey SM, Cosgrave EM, Killackey E, Stanford C, Godfrey K, Buckby J. 2005. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry. 39:964–971.

ACKNOWLEDGMENTS

There are many people I want to thank for their contribution to these studies and to my PhD formation.

Christopher Timmerman, Simone Battaglia, Giulia Petrillo, Irene Sciulli, Irene Silvestri, Daniele Migliorati, Sara Moroni, Francesca Casadei, Chiara Lancioni, Carina Tudor-Sfetea, Cindy Tu, Gwyneth Zai and Naresh Subramaniam for helping in different stages of the experiments presented in this thesis.

Prof. Mauro Ursino, Prof. Elisa Magosso and Veronica Forcelli for the collaboration with the Department of Engineering of the University of Bologna.

Paula for being a stimulating person and friend, for our infinite walks in which we shared ideas and views about life and science.

Francesca, Paolo, Riccardo, Neil, Brianna, Manu, Rob, Cate, Marco, Sara and all the people from the CNC of Cesena.

Elena, Khatereh and Antonio for sharing with me each single step of this PhD, in bad and good times.

Davide for supporting my work, my choices and everything I am and do (and for preparing task illustrations, as well ☺).

Martin Maier and Gonzalo Arrondo, for teaching me almost all I know about EEG and fMRI analysis and for being such nice and patient friends and co-supervisors.

Prof. Trevor Robbins and Dr. Graham Murray for giving me the chance to be a part of their invaluable research teams, for being supportive and truly inspirational supervisors.

Finally, a very special thank goes to Professor Giuseppe di Pellegrino for being a great scientist and supervisor, for teaching me to always look above and beyond evidence, and for believing in me.

Thank you all.

Fundings

These works were supported by grants from the University of Bologna awarded to Prof. Giuseppe di Pellegrino: Ricerca Fondamentale Orientata; Ministero Istruzione Universita` e Ricerca (PRIN 2010, protocol number: 2010XPMFW4_009).

These works were also supported by grants from the University of Cambridge awarded to Dr. Graham Murray and Prof. Trevor Robbins: MRC Clinician Scientist award (G0701911); Brain and Behaviour Research Foundation Young Investigator; Isaac Newton Trust; NIHR BRC; BCNI; Wellcome Trust; Addenbrooke's Hospital.