Spring 5-11-2019

AN ERP INVESTIGATION OF REWARD SIGNALS FOR DIFFERING CLASSES OF REINFORCING STIMULI

Darin R. Brown
University of New Mexico - Main Campus

Recommended Citation

Brown, Darin R.. "AN ERP INVESTIGATION OF REWARD SIGNALS FOR DIFFERING CLASSES OF REINFORCING STIMULI." (2019). https://digitalrepository.unm.edu/psy_etds/280

Darin R Brown

Candidate

Psychology

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

James F. Cavanagh, Ph.D., Chairperson

Kent A. Kiehl, Ph.D.

Jeremy Hogeveen, Ph.D.

Philip A. Gable, Ph.D.

AN ERP INVESTIGATION OF REWARD SIGNALS FOR DIFFERING CLASSES OF REINFORCING STIMULI

BY

DARIN R. BROWN

B.A., Psychology, California State University, Los Angeles, 2011

M.A., Psychology, California State University, Los Angeles, 2014

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Psychology

The University of New Mexico
Albuquerque, New Mexico

May, 2019

ACKNOWLEDGEMENTS

First, I would like to acknowledge Joey, Johnny, Tommy, and Dee-Dee, for putting a song in my heart and teaching me how to kick through walls and keep going. Next, the biggest thanks to my mentor Jim Cavanagh, for always keeping a door open and a glass ready for me. Your mentorship and friendship have been, and are, invaluable, and I will never be able to truly express my appreciation. Additionally, the decade-long support I've received from my Master's program mentor, Joel Ellwanger has been critical to all the good things I get to look forward to. I would also like to thank my dissertation committee for all the split-second mobilizations, eating all my treats, and of course, providing aid and guidance with a project that I've worked towards my entire adult life. Also, to all the people who wrote me letters of support, took time to meet with me about numerous philosophical agendas, or simply modeled what a good academic should be, I can't thank you enough. I would also like to acknowledge my amazing graduate cohort for all the love and support, my rock-star lab mates Jacki, John, and Trevor, and of course all my wonderful assistants who worked for me over the years. I could not have accomplished any of this without you all. You saved my life, all of you. I would also like to thank my amazing friends who kept their air mattresses and spare rooms open when I visited. The greatest of all thanks to Casey, Pablo, and Ashley, for loving me and supporting me even when I left you all out in Los Angeles. I would also like to acknowledge my greatest supporter, Kathryn. Thank you for keeping my head above water, for dragging me, kicking and screaming, across all the hot coals, and for supporting me when you did not have to. Any joy I experience from here on is thanks to you. Lastly, I would like to thank my family. Your support, and love, and what I can only assume as never yielding patience has kept me alive. You've sacrificed so much so I can achieve my dream. Please know I take it seriously and I will strive to live a life worthy of your time and sacrifices.

AN ERP INVESTIGATION OF REWARD SIGNALS FOR DIFFERING CLASSES OF REINFORCING STIMULI

by

Darin R. Brown

B.A., Psychology, California State University, Los Angeles, 2011
M.A., Psychology, California State University, Los Angeles, 2014
` Ph.D., Psychology, University of New Mexico, 2019

ABSTRACT

In order to successfully traverse an actively complex environment, an agent is required to learn from the consequences of their actions. For over a century, models of behavior have been developed demonstrating these consequence-based learning systems. More recently, underlying biological systems have been found to adhere to these constructs of learning. The electroencephalographic signal known as the Reward Positivity (RewP) is thought to reflect a dopamine-dependent cortical signal specific to reward receipt. Importantly, this signal has been shown to adhere to an axiomatic (rule-like) positive reward prediction error, whereby it is evoked following outcomes that are better than expected. These features of the RewP make it a candidate marker for clinical populations, such as major depressive disorder, substance use disorder, and Parkinson's disease. Although recent experimental endeavors have highlighted key characteristics of the generation and modulation of the RewP, a major understudied feature of the RewP in humans is the link between hedonic experiences and reward processes, and how these interact to modulate learning. This dissertation aims to probe this overlooked hedonic aspect of RewP generation through the use of emotionally evocative image rewards. The first aim addresses methodological issues relating to the use of complex, ecologically valid stimuli in EEG experimentation. The second aim investigated techniques for rectifying these methodological issues. Lastly, the third aim investigated the use of emotionally salient images as rewards in a reinforcement learning paradigm.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

Learning from the consequences of past events is essential for successful survival. Numerous neural imaging methodologies have highlighted specific neural phenomena that appear active during positive and negative outcomes that lend themselves to the optimization of behavior. Due to its near perfect temporal resolution, electroencephalogram (EEG) in particular is well suited to assess these early outcome phenomena and how these signals relate to learning and decision-making. Recently, the EEG signal known as the reward positivity (RewP) has gained a great deal of attention as a signal of early reward processing. However, a major understudied feature of the RewP in humans is the link between hedonic experiences and reward processes, and how these interact to modulate learning. A better understanding of these features can offer important insights into value-based decision-making, detailed models of learning, and clinical practices for affect related disorders such as substance use disorder, Parkinson's disease, and depression. In the current report, I explored how these complex reward types may influence the RewP, and how these signals operate in reinforcement learning paradigms.

Reinforcement Learning: Theories and Biological Systems

In order to successfully traverse an actively complex environment, an agent is required to learn from the consequences of their actions. The successful utilization of this consequence-based learning system would require an agent to maximize positive outcomes while simultaneously minimizing negative outcomes. For over a century, animal models have been developed demonstrating these consequence-based learning systems (Thorndike, 1898). Thorndike coined these early behavioral effects as the *Law of*

Effect, whereby, once the appropriate action-effect association has been made, behaviors that led to pleasant or rewarding outcomes are repeated, while negative outcomes or punishments are lessened or stopped. More recently, underlying biological systems have been found to adhere to these constructs of learning. Research investigating midbrain dopamine projections has revealed specific trends in its activation that directly relate to feedback learning (O'doherty et al., 2004; Schultz, Dayan, & Montague, 1997). In these models, dopaminergic projections from the ventral tegmental area (VTA) and substantia nigra transport information related to an outcome and its deviation from what was expected to happen. These signals have become known as signals of reward prediction error (RPE).

RPE signals have been shown to influence activation in neural sites such as midbrain regions (Zaghloul et al., 2009), striatum (Delgado, 2007; Mcclure et al., 2004; O'Doherty, 2004), and cortex (Gehring & Willoughby, 2002; Smith et al., 2009; Xue et al., 2008). These systems of reward processing appear to be especially sensitive to specific features of reward, such as its value, the likelihood that it was the expected outcome, the reward's valence (good or bad), how motivated the agent was to receive it, and the agency employed to receive the reward (Caplin & Dean, 2008; Holroyd & Coles, 2002; Schultz et al., 1997; Talmi, Fuentemilla, Litvak, Duzel, & Dolan, 2012).

Much of our understanding of these reward-related learning systems are derived from animal models (Schultz, 2015), where some specific behavior leads to the animal receiving food (Pavlov, 2010; Thorndike, 1898) or juice (Schultz et al., 1997) or some other hedonic outcome. These early animal models demonstrate how the mammalian brain learns to predict positive outcomes and to form associations between behaviors and these outcomes. However, although Pavlov's dog or Thorndyke's cat were able to successfully learn these associations, it is important to note the obvious: that animals cannot inform the experimenter on the degree to which the reward was enjoyed. Indeed, many models of reinforcement learning ignore (or simply cannot probe) these underlying emotion-related processes, because most emotional states can only be inferred, not confirmed in animals (Mendl, Burman, & Paul, 2010). Although what might be thought of as minutia in models of behavior, gaining an understanding of how these underlying emotional responses influence reward systems and signals could lead to greater understandings of clinical groups found to possess aberrant reward processing. For example, people with major depressive disorder show decreased neural signals in reward learning tasks, yet have no problem learning from them (Cavanagh, Bismark, Frank, & Allen, 2019), suggesting some underlying emotional-motivational modulation in these signals. An important research direction for gaining an understanding of these motivational state-trait aspects of reward would be to use highly emotional stimuli as rewards. One study in particular utilizing functional magnetic resonance imaging (Sabatinelli, Bradley, Lang, Costa, & Versace, 2007) found that neurological reward processing centers were activated for positively valenced pictures but not for aversive pictures. Indeed, other imaging methodologies would benefit from using complex emotional stimuli as rewards. EEG is an excellent tool for the study of these reward-emotion interactions due to its sensitivity to real-time canonical neural computations (Cavanagh et al., 2019). A component of EEG called the RewP appears to represent an EEG signal of positive RPE. Investigations utilizing emotional stimuli to elicit the RewP would further the understanding of these time-sensitive motivational processes.

The Reward Positivity

The generation of the RewP appears to adhere to an axiomatic (rule-like) function of +RPE, in that it is evoked when outcomes are better than expected (Baker & Holroyd, 2011; Holroyd & Umemoto, 2016). The RewP has been shown to reach maximal amplitudes over frontal-central sites usually 200 – 350ms following reward receipt (see Figure 1). The established occurrence of the RewP is in the exact time window where the N2 ERP component otherwise exists (Holroyd & Umemoto, 2016). Similar to other feedback-related signals (e.g. the feedback-related negativity or the error-related negativity), the RewP appears to reflect modulation of a control function from anterior cingulate cortex (ACC) signals (Baker & Holroyd, 2011; Holroyd & Umemoto, 2016). This specific feature of the RewP has been postulated to be the product of the inhibition of apical dendrites of motor neurons in the ACC by midbrain dopaminergic projections following an outcome that was better than expected. Although this theory of its generation has yet to be formally investigated, there is some evidence for this theory. First, similar to phasic dopamine firing, the RewP conforms to an axiomatic +RPE (Ait Oumeziane & Foti, 2016; Holroyd & Coles, 2002; Holroyd, Krigolson, & Lee, 2011; Proudfit, 2015). The timing of midbrain dopamine firing coincides with the time of the RewP activation (Luque, López, & Marco-pallares, 2012; Sambrook & Goslin, 2015; Talmi et al., 2012). Structural MRI studies have also shown that grey matter volume in the brain region of the dopamine nuclei is positively correlated with RewP amplitudes (Sambrook & Goslin, 2015; Zaghloul et al., 2009). Another piece of evidence for this theory is that individual differences in RewP amplitude are associated with individual differences in the expression of the D4 receptor (Carlson, Foti, Harmon-Jones, & Proudfit, 2015). Also, RewP amplitudes, but not the amplitudes for other potentials evoked during a virtual T-maze task were selectively blocked by a D2/D3 agonist (Baker, Stockwell, Barnes, Haesevoets, & Holroyd, 2016). Similarly, animal studies have shown that primates that receive a dopamine antagonist produce attenuated RewP-type waveforms (Santesso et al., 2009). Lastly, RewP amplitudes appear to be positively correlated with BOLD signals in the striatum (a signal believed to represent an axiomatic RPE; Vezoli, Procyk, Cell, & Lyon, 2009). Taken together, there appears to be quite a large amount of converging evidence endorsing this theory of RewP generation.

The RewP has been shown to be sensitive to experimental and individual factors, such as if the subject is motivated (Threadgill & Gable, 2016), angry (Angus, Kemkes, Schutter, & Harmon-Jones, 2015), or if the feedback is presented with a pleasant picture (Brown & Cavanagh, 2018). RewP amplitudes have been shown to be diminished in people with depression (Foti, Kotov, Klein, & Hajcak, 2011) or in people who adhere to authoritarian parenting styles (Nelson, Perlman, Klein, Kotov, & Hajcak, 2016; Proudfit, 2015). Surprisingly, RewP amplitudes appear to remain stable throughout development, suggesting that reward systems occur early in human development (Levinson, Speed, Nelson, Bress, & Hajcak, 2017). Importantly, the RewP appears to be sensitive to contextual factors that motivate behavior, suggesting that this signal is sensitive to global and not local reward context. In order to test this, Kujawa and colleagues (2013) had participants perform a simple gambling task where the value of the null condition differed between experimental blocks (win or lose). They discovered that evoked potentials for the null condition remained stable regardless of the block it was in. This suggests that the RewP is sensitive to all possible outcomes in a given policy state (Lukie, Montazer-Hojat, & Holroyd, 2014). Taken together, the evidence suggests that investigations of these reward signals using complex emotionally-salient rewards may offer important insights into learning paradigms and clinical populations (Cavanagh et al., 2019).

In addition, EEG investigations of reward signals have revealed a strong relationship between the RewP and underlying delta band activity (Bernat, Nelson, & Baskin-Sommers, 2016; Bernat, Nelson, Holroyd, Gehring, & Patrick, 2008). Indeed, many common ERP components reflect the summation of multiple frequency components (Bernat et al., 2008; Cohen, 2014). Two frequency bands in particular have received a great deal of attention regarding feedback-related neural electrical signals: theta band power (for losses) and delta band power (for wins). Importantly, delta band power has been shown to adhere to characteristics of positive RPE (Cavanagh, 2015). Furthermore, a recent study examining local field potentials sampled from the substantia nigra pars compacta in non-human primates revealed response patterns of dopamine neurons correlated with delta band frequency oscillations (Pasquereau, Tremblay, & Turner, 2019). These local field potentials were tuned only to reward-related processes and were unaffected by other events in the study such as movement. Indeed, a major benefit of spectral decompositions is its ability to identify separable overlapping neural activity.

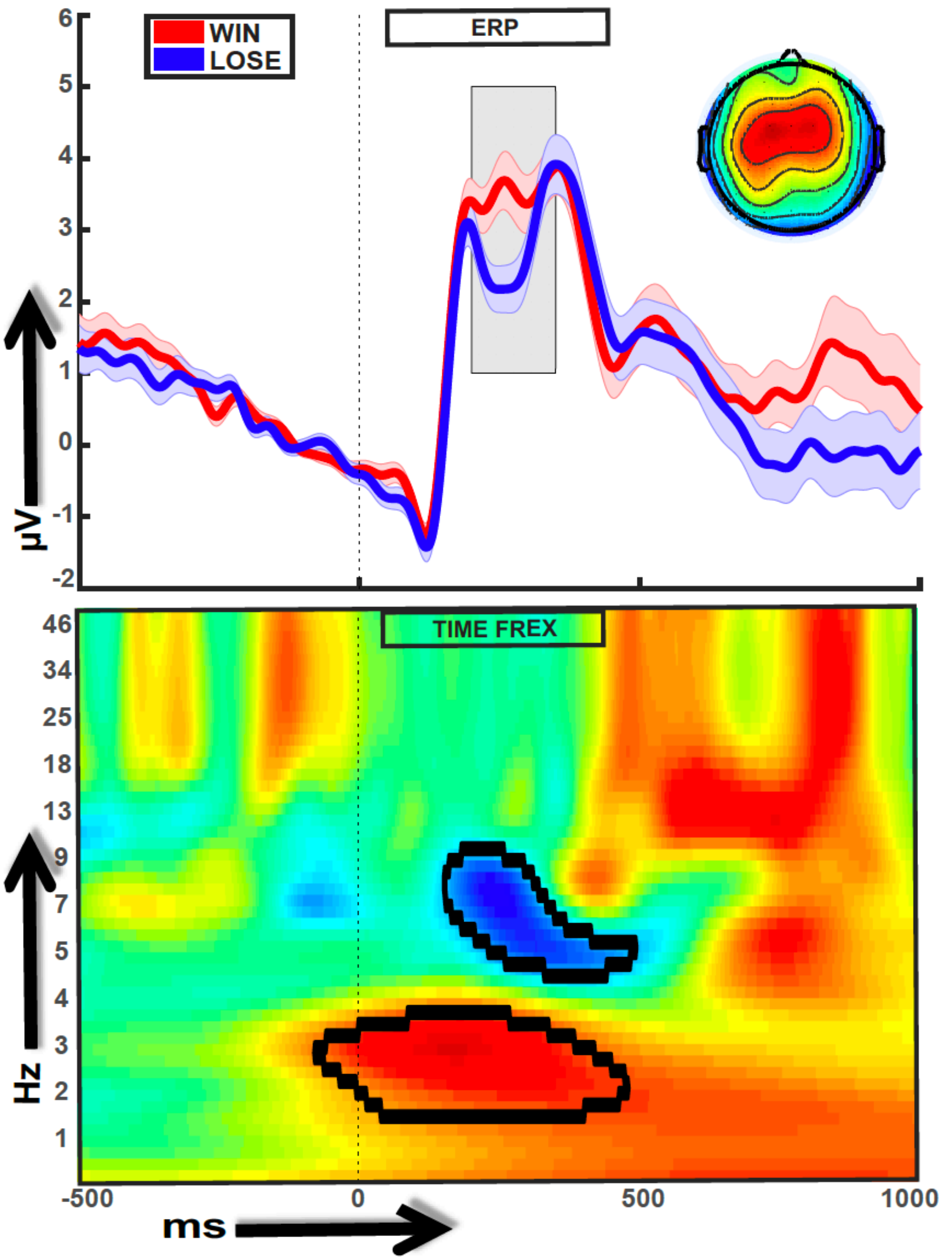

Figure 1. ERP and temporal decomposition plots following task feedback. The grey box (TOP ROW) highlights the general morphology of the RewP (in red) at electrode site Cz. The RewP is characterized by a “shelf-like” morphology connecting the P2 and P3 ERP components following better than expected feedback. Loss feedback (in blue) typically contains an enhanced N2 component. This feature is what characteristically disassociates these two feedback-related neural computational signals. Topographic maps for win minus lose feedback (scaled: ± 2.5 μV) highlight voltage distribution topographies for the RewP, which are maximal at central dorsal electrode sites. The time frequency plot (BOTTOM ROW) displays frequency power at electrode site Cz for win feedback minus lose feedback. Importantly, delta band power (~1 – 4 Hz) is more pronounced following feedback signifying a win (highlighted in red) while theta power (~5 – 7 Hz) is more pronounced following feedback related to a loss (highlighted in blue).

Clinical Applications

A recent review (Holroyd & Umemoto, 2016) has suggested that many aberrant behaviors and psychological disorders can be traced to maladaptive ACC activation. In this review, the authors provide detailed evidence for the role the ACC plays in disorders such as Parkinson’s disease, depression, anxiety, schizophrenia, and substance use disorder. Recent directions by the United States National Institute of Mental Health have led to approaching mental disorders from a Research Domain Criteria framework, aimed to “develop, for research purposes, new ways of classifying mental disorders based on behavioral dimensions and neurobiological measures” (Insel et al., 2010; National Institute of Mental Health, 2015). This approach is aimed to highlight basic components of symptomology that are present across disorders. Due to its relationship with dopaminergic tone, the RewP may be a strong candidate for such an approach.

Recent investigations of the RewP in clinical populations have already shown that this signal responds meaningfully across disorders. Schizophrenia, for example, has been associated with aberrant ACC function that may be due in part to deregulated dopamine expression (Seeman & Kapur, 2000). Past research has provided evidence that people suffering from schizophrenia have maladaptive associative learning rates when compared to healthy controls (Nestor et al., 2014), yet respond normally to rewards (Gold, Waltz, Prentice, Morris, & Heerey, 2008). Indeed, during a probabilistic learning task, RewP amplitudes remained stable between those with schizophrenia and normal healthy controls (Morris, Heerey, Gold, & Holroyd, 2008), suggesting that this clinical group has difficulty making associations between a reward’s value and motivated behavior. In contrast, people with major depressive disorder have been shown to have blunted RewP signals (Proudfit, 2015; Proudfit, Bress, Foti, Kujawa, & Klein, 2015) yet are able to learn from reward presentations (Cavanagh et al., 2019; Kumar et al., 2008), suggesting a maladaptive hedonic appreciation of the rewarding stimuli in this clinical group.

The RewP may also be useful as a diagnostic tool in diseased populations. Indeed, a recent study showed that early stage Parkinson's disease correlates with blunted RewP amplitudes; however, RewP amplitudes normalized during the progression of the disease (Brown, Pirio-Richardson, & Cavanagh, in review). This signal blunting during early stages of the disease might suggest an early indicator of disease onset. However, future longitudinal studies must be conducted to further unpack this effect. Another recent study revealed that blunted RewPs were associated with maternal major depression (a major risk factor for depression development: Kujawaa, Proudfit, Laptook, & Klein, 2016). Additionally, blunted RewP amplitudes have been suggested to be a major predictor of major depressive disorder (Proudfit, 2015). This feature of the RewP may have more to do with emotion-related hedonic aspects of reward receipt (Cavanagh et al., 2019) in that the blunted RewP seen in major depression is specifically related to anhedonia (Bress, Foti, Kotov, Klein, & Hajcak, 2013). Again, emotional processes appear to modulate this signal.

Recent studies have revealed that the RewP is sensitive to emotional manipulation (Brown & Cavanagh, 2018; Threadgill & Gable, 2016). These studies suggest that hedonic aspects of reward processing can influence the generation of this signal, motivating future studies to use emotional rewards for clinically depressed individuals or drug-specific rewards (for example pictures of alcohol) for people with substance use disorders (Verdejo-Garcia, Perez-Garcia, & Bechara, 2005). However, an important limitation in the RewP literature is its over-reliance on one specific class of reward. Much of what is known about RewP dynamics relates to a participant winning or losing conditional types of extrinsic rewards (i.e., money or points). Studies using animal models and fMRI, on the other hand, have made great strides in investing reward signals by utilizing both primary rewards (e.g., food and liquid rewards) and secondary rewards (e.g., points and money). One fMRI study in particular (Beck, Locke, Savine, Jimura, & Braver, 2010) revealed differing neural pattern activation between these two reward types, whereby transient BOLD activation effects were revealed for primary rewards but not for secondary rewards. The results of this study also revealed that reward processing centers were dissociable between reward types: secondary rewards were processed in cognitive control centers, such as the ACC and posterior cingulate cortex, and primary rewards were associated with sustained activation in sub-cortical regions such as the striatum and amygdala. EEG studies have historically employed only a single type of rewarding feedback (e.g., money or points) in order to investigate reward-related neural dynamics (Delgado, Locke, Stenger, & Fiez, 2003; Foti & Hajcak, 2009). Needless to say, this over-reliance on a single class of reward greatly limits the understanding of the boundary conditions and generalizability of these studies. The current series of studies aims to investigate the RewP and its underlying frequency components using complex, emotionally salient rewards.

The Present Studies

The present series of studies aimed to investigate the generation of the RewP for complex emotional rewards. A recent report revealed that the RewP evoked for emotional images is occluded by a signal of novelty (Brown & Cavanagh, 2018). The first two studies in the current series aimed to address and provide solutions to this methodological issue by utilizing spectral decomposition techniques. The third study in the series aimed to leverage RPE characteristics present in the RewP as an additional workaround. The final two studies aimed to investigate learning rates for differing types of rewards.

Chapter 2

NOVEL REWARDS OCCLUDE THE REWARD POSITIVITY, AND WHAT TO DO ABOUT IT

Abstract

The EEG response known as the Reward Positivity (RewP) appears to faithfully signal an axiomatic positive reward prediction error. This quality suggests that it reflects a core computational process in reinforcement valuation. Yet the RewP is also modulated by state and trait affect, suggesting that it has a more complex computational role than simple prediction error signaling. Here I conducted a series of experiments designed to disentangle the emotional aspects of reward processing, the nature of the interaction with control signals, and the culmination of these phenomena in the RewP. In the first two experiments I successfully occluded the RewP with an enhanced N2 component by presenting novel visual stimuli alongside rewards. This enhanced N2 and accompanying frontal midline theta power suggest that signals of control-evoking surprise occlude the RewP. In a third study, I parsed reward predictability, novelty, and the experience of affective valence using a fully informative outcome indicator prior to reward receipt. Participants were able to win pleasant novel images or points, but before the onset of the rewarding stimulus they were presented with a colored shape indicating whether or not they won, and what type of reward they would receive. The RewP successfully migrated to the first indication of reward (the colored shape), without the confound of novelty-induced N2 responses. Comparisons between win conditions revealed a larger RewP for cues that predicted pleasant images over points. Taken together, these findings suggest that the RewP reflects a combination of computational evaluation and affective valuation in the evaluation of reinforcement, and that novel stimuli can be used to modulate the RewP if appropriate methodological constraints are taken into account.

Experiment 1

A fundamental feature of reward is the intrinsic motivational value of hedonic pleasure (Berridge & Robinson, 1998; Schultz, 2000, 2015). Yet many aspects of reward are studied using probabilistic or deterministic reinforcement learning paradigms, which may be good for investigating anticipation or expectation but don't address other aspects of reward such as hedonic value or liking. Although prior research has identified neural signatures relating to varied aspects of reward processing in humans ((Berridge, 1996; Bunzeck, Doeller, Dolan, & Duzel, 2012; Finlayson, King, & Blundell, 2007; Kakade & Dayan, 2002; Smith & Berridge, 2005), electroencephalogram (EEG) experiments have suffered from confounding methodological issues due to novel, visually complex rewards. In our recent report, I detailed how novel pleasant images evoked an N2 instead of a Reward Positivity (RewP), yet with appropriate controls they appeared to enhance an underlying process suggestive of a RewP (Brown & Cavanagh, 2018). This methodological confound of novelty vs. reward responsivity hinders our understanding of how reward signals may be modulated by ecologically relevant stimuli. The use of idiosyncratic rewarding stimuli can assist in research on aberrant valuation in clinical populations, for example presenting alcohol (George et al., 2001) or cigarette cues (David et al., 2005) to people with substance use disorders, or emotionally evocative images to people who are depressed (Tremblay et al., 2005). These complex, ecologically relevant stimuli are necessary for a sophisticated understanding of motivational drives in learning and decision-making. In the current series of studies, I aimed to probe this methodological issue by formally investigating the interaction between reward and novelty in the EEG.

The RewP is a positive going EEG deflection appearing roughly 250ms following reward receipt over fronto-central sites. Recently, the RewP has received a great deal of attention as a neural signal of axiomatic (rule-based) positive reward prediction error (RPE), a signal representing a special case of surprise evoked when an outcome is better than expected (Brown & Cavanagh, 2018; Cavanagh, 2015; Cockburn & Holroyd, 2018; Heydari & Holroyd, 2016; Holroyd, Krigolson, & Lee, 2011; Meadows, Gable, Lohse, & Miller, 2016; Proudfit, 2015). Typically, the RewP has been evoked with point (Brown & Cavanagh, 2018), money (Bellebaum & Daum, 2008; Talmi, Atkinson, & El-Deredy, 2013), abstract icon (Angus, Kemkes, Schutter, & Harmon-Jones, 2015; Proudfit, 2015), or linguistic feedback (Cavanagh et al., 2019) indicating a successful decision. In fact, with very few exceptions (Angus et al., 2015; Heydari & Holroyd, 2016; Talmi, Atkinson, & El-Deredy, 2013), research involving the RewP has not explored its sensitivity to other forms of rewarding feedback. Unfortunately, the over-reliance on a single type of trial feedback greatly limits the understanding of the boundary conditions of the RewP.

The canonical occurrence of a RewP is in the exact time window where an N2 otherwise exists (Holroyd, Pakzad-Vaezi, & Krigolson, 2008). This specific feature of the RewP has been postulated to be the product of the inhibition of apical dendrites of motor neurons in the anterior cingulate cortex by midbrain dopaminergic projections following an outcome that was better than expected. Although the dopaminergic aspect of this theory has not been formally tested, numerous studies have shown that N2 amplitudes are evoked for novelty-induced control, including negative outcomes but also positive outcomes if they are rare or control-demanding (Baker & Holroyd, 2011; Oliveira, McDonald, & Goodman, 2007; Cavanagh, Figueroa, Cohen, & Frank, 2012).

This feature of the N2 explains why novel images evoked an enhanced N2 component for pleasant picture rewards in our previous study. The current study aimed to investigate if the induction of irrelevant novel stimuli during reward receipt will affect the morphology of the RewP by the N2 component.

There are numerous methodological techniques that may be useful for isolating overlapping spatio-temporal signals. The RewP and the N2 component don't only appear within the same temporal window, but also over the same fronto-central spatial locations on the scalp. However, these different components have different frequency representations in the delta (1-4 Hz) and theta (4-8 Hz) bands, respectively (Bernat et al., 2008). Spectral decomposition may be a fruitful technique for the isolation and examination of these overlapping frequency bands during novel rewards.

In the current report, I conducted a series of studies aimed to modulate the RewP through the pairing of rewarding stimuli with novel shapes and images. I hypothesized that the simple inclusion of irrelevant novel images would occlude the typical morphology of the RewP signal due to an enhanced N2. Then I tested two potential solutions to this methodological confound: first by leveraging spectral decomposition in order to parse condition-specific frequency bands that are sensitive to novelty vs. reward, and second, with a methodological manipulation of reward expectation.

Method

Participants

In all experiments, participants were recruited from the University of New Mexico subject pool. Students received class credits for participation. Participants were excluded from participation if they had a history of head injury that resulted in loss of consciousness for more than five minutes, had a history of epilepsy, had a history of any psychiatric or neurological disorder, or was currently on any psychiatric or neurological drugs. The Institutional Review Board of the University of New Mexico approved the study protocol. In Experiment 1 I recruited 30 participants (17 females) with a mean age of 20.47 (SD = 6.13).

Procedure

During the experiment participants played a simple forced choice task. The task was programmed in Matlab using Psychtoolbox (Brainard, 1997). On each trial, participants were presented three doors and were asked to guess which door they believe a reward was behind (Figure 2). Participants made a choice using an appropriate button press on the keyboard. A fixation cross was then displayed for a duration selected from a uniform distribution of 550 to 1050ms. Finally, participants received their win feedback (+1) or no-win feedback (a yellow bar). During half of the trials, the point feedback was presented in front of a novel shape. The participants were informed that the shape didn't mean anything, and that they should only focus their attention on the points. A new trial started automatically following a 1000ms inter-trial interval. Together, there was 80 trials total (20 point-only wins, 20 point-only no-wins, 20 point-shape wins, 20 point-shape no-wins). The average time for task completion was 10 minutes.

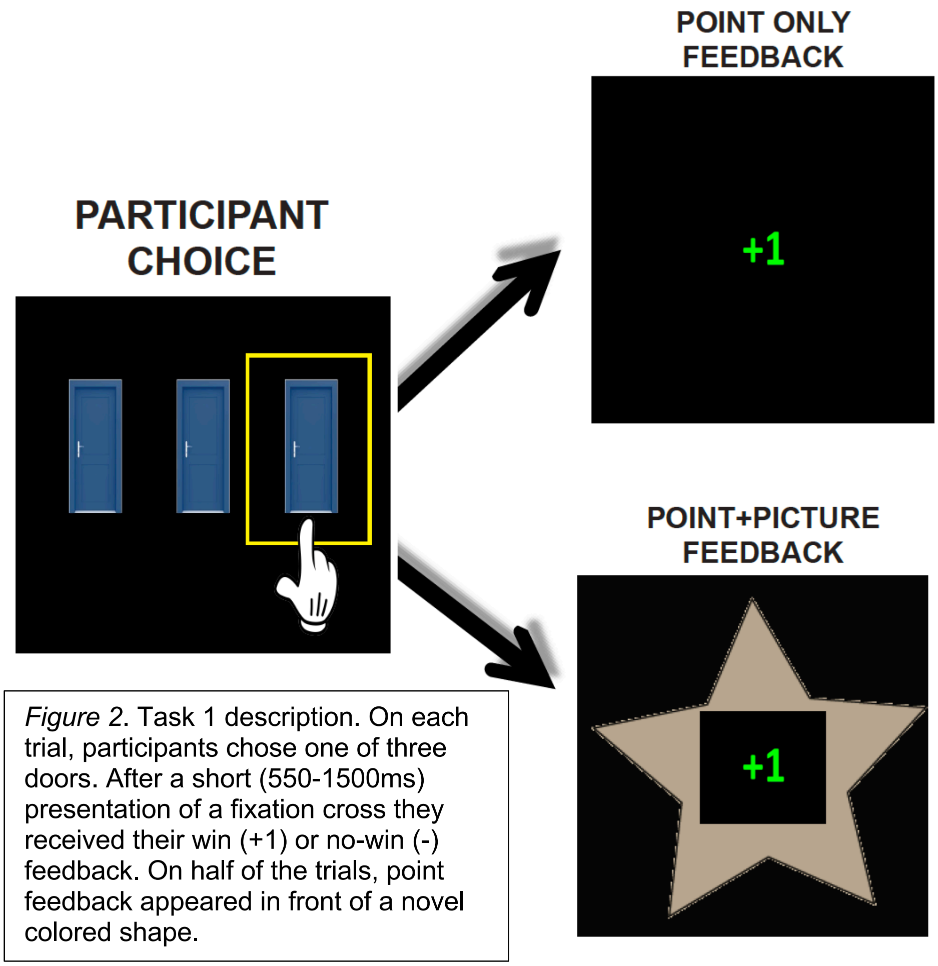

Figure 2. Task 1 description. On each trial, participants chose one of three doors. After a short (550-1500ms) presentation of a fixation cross they received their win (+1) or no-win (-) feedback. On half of the trials, point feedback appeared in front of a novel colored shape.

Data acquisition and preprocessing

Electrophysiological data were collected with a 64Ag–AgCl electrodes embedded in a stretch-lycra cap with a sampling rate of 500Hz with low and high cutoffs at .01-100 Hz. CPz served as the reference electrode and FPz as the ground electrode. Data was recorded with a Brain Vision system (Brain Products GmbH, Munich, Germany). Vertical electrooculogram (VEOG) activity generated by blinks was recorded by two auxiliary electrodes placed superior and inferior to the left pupil.

All EEG processing was conducted in EEGlab (Delorme & Makeig, 2004). First, CPz was re-created via computation of the average reference (EEGlab function pop_reref.m). Very ventral electrodes (FT9, FT10, TP9, and TP10) were then removed, as they tended to be unreliable. The average reference was then recomputed for the remaining 60 electrodes. ERPs were filtered between .01 to 20 Hz prior to averaging. In order to capture all trial events, data was epoched around feedback screen onset (-1000 to 3,000ms). Using statistical deviations from the mean for each EEG channel, FASTER (Nolan, Whelan, & Reilly, 2010) identified artifacts in each epoch for later rejection. Eye blink activities were removed following ICA (runica; Makeig, Bell, Jung, & Sejnowski, 1996). Epochs were then baseline corrected (-200 to 0 ms before feedback onset) and averaged to calculate event related potentials (ERP). The RewP was quantified at electrode site Cz and was measured between conditions within a 200 – 350ms window post feedback onset.

Time-frequency measures were computed using custom-written MATLAB functions (Cavanagh, Cohen, & Allen, 2009) by multiplying the fast Fourier transformed (FFT) power spectrum of single trial EEG data with the FFT power spectrum of a set of complex Morlet wavelets (defined as a Gaussian-windowed complex sine wave: $e^{i2\pi tf}e^{-t^{\wedge}2/(2x\sigma^{\wedge}2)}$, where t is time, f is frequency (which increase from 1–50 Hz in 50 logarithmically spaced steps) and defines the cycles of each frequency band, set according to $4/(2\pi f)$), and taking the inverse FFT. The end result of this process is identical to time-domain signal convolution, and it resulted in estimates of instantaneous power (the magnitude of the analytic signal), defined as Z[t] (power time series: $p(t)=real[z\ (t)]^2 + imag[z(t)]^2$). Each epoch was then cut in lengths (-500 to 1,500ms). Power was normalized by conversion to a decibel scale (10 x log10[power(t)/power(baseline)]), allowing a direct comparison of effects across frequency bands. To account for spectral smearing (Cohen, 2014) frequency power plots was measured between conditions within a 200 – 400ms window post feedback onset.

Statistical Analyses

Greenhouse-Geisser adjusted ANOVAs and planned comparison decompositions were used for data analyses. Reports of effect size for ANOVA are partial-η^2, while planned comparison effect sized are reported as Cohen's *d.*

Results

ERP Results

Figure 3 reveals the grand average waveforms for stimulus locked ERPs. The 2(VALENCE: win vs. no-win) × 2(NOVELTY: point-only vs. point-shape) ANOVA revealed a significant VALENCE (win > no-win) main effect ($F(1,29) = 8.195$, $p = .008$, $\eta^2 = .220$) but no NOVELTY main effect ($F(1,29) = 0.033$, $p = .856$, $\eta^2 = .001$). The ANOVA revealed a significant interaction ($F(1,29) = 6.985$, $p = .013$, $\eta^2 = .194$). Bonferoni protected multiple comparison *t*-tests revealed a significant win > no-win difference in point-only condition ($t(29) = 4.244$, $\mathrm{p} < .001$, $d = .723$) but not in the point-shape condition ($t(29) = .327$, $\mathrm{p} = .746$, $d = .056$).

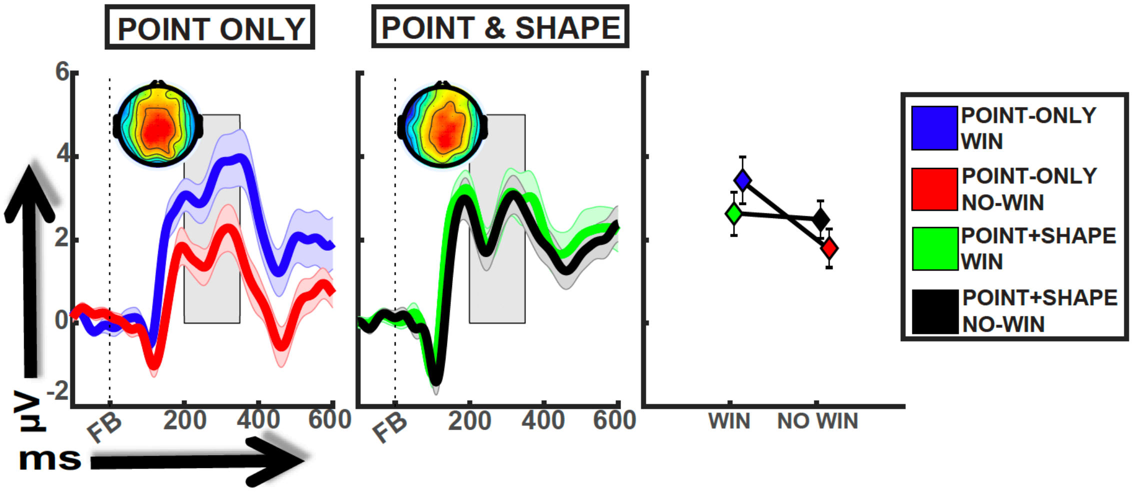

Figure 3. Feedback evoked ERPs for Experiment 1. These reveal statistical differences between win and no-win outcomes when participants were presented point-only feedback, but these were wiped out when feedback was paired with novel shapes. Topographic maps for win minus no-win feedback (scaled: ± 2.5 μV) reveal similar distributions of voltage differences during the RewP time window between the point-only and shape-point conditions. Line graphs represent averaged voltage over the RewP time window (200-350ms) and reveal a significant outcome by valence interaction.

I also employed a spatial filter (surface Laplacian) to investigate if this filtering technique could isolate these signals (see Figure 4). Generated ERPs after applying a Surface Laplacian spatial filter was not able to disseminate valence or novelty effects. Surface Laplacian is commonly conducted to emphasize electrical activity while filtering out noise related to spatial distribution of the signal (Carvalhaes & De Barros, 2015; Cohen, 2014). The 2(VALENCE: win vs. no-win) × 2(NOVELTY: point-only vs. point-shape) ANOVA failed to reveal a main effect for either VALENCE ($F(1,29) = .029$, $p = .866$, $\eta^2 = .001$), NOVELTY ($F(1,29) = 1.090$, $p = .305$, $\eta^2 = .036$), or an interaction ($F(1,29) = 2.174$, $p = .151$, $\eta^2 = .070$). Taken together, it appears that the surface Laplacian is ill-equipped for isolating the reward-specific component in the RewP.

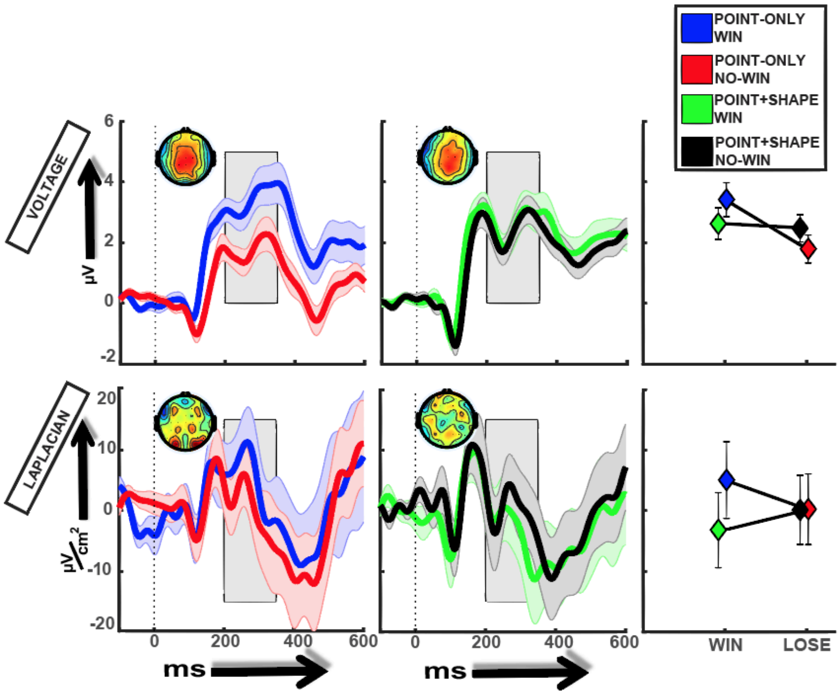

Figure 4. Experiment 1 ERP and topographic maps plots following voltage and surface Laplacian filtering. Voltage ERPs (TOP ROW) for trial feedback reveal a significant differentiation between win and no-win feedback in the point-only condition but not in the point-shape condition. Generated ERPs after applying a Surface Laplacian spatial filter (BOTTOM ROW) was not able to disseminate these effects.

Time Frequency Analysis Results

Figure 5 depicts the grand averages for feedback-locked time frequency decompositions. A 2(VALENCE) x 2(NOVELTY) repeated measures ANOVA was conducted on delta power (1–4 Hz) over the time frequency time window of 200 – 400ms. This analysis revealed a main effect for VALENCE (win > no-win: $F(1,29) = 7.446$, $p = .011$, $\eta^2 = .204$). The ANOVA failed to reveal a significant NOVELTY main effect (point-shape > point-only; $F(1,29) = 3.478$, $p = .072$, $\eta^2 = .107$) or an interaction ($F(1,29) = .218$, $p = .644$, $\eta^2 = .007$).

The 2(VALENCE) x 2(NOVELTY) repeated measures ANOVA on theta power (4–7 Hz) revealed a significant NOVELTY main effect (point-shape > point-only;

$F(1,29) = 10.118$, $p = .003$, $\eta^2 = .259$) but no main effect for VALENCE ($F(1,29) = 1.518$, $p = .228$, $\eta^2 = .050$), or interaction ($F(1,29) = .144$, $p = .707$, $\eta^2 = .005$).

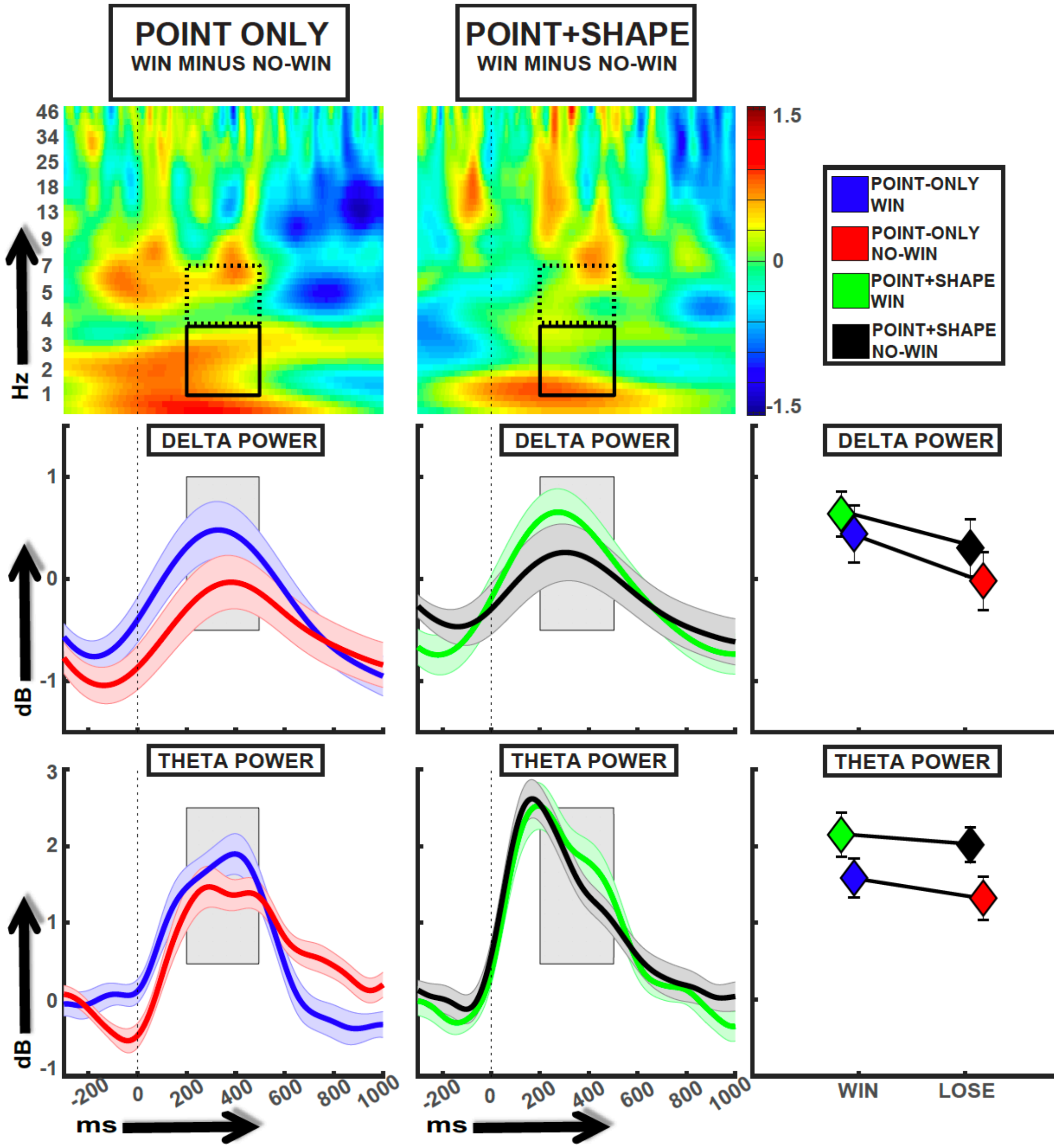

Figure 5. Experiment 1 Time-Frequency power plots. (TOP ROW) Time-frequency power for win minus no-win conditions (panel 1 and 2). The black boxes outline the time and frequency range used for analysis (solid box = delta; dotted box = theta). Increased delta band power (1–4 Hz) is evident for win conditions over that of no-win condition. (MIDDLE ROW) Delta band power envelopes and line plots display main effects for valence and novelty. (BOTTOM ROW) Theta band power envelopes and line plots display only a main

Experiment 1 Discussion

These results revealed that the inclusion of novel stimuli interfered with the detection or generation of the RewP, but reward-locked delta power was unaffected. Furthermore, there was increased theta power, an index of general surprise (Cavanagh & Frank, 2014), for win feedback paired with novel shapes, demonstrating a double dissociation between the two frequency ranges. Taken together, these results support our hypothesis that signals of novelty occlude the detection of EEG signals of the RewP and that spectral decomposition is a viable technique for isolating these separable neural signatures.

Experiment 2

In Experiment 1 it was revealed that novel shapes evoked an enhanced N2 component which occluded the RewP, and frequency decompositions were able to isolate these distinct processes. I suspect this effect was produced by the outcome-specific surprise evoked for the inclusion of novel shapes with rewarding feedback. However, in that task there were two levels of surprise: (1) the trial-specific novelty surprise evoked for the novel shapes, and (2) the trial-to-trial outcome surprise of points-only vs. point-shape presentations. This additional level of surprise may have further enhanced this signal of novelty. To replicate this finding as well as extend it by isolating the effect to a single type of surprise, I designed a second experiment eliminating the trial-to-trial outcome surprise while still maintaining the trial-specific novelty by presenting the type of feedback (point-only and point-shape) in separate blocks.

Method

In Experiment 2 30 participants (20 female) were recruited with a mean age of 20.67 (SD = 3.77). Experiment 2 was the same as Experiment 1, but outcome conditions (point-only and point-shape) were blocked together so that participants would only be presented with a block of point-only feedback, and then after a short break, they would receive a block of feedback in the point-shape condition. The order of the outcome blocks was counter balanced in order to avoid primacy effects. All experimental protocols and data processing procedures were similar between Experiment 2 and Experiment 1.

Results

ERP Results

Figure 6 shows the grand average waveforms for stimulus locked ERPs. First, I examined if the ordering of the experimental blocks affected RewP amplitudes. There was no block order main effect for participants who received the point only condition first (n=19) versus those who received the point-shape condition first (n=11; $F(1,28) = .246$, $p = .624$, $\eta^2 = .009$) suggesting that block order effects did not affect the amplitude of the RewP. The 2(VALENCE: win vs. no-win) × 2(NOVELTY: point-only vs. point-shape) revealed a significant VALENCE (win > no-win) main effect ($F(1,29) = 22.598$, $p < .001$, $\eta^2 = .438$) as well as a NOVELTY (point-only > point-shape) main effect ($F(1,29) = 4.782$, $p = .037$, $\eta^2 = .142$), but not an interaction ($F(1,29) = .467$, $p = .500$, $\eta^2 = .016$) suggesting that the RewP was sensitive to trial-specific novelty. However, a post-hoc analysis comparing RewP amplitudes between the two win conditions (point-only and point-shape) revealed a significant difference between the RewP evoked in these conditions ($t(29) = 2.230$, $p = .034$, $d = .423$), whereby RewPs evoked in the point-only condition were significantly larger than those evoked in the point-shape condition. Visual inspection of these signals reveals that the RewP for point-shape pairs had an enhanced N2 component. This suggests that the removal of the trial-to-trial outcome surprise in this task was not enough to completely remove the N2 component from overlapping the RewP.

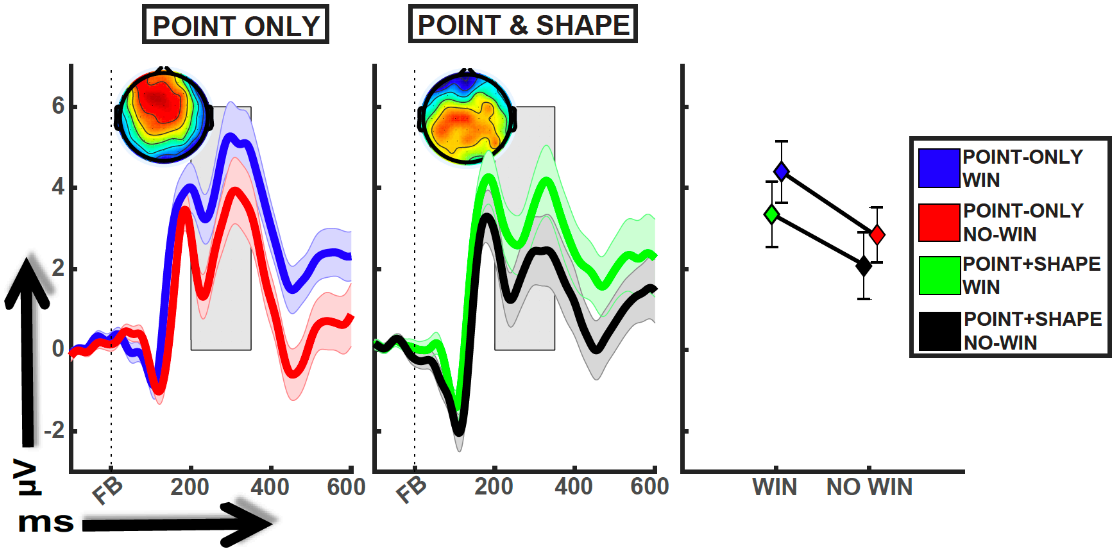

Figure 6. Feedback evoked ERPs when outcome types (point-only and point-shape) were blocked. Here we modified the degree of feedback novelty by keeping the outcome type consistent within blocks. This blocked designed reveal win vs. no-win differentiations were present for both point-only and point-shape outcomes. Topographic maps for win minus no-win feedback (scaled: ± 2.5 μV) reveal voltage distribution differences between the two outcome types, whereby win minus no-win voltage distribution was more anterior in the point-only condition while voltage distribution was more posterior in the point-shape conditions. This difference in voltage distributions is most likely a product of the visual gain involved in the processing of the colored shapes in the novel conditions.

Time Frequency Analysis Results

The ANOVA (Figure 7) results conducted on delta power replicated the effects from Experiment 1, whereby there was a main effect for VALENCE (win > no-win: $F(1,29) = 5.384$, $p = .028$, $\eta^2 = .157$ but no main effect for NOVELTY (point-shape > point-only; $F(1,29) = .568$, $p = .457$, $\eta^2 = .019$) or an interaction ($F(1,29) = .218$, $p = .644$, $\eta^2 = .007$). Surprisingly, the ANOVA on theta power did not reveal a significant NOVELTY main effect ($F(1,29) = 2.721$, $p = .110$, $\eta^2 = .086$) or VALENCE main effect

($F(1,29) = .004$, $p = .952$, $\eta^2 < .001$), nor was there an interaction ($F(1,29) = .056$, $p = .815$, $\eta^2 = .002$).

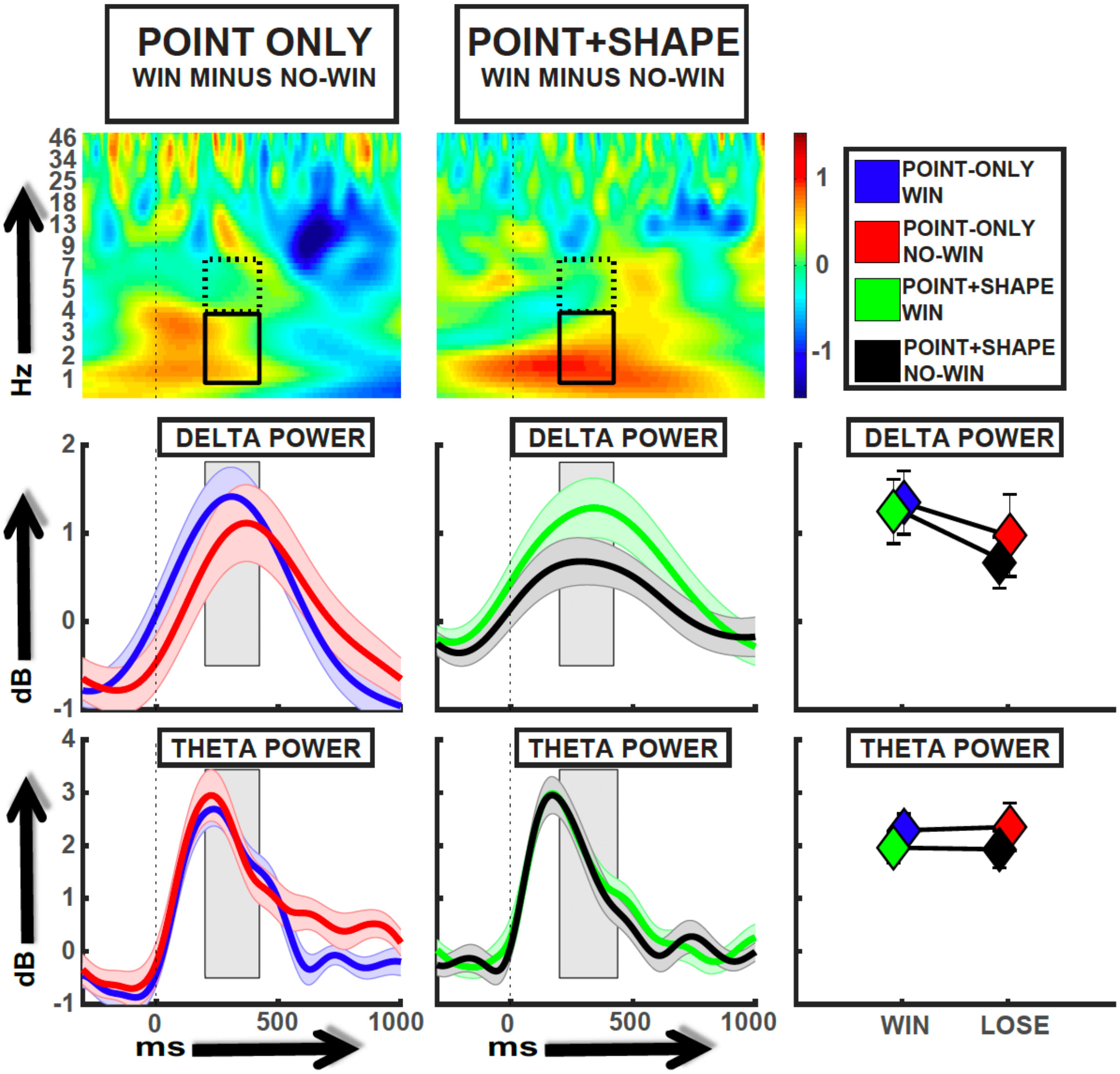

Figure 7. Experiment 2 Time-frequency power plots. (TOP ROW) Time-frequency power for win minus no-win conditions (panel 1 and 2). The black boxes outline the time and frequency range used for analysis (solid box = delta; dotted box = theta). Increased delta band power (1–4 Hz) is evident for win conditions over that of no-win condition. (MIDDLE ROW) Delta band power envelopes and line plots display main effects for valence and novelty. (BOTTOM ROW) Theta band power envelopes and line plots display no main effects or interactions.

Experiment 2 Discussion

In Experiment 2 the type of feedback received by the participants was differed by blocks in order to diminish trial-to-trial outcome surprise. The results reveal that the signal generated for novel shapes did not completely occlude the RewP after removing the trial-to-trial surprise and thus the signal for novel win feedback could be differentiated from the signal for novel no-win. However, in Experiment 2, I did see an enhanced N2 component in the point+shape win condition that was different from the point-only win condition. Spectral decomposition replicated Experiment 1 whereby delta power was larger for wins than no-wins. However, the increased theta power for novel condition disassociation I saw in Experiment 1 did not replicate in Experiment 2, likely due to the reduction of trial-to-trial outcome surprise. Taken together, it appears that the manipulation of surprise can lessen, but not eliminate this overlap between neural signals.

Experiment 3

In Experiment 1 I successfully occluded the RewP with the inclusion of novel shapes. I partially ameliorated this effect in Experiment 2 by presenting the feedback (standard and novel) in experimental blocks, thus limiting the trial-by-trial surprise in the viewing of novel rewarding stimuli. Even though the RewP evoked for novel rewards was not completely occluded by the novel shape pairings, I did see an enhanced N2 component in the novel reward conditions that was different from the RewP evoked for point-only rewards. These results suggest an inherent methodological limitation relating to the use of complex stimuli as reinforcing stimuli. Our earlier report suggested that intrinsically motivating rewards can still modulate the underlying RewP (Brown & Cavanagh, 2018), but I did not identify a way to isolate this variance to reward-related processing. However, by leveraging principles of reinforcement learning there may be options for the use of these complex rewards in experiments aimed to investigate the RewP.

A foundational feature of reinforcement learning is the migration of the RPE to the first reliable predictor of reinforcement (Sutton & Barto, 2018). A neural instantiation of this phenomenon is illustrated by the seminal work by Schultz and colleagues, where midbrain dopamine firing propagates back in time to the presentation of a cue predicting reward receipt without a response to the actual reinforcement itself (Schultz, Dayan, & Montague, 1997). The migration of signals to cues predicting reward receipt has also been documented in ERP experiments (Angus et al., 2015; Holroyd et al., 2011; Potts, Martin, Burton, & Montague, 2006), but not with hedonically meaningful stimuli. In this third experiment, I aimed to investigate if this signal of positive RPE would migrate to a predictive cue for different types of rewarding stimuli (points vs. pleasant pictures). I hypothesized that the RewP would be evoked for cues predicting point or picture rewards, and that it would be larger for preferred picture rewards. Additionally, I hypothesized that frontal midline delta would also migrate to the reward predicting cues while frontal midline theta would migrate to cues predicting no-win outcomes.

Method

Participants were 30 students (17 female) recruited from the University of New Mexico subject pool, with a mean age of 21.02 (SD = 3.52). Inclusion and exclusion criteria were the same as Experiments 1 and 2. First, in order to choose idiosyncratic picture rewards for the main task, participants completed a short image-rating task. On each trial participants were presented with an image and asked to rate how pleasant they found it from 1(not pleasant at all) to 9(extremely pleasant). Pictures were drawn from 1 of 5 affective image categories: male models, puppies, nature scenes, nude women, and babies (Brown & Cavanagh, 2018). The picture class that was rated as most pleasant was used as picture rewards in the following task. Picture classes were based on images that were rated as highly pleasurable from the International Affective Picture System (IAPS: Lang, Bradley, & Cuthbert, 1997). Standardized ratings of valence (1=negative to 9=positive) from the IAPS technical manual were compared by gender. The reward classes used for the current study were chosen from the most occurring image themes. All images were selected from internet searches. All images were inspected in task-presentation conditions in order to assess no image appeared blurry and that sizing was equal for all images. In order to assess if a participant's top picture selection was truly pleasant, an additional 6th category was made up of affectively neutral images (e.g. lightbulb, door knob, etc.) from the IAPS. All participants rated their top choice as more pleasant than the neutral picture class.

Following the image rating task, participants performed forced choice task similar to that of Experiment 1 and 2 (Figure 8). This iteration of the task differed in two ways: (1) participants would either receive point win (+1), picture win (1st choice from the image rating task) or no-win (-) feedback, and (2) before the presentation of the trial feedback, participants would receive a colored cue (either a star or a square) indicating what type of reward (point of picture) and whether they won or lost (green cue = win, red cue = no-win). Colored cues were presented to the participant for 1000ms followed by a second fixation screen. The second fixation screen remained on the screen for 550 to 1050ms and afterwards participants received their trial feedback. The task took an average of 22 minutes to complete.

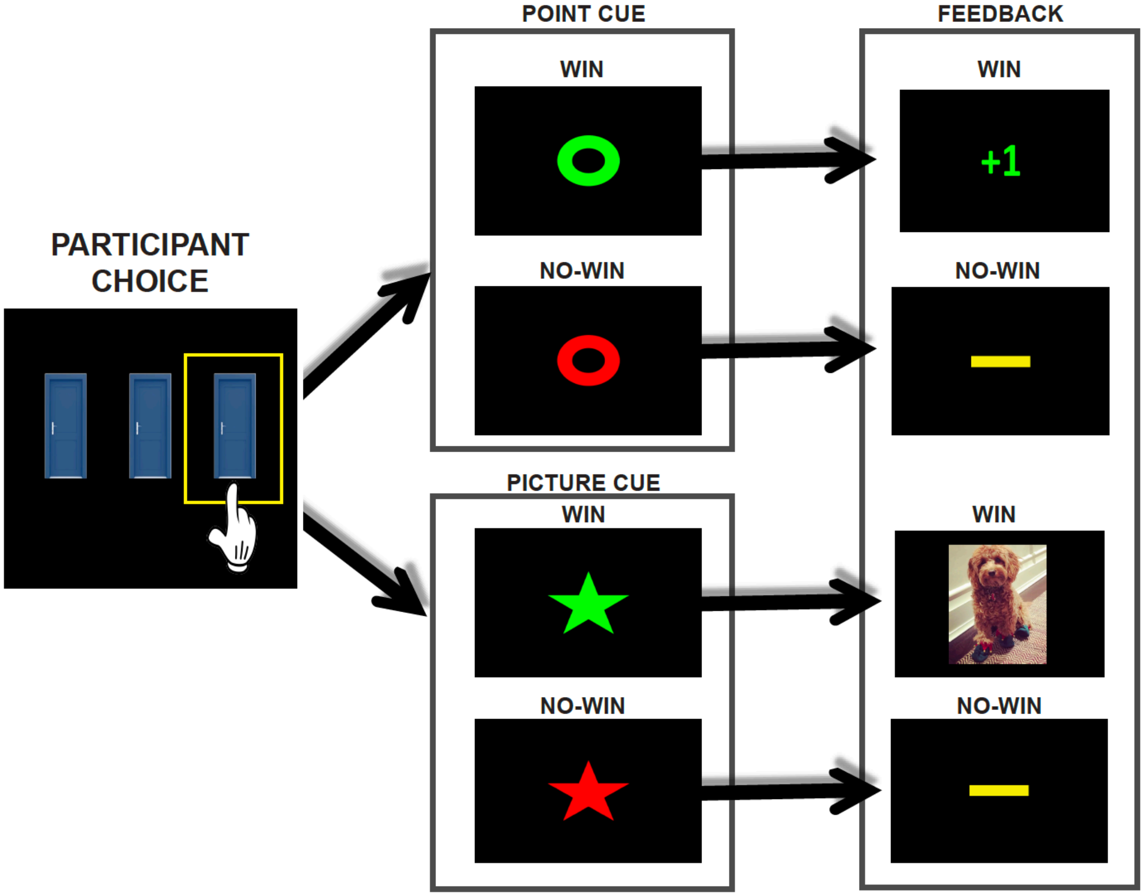

Figure 8. Experiment 3 task. After choosing a door, participants were presented with a one of two-colored cues. The shape of the cue indicated the type of reward the participant earned while the color indicated the whether they would receive win or no-win feedback. After a short delay, participants received the feedback.

After the forced choice task, participants completed a third task in order to assess feedback preferences between picture and point rewards. This image preference task was similar to our prior study (Brown & Cavanagh, 2018) where two pictures (one from each picture class) were presented on either side of the screen. Participants were instructed to choose the object that “would have been the better feedback” during the door task. Participants were also presented point feedback along with the picture feedback. The point feedback presented was the same point outcomes that was presented during the forced choice task (+1 and a yellow bar). Additionally, participants were also presented a large point feedback (+5) that not presented to them during the forced choice task to act as a manipulation check. In the image preference task, all image classes and point feedback types were compared with each other 4 times each for a total of 112 trials and took about 7 minutes to complete.

Results

Picture Rating Task and Image Preference Task Results

Figure 9a depicts the bar plots for mean pleasant ratings for the image rating task. I conducted a one-way ANOVA comparing the pleasant ratings across the six levels of image class (1st Choice, 2nd Choice, 3rd Choice, 4th Choice, 5th Choice, and neutral). The ANOVA revealed a significant effect across the image classes ($F(2.850, 82.656) = 104.516$, $p < .001$, $\eta^2 = .783$). Multiple comparisons revealed pleasant ratings across all image classes were significantly different from each other (all $p < .001$) except the ratings made for the participant's 5th choice and neutral image classes ($p = .081$). These findings verify that our participants consistently found their top choice more pleasant than emotionally neutral images.

Next, I conducted an analysis comparing the feedback preferences from the choice task. Here, I conducted a 3(POINT: +5, +1, Null) x 5(CLASS CHOICE: 1st Choice, 2nd Choice, 3rd Choice, 4th Choice, 5th Choice) ANOVA. The ANOVA revealed a significant POINT main effect ($F(1.496, 43.395) = 60.670$, $p < .001$, $\eta^2 = .677$), whereby +5 was chosen more than +1 ($p < .001$) and null ($p < .001$), and +1 was chosen more than null ($p < .001$). There was also a CLASS CHOICE main effect ($F(3.265, 94.680) = 58.972$, $p < .001$, $\eta^2 = .670$), whereby a participant's picture choices all significantly differed from their top choice (all $p < .005$) except for the participants 2nd choice ($p = .068$). Importantly, there was a significant interaction ($F(6.510, 188.798) = 6.936$, $p <$

.001, $\eta^2 = .193$). Figure 9b depicts these preference trends. Importantly, >80% of participants preferences were for the 1st choice picture class over +1 point (which was used in the main task). This suggests that participants preferred receiving the images in their first choice set more than receiving the +1 feedback.

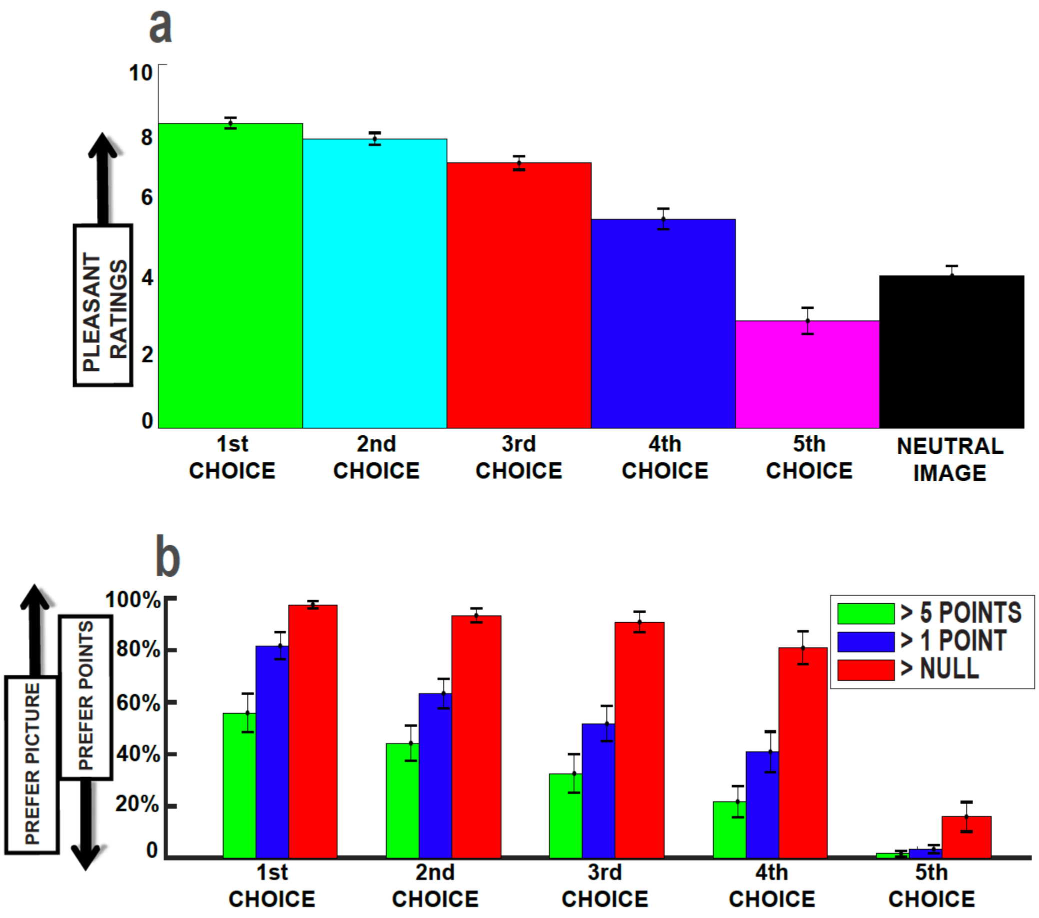

Figure 9. Results from the Picture Rating Task and Image Preference Task. Picture reward class was chosen based on the pleasantness scores rated by the participants. Top panel (a) displays the mean ratings of pleasantness across participants ranked image classes. The image class that was rated the most pleasant was later used in the forced choice task. Importantly, all images classes (except for the fifth-choice class) was rated as significantly more pleasant than the emotionally neutral class of images. (b) Displays results for the Image Preference Task. The majority (>80%) of participants preferred receiving their first-choice picture than the point feedback (+1) in the forced choice task.

ERPs Results for Predictive Cues

Figure 10 shows the grand average waveforms for ERP activities locked to the predictive cues. The 2(VALENCE: win vs. no-win) × 2(REWARD TYPE: point vs. picture) ANOVA revealed a significant VALENCE (win > no-win) main effect (*F*(1,29) = 6.752, $p = .015$, $\eta^2 = .189$) as well as an REWARD TYPE (picture > point) main effect (*F*(1,29) = 6.010, $p = .020$, $\eta^2 = .172$) without a significant interaction (*F*(1,29) = 1.621, $p = .213$, $\eta^2 = .053$). Taken together these results support our hypothesis in that the RewP successfully migrated to the first instance of trial feedback and correctly differentially win from no-win feedback for both points and picture.

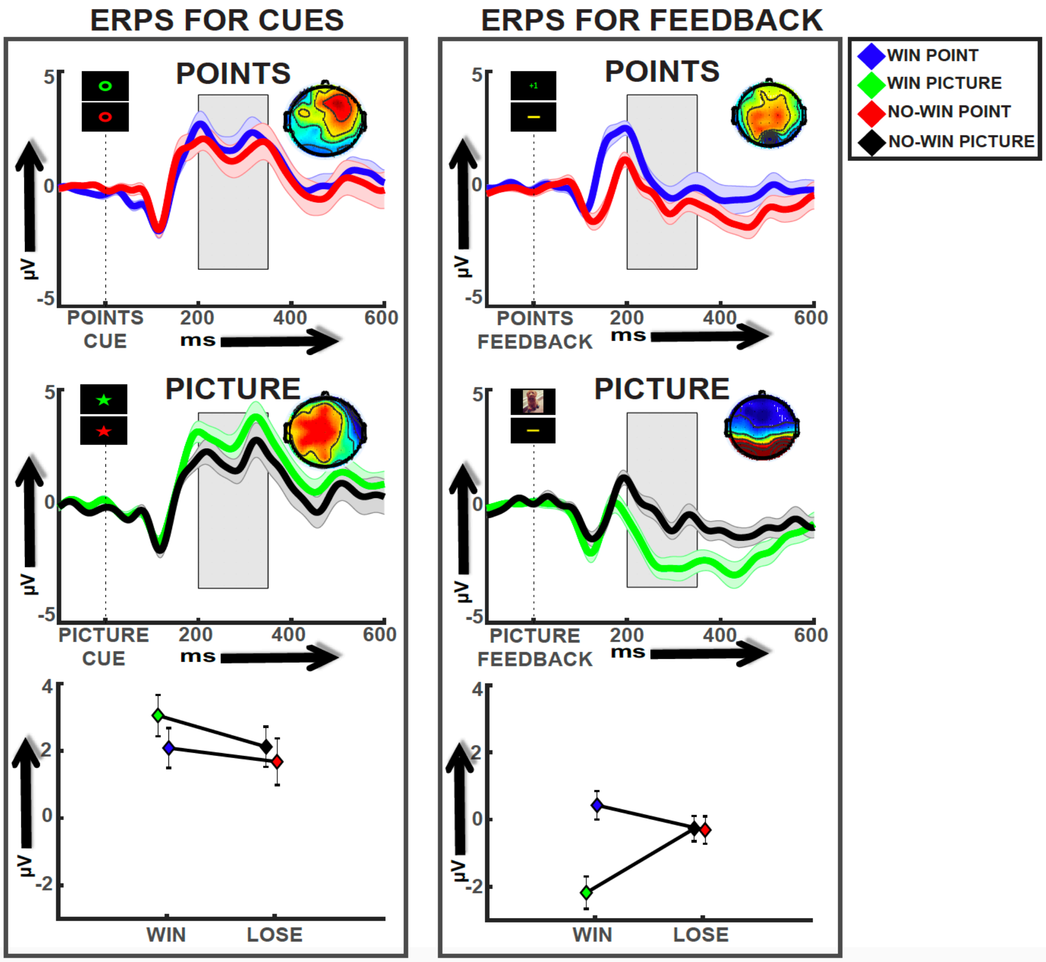

Figure 10. ERPs for reward cues (left) and trial feedback (right). In adherence with our hypothesis, the RewP migrated to the first instance of reward receipt, the predictive cue. This effect was true for both point and picture feedback. Interestingly, once striped of the expectation of reward, the ERP's for point reward exhibited an enhance P2 but no RewP, while the picture reward exhibited an enhanced negativity. Topographic maps for win minus no-win cue and feedback (scaled: ± 2.5 μV) reveal similar distributions of voltage differences during the RewP time window between the cues and point feedback. The voltage distribution for the picture feedback were greatest over occipital sites most likely due to the visual complexity of the images; with an average reference this posterior effect caused the large dip (green) at Cz in the time window here.

Experiment 3 ERP Results for Trial Feedback

The ANOVA examining ERP amplitudes during the RewP time window (200-350ms) for trial feedback failed to reveal a main effect for VALENCE ($F(1,29) = 2.980$, $p = .095$, $\eta^2 = .093$), however there was an REWARD TYPE main effect(point > picture: $F(1,29) = 20.286$, $p < .001$, $\eta^2 = .412$) as well as an interaction ($F(1,29) = 18.426$, $p < .001$, $\eta^2 = .389$). Bonferoni protected multiple comparisons revealed significant ERP difference between win feedback (point > picture: ($t(29) = 5.012$, p < .001, $d = .973$) but not for no-win feedback ($t(29) = .148$, p = .883, $d = .028$). Visual inspection of the ERPs evoked for trial feedback (Figure 10) reveal an enhanced voltage negativity for picture rewards, mirroring the ERPs from a previous study involving picture feedback (Brown & Cavanagh, 2018). As hypothesized, ERPs evoked by the feedback did not have a RewP.

Time Frequency Analysis Results for Predictive Cues

Figure 11 depicts the grand averages for time frequency decompositions locked to the predictive cue. A 2(VALENCE) x 2(REWARD TYPE) repeated measures ANOVA was conducted on frontal midline delta power and revealed a significant main effect for VALENCE (win > no-win; $F(1,29) = 9.140$, $p = .005$, $\eta^2 = .240$), but no REWARD TYPE main effect ($F(1,29) = 1.241$, $p = .274$, $\eta^2 = .041$) or interaction ($F(1,29) = .001$, $p = .980$, $\eta^2 < .001$). The 2(VALENCE) x 2(REWARD TYPE) repeated measures ANOVA conducted on frontal midline theta power failed to reveal either a VALENCE or REWARD TYPE main effect ($F(1,29) = .578$, $p = .453$, $\eta^2 = .020$; $F(1,29) = 1.972$, $p = .171$, $\eta^2 = .064$, respectively), or an interaction ($F(1,29) = .771$, $p = .387$, $\eta^2 = .026$).

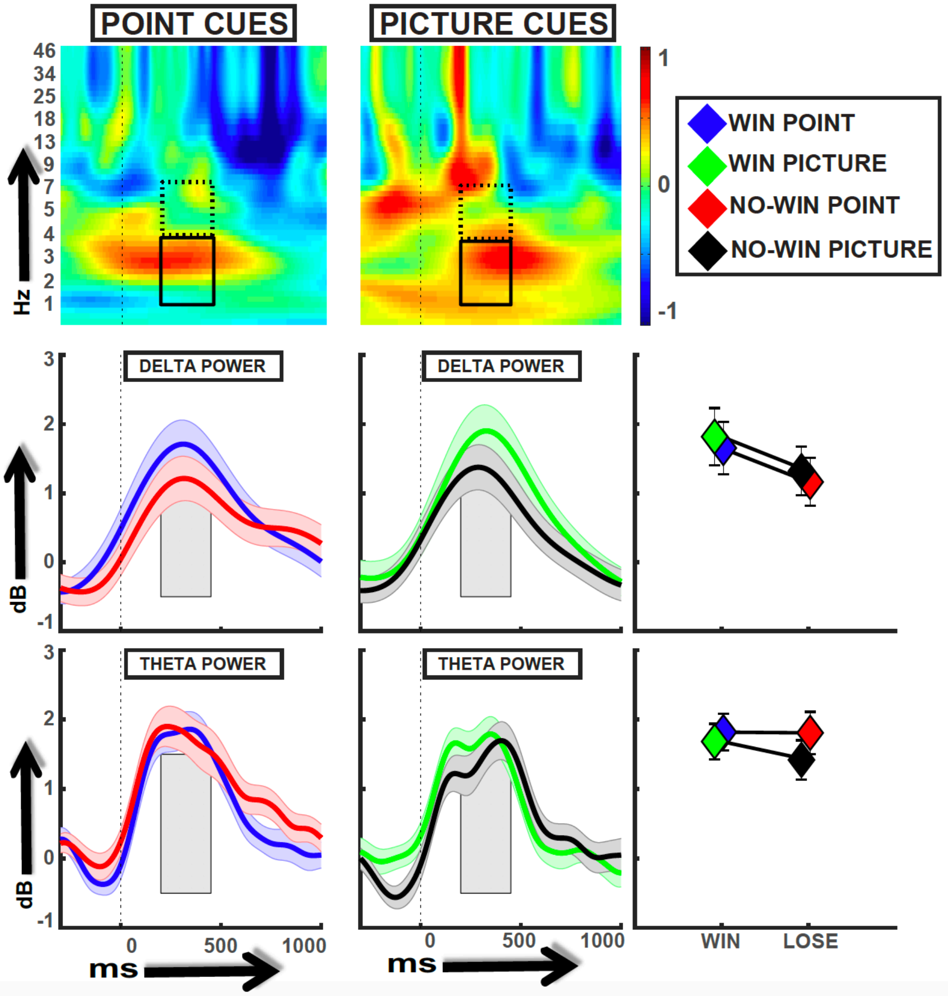

Figure 11. Time-frequency power plots for reward cues. Time-frequency power for win minus no-win conditions Increased delta band power, outlined by the black boxes, is evident for win conditions over that of no-win condition, yet these did not differ between cues predicting different reward types in delta power plots (MIDDLE ROW). There were no differences in the theta power for cues predicting rewards (BOTTOM ROW).

General Discussion

In the current study I investigated if the RewP would be occluded by signals of novelty. In the first experiment, trial feedback (+1 and -) was presented by itself or in front of a novel shape. These reward-shape pairings evoked an enhanced N2 component thus occluding the RewP. In a follow up experiment, I aimed to remove the trial-to-trial outcome surprise between the conditions of interest (point vs. point+shape), while still maintaining the trial-specific novelty of the shapes by presenting feedback (point only, point+shape) in a blocked design. In this experiment, win and no-win ERPs were able to be disassociated in the novel feedback condition following the removal of the trial-to-trial outcome surprise. However, the novel point+shape rewards still evoked an enhanced N2 component that differed from the point only rewards. Taken together these results reveal a major methodological complication for examining the RewP in experimental paradigms using complex, ecologically valid rewards. In a third experiment I rectified this issue by leveraging principles of reinforcement learning and presented a colored cue indicating a win or no-win outcome before presenting participants with novel, idiosyncratically chosen positive images. As hypothesized, the RewP as well as frontal midline delta migrated to the colored cue, whereas the N2 and theta power did not, thus revealing a solution to this methodological complication.

The application of spectral decomposition was able to isolate frequency bands (delta and theta) related to distinct information content in these interacting outcomes. In Experiment 1, I saw a double dissociation between our two frequency bands of interest; where delta power was able to disassociate outcome valences (win > no-win), and theta was able to disassociate novelty conditions (point+shape > point-only). Taken together, both analytic (time frequency analysis) and methodological (the migration of the reward signal to the cue) solutions were able to address this issue of novelty occluding the RewP signal. I also employed a spatial filter (surface Laplacian) in order to asses if this filtering technique would be able to differentiate these signals. I saw that this filtering technique was not only unable to disentangle novelty-reward interactions, but it was unable to distinguish win from no-win outcomes for the point-only condition, suggesting this technique is not ideal for examining the RewP. These findings suggest new directions for EEG-related reward experimentation utilizing novel ecologically valid rewards.

Novelty can be thought of a special type of surprise (Barto, Mirolli, & Baldassarre, 2013; Ferrari, Bradley, Codispoti, & Lang, 2010) that relies on a system of comparing the context and content of a set of expectations that have been violated. A leading hypothesis describing novelty-reward interactions centers around the theory that novel environments promote reward searching circuitry in the brain (Bunzeck et al., 2012; Kakade & Dayan, 2002), leading to intrinsic motivational behaviors of an agent. The promotion of exploratory behaviors is supported by computational models and theories that relate to reinforcement learning paradigms (Sutton & Barto, 2018). Similarly, signals of RPE can also be thought of as a special type of surprise whereby outcomes that are better or worse than expected evoke such signals. Functional neuroimaging studies have implicated specific neural structures as key-players in the maintenance of these prediction errors such as the orbitofrontal cortex, the amygdala, and the striatum (O’Doherty, 2004). Activation of these same neural areas appear sensitive to the processing of novel stimuli as well (Bunzeck, Dayan, Dolan, & Duzel, 2010; Bunzeck & Düzel, 2006) suggesting a neural cross-talk during the processing of rewarding stimuli and stimuli that are new.

By leveraging principles of reinforcement learning and pairing the attainment of reward with a cue predicting its receipt, I was able to examine RewP in the context of emotionally informative rewards, without the signal being superseded by neural signals of surprise and novelty. These findings open up the possibility for new testable hypotheses relating to complex rewards that promote intrinsic motivational states. In Experiment 3 ERPs were larger for cues predicting images than for cues predicting points. Although this effect wasn’t specific to win conditions only, I suspect the main effect to this type of outcome was a product of the recruitment additional motivational circuitry in their processing. These findings suggest an important direction for investigations utilizing motivationally relevant stimuli as rewards. Research investigating substance use disorder and major depression have found imagery relating to these clinical population (i.e. alcohol cues or emotional images) recruit distinct motivational processes (David et al., 2005; George et al., 2001; Tremblay et al., 2005). Recent evidence suggests that depressed symptomology predicted a smaller RewP yet this did not affect the ability to learn from reward (Cavanagh et al., 2019; Kumar et al., 2008). The authors suggest this diminishment of the RewP may reflect aberrant valuation (e.g. liking) in the context of maintained prediction error signaling for learning. Future investigations may adopt the use of these complex rewards in order to further identify interactions of motivational states and reward signals in order to gain a better understanding of how reinforcement and valuation systems are aberrant in clinical populations.

Chapter 3

AN EEG INVESTIGATION OF REWARD BASED LEARNING SIGNALS FOR DIFFERING TYPES OF REWARD

AKA

IT’S A MANS WORLD: AN EEG INVESTIGATION OF GENDER DIFFERENCES OF REWARD LEARNING

Abstract

Investigations of motivational aspects of emotion and how it relates to reward learning can offer important insights into both healthy and diseased population. Recent accounts have shown specific clinical groups evoked blunted neurological reward signals yet are still able to learn from rewards. This suggests an emotional feature of rewards which responds to hedonic aspects of reward attainment. The reward positivity (RewP) and delta band activity have been shown to be reliably evoked following outcomes that are better than expected. Although studies have revealed that the use of ecologically-relevant rewards, such as pleasant images, tend to produce signals that occlude the RewP, underlying delta band power has been shown to be free of these methodological confounds. The present series of studies aimed to address how individuals learn when provided differing types of rewarding feedback. During these tasks, participants performed a reinforcement learning (RL) task, where they had to learn to win rewards while avoiding punishments. During each trial participants received either points or pictures as feedback. Following the RL task, participants performed a BuyBack task were given the chance to trade in the points they earned in order to see positive instead of negative pictures. RewP and delta band power activation was also investigated. The results revealed that female participants were worse than males at learning from picture types of rewards, and these trends were also seen in the EEG data. A follow up study was conducted on only female participants which replicated the effects from Experiment 1. In both tasks, a major predictor of a participants learning accuracy for picture rewards was the amount of points traded in during the BuyBack task, suggesting that learning from rewards relies on hedonic motivational processes related to the actual reward.

Experiment 1

Signals of reward prediction error (RPE) are evoked for a special case of surprise following outcomes that are both more surprising than expected (Krigolson, 2017; Rescorla & Wagner, 1972; Schultz et al., 1997) and also coded as either better or worse than expected. These signals appear sensitive to trial-and-error outcomes, making them candidate biological markers for reinforcement learning (RL) paradigms. The reward positivity (RewP) in particular has gained a great deal of attention as a signal of positive RPE, in that it is only evoked to rewards and it is larger when outcomes are better than expected. Over the past decade, numerous studies have elucidated important features of the RewP and its utility as a clinical tool (Cavanagh et al., 2019; Proudfit, 2015). However, many experimental paradigms used to evoke the RewP have relied on a single type of reward indicating good-versus-bad outcomes. This overreliance on a single type of simple rewarding outcome limits the understanding of the boundary conditions of the RewP and ignores additional motivational processes, which may modulate this signal independent of expectation violation, as is seen in disorders like depression (Cavanagh et al., 2019), schizophrenia (Schneider, Gur, Gur, & Shtasel, 1995), or Parkinson's disease (Benke, Bosch, & Andree, 1998). Here, I aimed to investigate specific features of the RewP in an RL paradigm where participants could earn visually complex, ecologically-valid reward types—emotional pictures.

The RewP is a positive going deflection appearing ~200ms over frontal central sites following reward receipt. Spectral decomposition of this signal has revealed that delta band frequency power (Bernat, et al., 2008) underlies its generation. Numerous studies have shown that the RewP is sensitive to the acquisition of simple forms of reward such as points (Bellebaum, Polezzi, & Daum, 2010; Wu & Zhou, 2009; Yeung &

Sanfey, 2004), or abstract icons such as fruit or colored arrows (Angus et al., 2015; Holroyd et al., 2008; Proudfit, 2015) . However, very few studies have attempted to examine how the RewP responds to stimuli closer to real world outcomes. A recent study attempted to evoke a RewP to pleasant images (Brown & Cavanagh, 2018) and found that pleasant images did not evoke a RewP but an enhanced N2 component instead. The researchers concluded that the RewP might be hidden by a signal of novelty (the N2). This methodological complication further hinders our understanding of temporally precise reward processing for these motivationally salient stimuli types.

Investigations of motivational aspects of emotion and how it relates to reward learning can offer important insights into aberrant processes in clinical populations. Indeed, recent investigations of hedonic types of reward have further added to the growing understanding of clinical phenomena such as autism spectrum disorder (Sabatino, Richey, Bodfish, Dichter, & Rittenberg, 2011), depression (Kumar et al., 2008), and alcohol use disorder (Schacht, Anton, & Myrick, 2014). EEG is an excellent tool for the study of these reward-emotion interactions due to its sensitivity to real-time canonical neural computations (Cavanagh et al., 2019; Fries, 2009). However, as mentioned above, there are numerous methodological issues to take into account when designing EEG experiments that utilize these ecologically relevant rewards.

Although a move towards real-world-like rewarding stimuli would expand our understanding of hedonic aspects of the RewP, there are several methodological complications for the use of this type of rewarding stimuli in EEG experiments. First, using complex emotional imagery as rewarding feedback presents a novelty-induced masking of the RewP signals (Brown & Cavanagh, 2018). Additionally, the use of emotional stimuli must also be appropriate for the subject. Much of the work relating to emotional processes in neuroscience has utilized images from the International Affective Picture System (Lang, et al., 1997). Although the IAPS is a well-established tool in the field, individual differences for emotional stimuli processing adds an additional layer of complexity in its use as reinforcing stimuli (Hamann & Canli, 2004). Finally, recent accounts have demonstrated gender differences in the perception and processing of emotional stimuli (Bradley, Codispoti, Sabatinelli, & Lang, 2001). These gender differences relating to emotional image processing may also present a complication for the use of this emotional picture types of reward.

In a simple image rating task, Bradley and collogues (2001) demonstrated that females tend to show a greater sensitivity towards negative stimuli, while males tend to be highly sensitive to appetitive positive images. Other researchers have replicated these findings (Calvo & Avero, 2009; Kemp, Silberstein, Armstrong, & Nathan, 2004; Wrase et al., 2003). These gender effects are further supported by results from an electrophysiological study (Kemp et al., 2004), whereby females showed greater steady state visually-evoked potentials to negative images than males. These gender differences in emotional processing may further complicate how males and females learn from rewards in emotion induction experimental paradigms. Indeed, a recent study revealed that during a risky decision-making task, stress induction led to better reward related decision making in male participants, and worse performance in females (Lighthall et al., 2012). Magnetic resonance imaging in these participants revealed increased activation in reward centers (dorsal striatum and anterior insula) in male participants during stress induction, while females showed decreased BOLD signals in these same neural areas. Taken together, males appear to perform better for outcomes related to learning to approach rewards, whereas females tend to perform better in situations that require the avoidance of punishment.

In the current study, I examined learning rates in a RL paradigm for differing types of rewarding feedback: points and pictures. Participants were required to learn the optimal response to a colored cue resulting in winning points (each amounting to 5 cents) or pleasant images. Additionally, participants had to learn to avoid punishing feedback (losing points or viewing disturbing negative images). It was predicted that the accuracy for learning from picture rewards would relate to how motivated a person was to obtain it. To this end, participants performed a task following the learning task where they were given the opportunity to trade actual money in order to see positive images instead of negative images. I predicted that the amount of points (money) a participant was willing to trade would relate to their accuracy to learn from cues predicting pleasant picture rewards. Based on previous findings demonstrating RewPs evoked for emotional pictures are blocked enhanced N2 component (Brown & Cavanagh, 2018) I also hypothesized that ERPs evoked for the emotional feedback will be blocked in the same manner. However, prior evidence has revealed that delta band power is unaffected by visually complex reward stimuli (Brown & Cavanagh, in prep). Here, I hypothesized that delta band power will be unaffected by this constraint and will provide us a neural measure of reward processing for these reward types. Finally, due to the observed gender differences in emotional processing, I predicted that males would learn the optimal response to cues predicting the winning of rewards better than females, while females would learn the optimal response to cues predicting avoidance of punishment better than males.

Method

Participants

In Experiment 1, I recruited 84 participants (48 females) with a mean age of 20.47 (*SD* = 6.13). In all experiments, participants were recruited from the University of New Mexico subject pool. Students received class credits for participation. Participants were excluded from participation if they had a history of head injury that resulted in loss of consciousness for more than five minutes, had a history of epilepsy, had a history of any psychiatric or neurological disorder, or were currently on any psychiatric or neurological drugs. The Institutional Review Board of the University of New Mexico approved the study protocol.

Procedure

Affective reward preference task

In order to choose idiosyncratically preferred images for picture rewards, a reward preference task was performed prior to the RL tasks. The task was programmed in Matlab using Psychtoolbox (Brainard, 1997). During this short two-alternative forced choice task, participants were presented two images drawn from 1 of 5 affective reward categories: male models, nude women, puppies, babies, or nature scenes. Image classes were based on images that were rated as highly pleasurable from the IAPS (Lang et al., 1997). Standardized ratings of valence (1=negative to 9=positive) from the IAPS technical manual were compared by gender. The reward classes used for the current study were chosen from the most occurring image themes. All images were selected from Internet searches (i.e. “Hi-Definition Puppy Images”). All images were inspected in task-presentation conditions in order to assess that no image appeared blurry and that sizing was equal among images. During the choice task, an image from one of the affective reward categories was presented on the left side of the screen while another image from a different affective reward category was presented on the right side. Participants were instructed to choose with a left or right button press which image of the two they preferred. There were 16 images in each category; each category was paired against each of the other categories four times (total of 40 trials). If two classes were chosen the same amount of times, a special “tie-breaker” trial was conducted which presented two never-before-seen images from the tied classes. The participant's top choice was selected for the RL task.

Reinforcement learning task

Figure 12 displays the sequence of events in an individual trial of the RL task. This consisted of a forced choice training phase followed by a subsequent testing phase. During the training phase, the participants were presented with a cue shape and were instructed to use an appropriate button press (left button or right button) in order to learn how to earn rewards. Each cue shape was associated with a different reward outcome (win-point, avoid-point, win-picture, avoid-picture). Participants were instructed that point feedback (green +1, yellow bar, red -1) would be calculated and after the task was complete, would be traded in for actual money (5 cents USD). Participants were informed that picture feedback would not affect the money they earned. Each cue had a 70% probability of reward. All training trials began with a jittered inter-trial interval between 300 and 700ms. The stimuli then appeared for a maximum of 4000ms and disappeared immediately after the choice was made. If the participant failed to make a choice within the 4000ms, “No Response Detected” was presented. Following a correct button press, participants would receive a green +1(correct win point condition), a yellow bar (correct avoid point condition), a pleasant picture selected from the participant’s top choice (correct win picture condition), or a picture of a chair (correct avoid picture condition). Incorrect button presses were followed by either a yellow bar (incorrect win point condition), a red -1(incorrect avoid point condition), a chair (incorrect win picture condition), or a negative picture feature mutilation (incorrect avoid picture condition). Note that the use of the terms "correct" and "incorrect" throughout refer to the feedback, not to the optimal or accurate response. All trial feedback was presented for 2000ms. During the testing phase, all possible cue shape pairs were presented eight times (120 trials total) and no feedback was provided. For the testing phase, participants were instructed to choose (with a left or right button press) the cue that led to the most rewarding outcome. After completing the testing phase, participants were informed of their point winnings. Critically, all participants were told that they won 66 points (equal to $3.30 USD) regardless of their actual performance. The task took ~28 minutes on average to complete.

BuyBack Task

Following the reinforcement learning task, participants performed an image rating task. Participants were informed that they would be exposed to 50 emotionally negative images. Critically, participants were informed that they would be able to swap a negative image out for a positive image, however, this swap would cost them 1 point (5 cents). Participants were informed that they could trade as much or as little of their points (up to 50) to see positive pictures. The task began once participants made their decision regarding how many points to swap out and took ~9 minutes to complete.

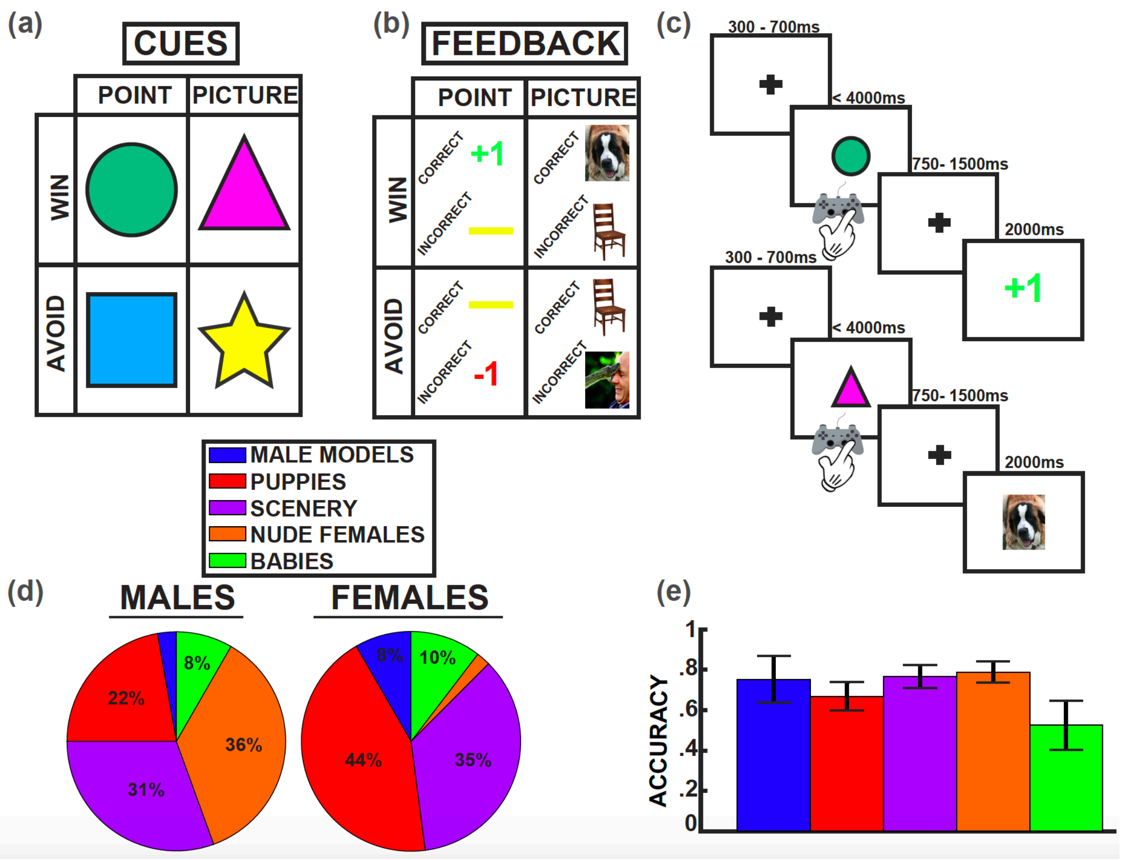

Figure 12. Reinforcement Learning Task. During the task, participants were presented with a cue (a) predicting a specific reward type (points or pictures) and outcome (win or avoid). Participants would make a button press (left or right) for each cue and would be presented feedback (b). (d) Display of typical experimental trials leading to point or picture feedback. Pie graphs (d) display the percentage of each reward type chosen by the participants for each gender group. Bar plots (e) display accuracy scores across picture reward types. There were no significant accuracy differences between the picture types.

Data acquisition and preprocessing

Electrophysiological data were collected with a 64Ag–AgCl electrodes embedded in a stretch-lycra cap with a sampling rate of 500Hz with low and high cutoffs at .01-100 Hz. CPz served as the reference electrode and FPz as the ground electrode. Data was recorded with a Brain Vision system (Brain Products GmbH, Munich, Germany). Vertical electrooculogram (VEOG) activity generated by blinks was recorded by two auxiliary electrodes placed superior and inferior to the left pupil.

All EEG processing was conducted in EEGlab (Delorme & Makeig, 2004). First, CPz was re-created via computation of the average reference (EEGlab function pop_reref.m). Very ventral electrodes (FT9, FT10, TP9, and TP10) were then removed, as they tended to be unreliable. The average reference was then recomputed for the remaining 60 electrodes. EEG data for ERPs was filtered between .01 to 20 Hz. Data was epoched around feedback screen onset (-1000 to 3,000ms). Using statistical deviations from the mean for each EEG channel, FASTER (Nolan et al., 2010) identified artifacts in each epoch for later rejection. Eye blink activities were removed following ICA (runica; Makeig, Bell, Jung, & Sejnowski, 1996). Epochs were then baseline corrected (-200 to 0ms before feedback onset) and averaged to calculate event related potentials (ERP). The RewP was quantified at electrode site Cz and was measured between conditions within a 200 – 350ms window post feedback onset.

Time-frequency measures were computed using custom-written MATLAB functions (Cohen, 2014) by multiplying the fast Fourier transformed (FFT) power spectrum of single trial EEG data with the FFT power spectrum of a set of complex Morlet wavelets (defined as a Gaussian-windowed complex sine wave: $e^{i2\pi tf}e^{-t\wedge 2/(2x\sigma\wedge 2)}$, where t is time, f is frequency (which increases from 1–50 Hz in 50 logarithmically spaced steps) and defines the cycles of each frequency band, set according to $4/(2\pi f)$), and taking the inverse FFT. The end result of this process is identical to time-domain signal convolution, and it resulted in estimates of instantaneous power (the magnitude of the analytic signal), defined as Z[t] (power time series: $p(t)=real[z\ (t)]^2 + imag[z(t)]^2$). Each epoch was then cut in lengths (-500 to 1,500ms). Power was normalized by conversion to a decibel scale (10 x log10[power(t)/power(baseline)]), allowing a direct comparison of effects across frequency bands. To account for spectral smearing (Cohen, 2014) frequency power plots were measured between conditions within a 200 – 400ms window post feedback onset.

Statistical Analyses

Greenhouse-Geisser adjusted ANOVAs and planned comparison decompositions were used for data analyses. Reports of effect sizes for ANOVA are partial-η^2, while planned comparison effect sizes are reported as *d.* Due to the non-normality of the distributions of traded points during the BuyBack task, correlations were conducted using Spearman's Rho (rho). Differences between correlation coefficients were compared using Fishers r-to-z transform and z-test.

Results

Behavioral Analysis

Figure 13 depicts accuracy scores across the all trials. A participant's response accuracy score was calculated using the average accuracy scores (correct button selection) from all the trials across experimental conditions in the RL task. I first assessed if participants' performance differed between image classes. I conducted a one- way ANOVA comparing response accuracy across the five image class conditions. This analysis failed to reveal a significant effect between the image classes ($F(4,79) = .955$, $p = .437$, $\eta^2 = .046$). Additionally, I compared points traded during the BuyBack task across image classes. This analysis also revealed no significant differences between the image classes ($F(4,79) = 1.029$, $p = .398$, $\eta^2 = .050$). Taken together, these results suggest that learning rates or the motivation to learn from image feedback did not differ between image classes.

The 2 (REWARD TYPE: point vs. picture) × 2(OUTCOME: win vs. avoid) × 2(GENDER: male vs. female) ANOVA for response accuracy revealed a significant REWARD TYPE (point > picture) main effect ($F(1,82) = 10.055$, $p = .002$, $\eta^2 = .109$), as well as a GENDER (male > female) main effect ($F(1,82) = 4.046$, $p = .048$, $\eta^2 = .047$), but no OUTCOME main effect ($F(1,82) = 1.602$, $p = .209$, $\eta^2 = .019$). The ANOVA also revealed a REWARD TYPE × GENDER interaction ($F(1,82) = 4.559$, $p = .036$, $\eta^2 = .019$), whereby response accuracy for males did not significantly differ between point and picture types of reward ($t(35) = .899$, $p = .375$, $d = .145$); however, there was a significant response accuracy difference (point > picture) for female participants ($t(47) = 3.538$, $p = .001$, $d = .560$). Lastly, in order to test if optimal response accuracies were above chance, I conducted a one-samples t-test comparing accuracy for cues predicting pleasant pictures against a baseline of 50% in female participants. This analysis revealed a non-significant effect ($t(47) = 1.926$, $p = .060$, $d = .505$), suggesting females accuracy scores were not better than chance. Taken together, it appears that females did not learn from the picture rewards. There were no other significant interactions in the ANOVA (all p's > .127).

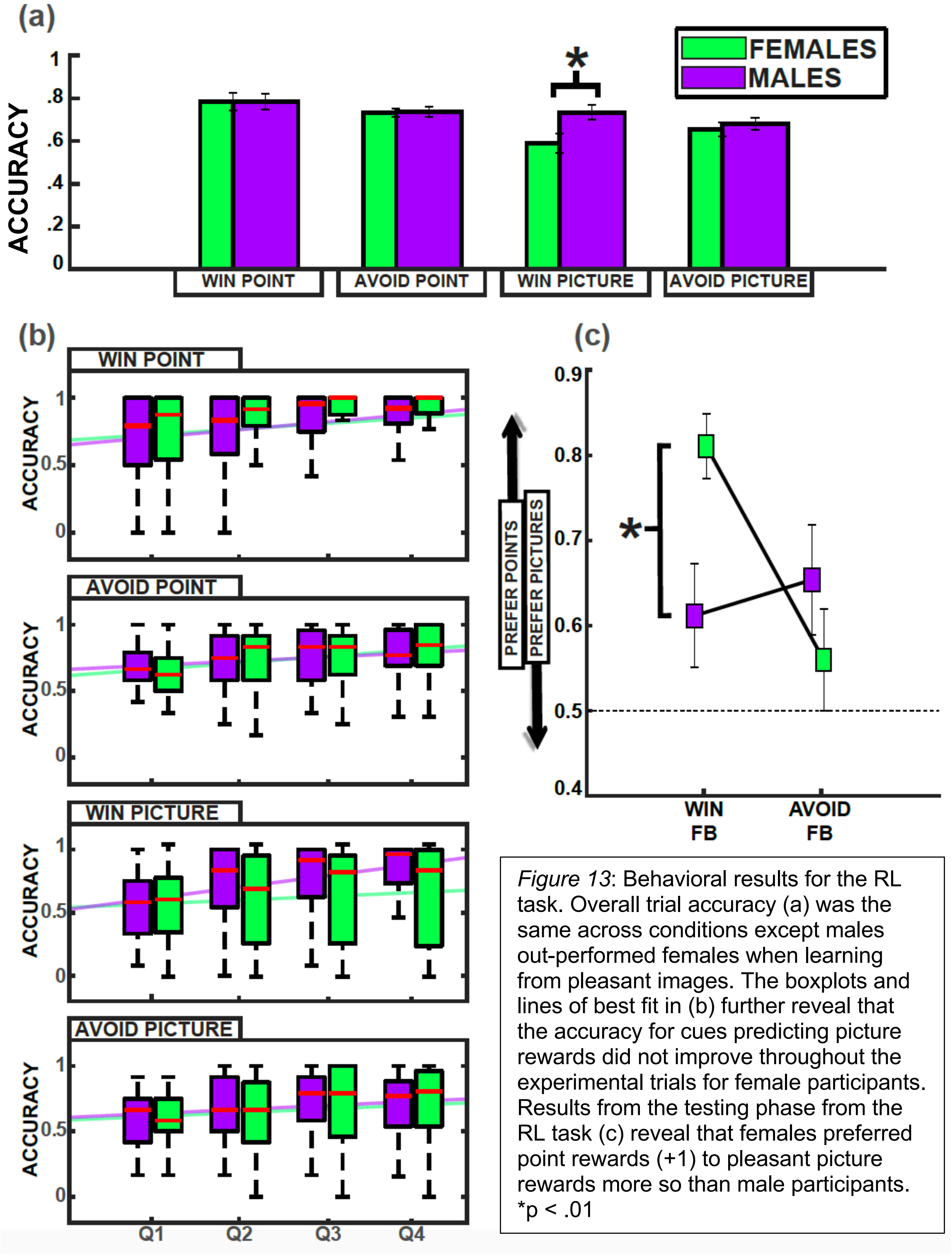

Figure 13: Behavioral results for the RL task. Overall trial accuracy (a) was the same across conditions except males out-performed females when learning from pleasant images. The boxplots and lines of best fit in (b) further reveal that the accuracy for cues predicting picture rewards did not improve throughout the experimental trials for female participants. Results from the testing phase from the RL task (c) reveal that females preferred point rewards (+1) to pleasant picture rewards more so than male participants. *p < .01

To further unpack this gender effect, a 2(OUTCOME) × 2(GENDER) ANOVA was conducted comparing point > picture cue preference from the testing accuracy phase of the RL task (Figure 13c). There was no OUTCOME main effect ($F(1,82) = 2.992$, $p = .087$, $\eta^2 = .035$), suggesting that cues predicting points and cues predicting pictures were seen as equally rewarding. However, there was an interaction ($F(1,82) = 5.896$, $p = .017$, $\eta^2 = .067$), whereby females preferred cues predicting winning points over cues predicting winning pictures more than males ($t(82) = 2.890$, $p = .005$, $d = .618$). The two gender groups did not differ in the preferences for cues predicting avoiding punishment ($t(82) = 1.058$, $p = .293$, $d = .225$).

As a final test to unpack this gender effect I examined relationships between response accuracy scores for winning picture rewards and the number of points participants were willing to trade during the BuyBack task between the two gender groups. This analysis revealed a significant correlation for females (rho(48) = .305, $p = .035$) but not for males (rho(36) = -.107, $p = .535$). An r-to-z transform was applied in order to compare the correlation coefficients revealing a trending to significance effect ($Z = 1.834$, $p = .066$). Figure 14 depicts these relationships.

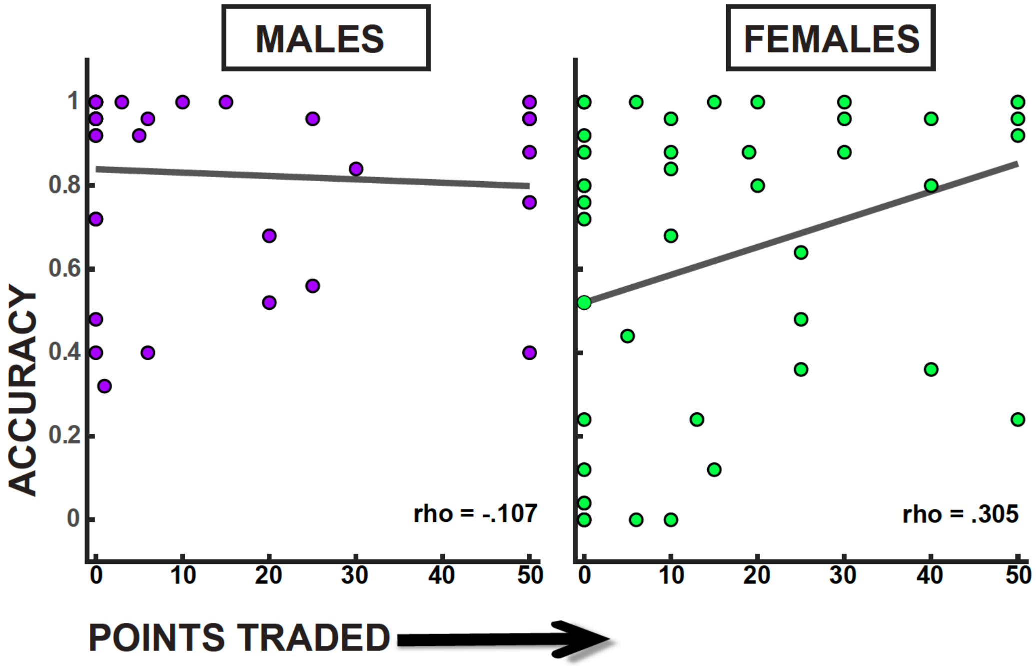

Figure 14. Correlations between accuracy for winning pictures in the RL task and points traded in during the BuyBack task. During the BuyBack task, participants were given the opportunity to trade in actual money (1 point = 5 cents) to see a positive image instead of a negative image. Female participants accuracy in the RL task for winning pictures was strongly related to the amount of money they were willing to trade in. This effect was not seen in male participants, whose accuracy scores were significantly higher than the female participants.

EEG Analysis

Figure 15 depicts ERPs and spectral decompositions plots to point and picture feedback. An analysis conducted comparing ERPs for reward feedback: +1 (correct win point), yellow bar (correct avoid point), pleasant picture (correct win picture), and chair image (correct avoid picture). This 2(REWARD TYPE: point vs. picture) × 2(OUTCOME: win vs. avoid) × 2(GENDER: male vs. female) ANOVA revealed a significant 3-way interaction ($F(1,82) = 6.563$, $p = .012$, $\eta^2 = .074$). In order to investigate this interaction, two 2-way ANOVA were conducted on the gender groups separately. The ANOVA for males revealed a significant REWARD TYPE main effect (point > picture: $F(1,35) = 7.613$, $p = .009$, $\eta^2 = .179$). There was no OUTCOME main effect ($F(1,35) = .438$, $p = .512$, $\eta^2 = .012$), or interaction ($F(1,35) = 2.150$, $p = .152$, $\eta^2 = .058$). The ANOVA for female participants revealed significant REWARD TYPE (point > picture) and OUTCOME (win > avoid) main effects ($F(1,47) = 140.891$, $p < .001$, $\eta^2 = .750$; $F(1,47) = 8.440$, $p = .006$, $\eta^2 = .152$, respectively). There was also a significant interaction ($F(1,47) = 29.917$, $p < .001$, $\eta^2 = .389$). In order to investigate this interaction, post-hoc t-test were conducted comparing ERPs for OUTCOME (win vs. avoid) condition across REWARD TYPE (point vs. picture) separately. These tests failed to reveal a significant difference for point rewards (t(47) = 1.704, p = .095, d = .243), however, there was a significant difference for picture reward (avoid > win: t(47) = 5.761, p < .001, d = .811).

The ANOVA for delta band power (300 – 1000ms) revealed a significant 3-way interaction ($F(1,47) = 8.440$, $p = .006$, $\eta^2 = .152$). In order to unpack this, two 2-way ANOVAs were analyzed on gender groups separately, however, this separation of gender groups failed to reveal any significant effects (all *p*'s > .132). Although there were no significant delta band effects, examination of the line plots for time frequency (Figure 15) reveal these effects are going in the expected direction, whereby delta band power was larger for pictures than points in male participants, yet smaller in female participants.

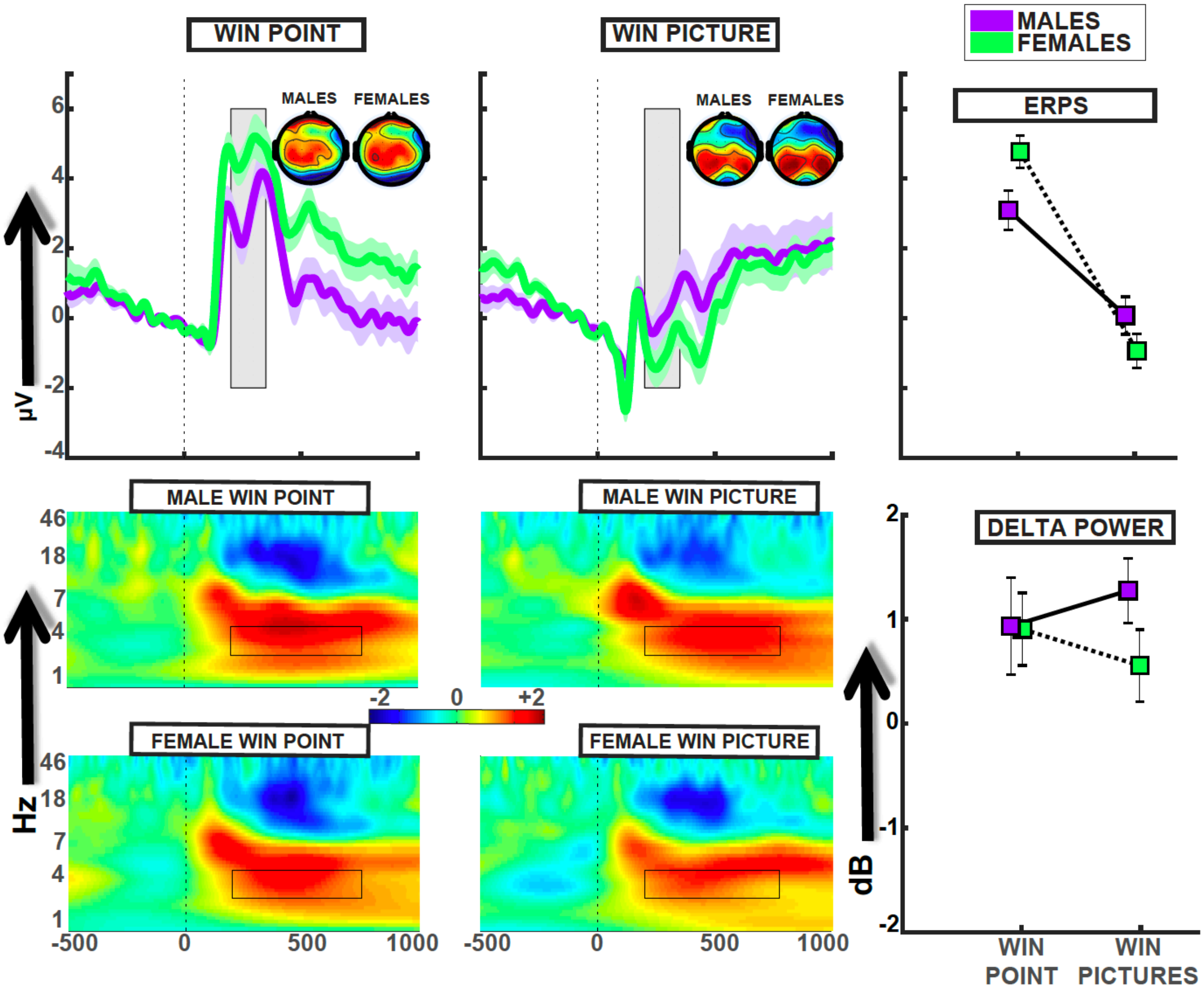

Figure 15. EEG dynamics for point and picture rewards. ERPs (TOP ROW) for reward outcomes reveal a significant smaller ERPs for picture rewards. (BOTTOM ROW) Although delta-band activity did not significantly differ between reward types, the effects are trending in the hypothesized direction, whereby there was more delta band power for picture rewards than point rewards in males, yet less delta band power for pictures in female participants.

Investigations of confounding variables

There may be underlying confounds influencing the ERP data due to the content of the images. To this end RewPs were compared across the picture types. This analysis revealed a significant difference in the RewP evoked for the different classes ($F(4,79) = 5.515$, $p = .001$, $\eta^2 = .218$), whereby ERPs evoked for puppies and scenery were significantly smaller than those evoked for nude women (vs. puppies $p < .001$; vs. scenes *p* = .001) and babies (vs. puppies *p* = .008 ; vs. scenes *p* = .029). From this effect, it may be a reasonable assumption that certain picture types were more rewarding than others. However, our analysis comparing accuracy and points traded during the BuyBack task revealed no significant differences, suggesting no one picture type was more rewarding than any other and that the picture type did not influence a participant's ability to learn from them. Indeed, comparisons of the underlying frequency band activity revealed no significant delta band power differences across picture types ($F(4,79) = .394$, $p = .812$, $\eta^2 = .020$). It was suspected that the differences revealed in the ERPs across the picture types related to differences of image content, not reward processing. To this end, I conducted a series of analyses probing these potential underlying variables. First, the individually selected picture types used for picture rewards were made up of images with humans (male models, nude women, and babies), puppies, and nature scenes. Interestingly, the differences seen in the ERPs across trials were significantly larger for images of nude women and babies. A well-studied ERP component, the N170, has been shown to be reliably evoked when viewing human faces and bodies (Blau, Maurer, Tottenham, & McCandliss, 2007; Hietanen & Nummenmaa, 2011; Rossion et al., 2000). The larger ERPs seen for images with human faces than non-humans may have been a product of an underlying N170 during image viewing. In order to address this, I tested the N170 component (at temporal electrode sites T7 and T8, between 150 - 190ms) between the picture classes. I found no N170 difference across class conditions ($F(4,79) = 1.385$, $p = .247$, $\eta^2 = .066$). Additionally, I examined the late positive potential (LPP) across the image classes. The LPP is a well-studied ERP component which is reliably evoked during the viewing of emotional images (Brown & Cavanagh, 2017; Gable & Harmon-Jones, 2010; Hajcak, Macnamara, & Olvet, 2010; Hajcak & Olvet, 2008). A recent study found that LPP amplitudes were larger for motivationally salient images (Weinberg & Hajcak, 2010). Here, I leverage the sensitivity of the LPP to motivationally salient stimuli to examine potentially content inducing salience difference across the picture types. This analysis (at parietal electrode site Pz, between 300 - 1000ms) revealed no difference across the image classes ($F(4,79) = .908$, $p = .464$, $\eta^2 = .044$). Taken together, there does not appear to be any electrophysiological differences across the picture types used in the study.

Another potential confounding variable in the study relates to the suboptimal learning rates of picture rewards in the female participants. Importantly, before I are able to suggest female participants were not motivated by the picture rewards, I must rule out an additional interpretation of these result. Indeed, one potential explanation for this effect may be that pleasant pictures were not motivating enough for the female participants to learn the optimal response to cues predicting them. A second explanation may be that female participants were unable to distinguish the win and no-win feedback in the picture condition. In this way they viewed the lose feedback in this condition (chairs) as equally pleasant as the positive emotion pictures. In order to address this issue, I conducted a second experiment with only female participants.

Experiment 2

In a second study I recruited only female participants (N=36). All experimental procedures were the exact same as Experiment 1, except participants were informed what to expect from the picture feedback. All participants received a scripted instruction from the experimenter saying “Positive feedback will be your favorite positive picture from the last task. Neutral feedback will be standard chairs. Negative feedback will be disturbing or graphic images”. This single manipulation in the instructions was done in order to make sure the participants knew that positive images were better outcomes than the neutral images.

Figure 16 depicts behavioral results for Experiment 2. The statistical analysis used for Experiment 1 was also used for Experiment 2 except gender was not used as a variable. The 2(REWARD TYPE: point vs. picture) × 2(OUTCOME: win vs. avoid) ANOVA for response accuracy failed to reveal a REWARD TYPE main effect (point > picture: $F(1,35) = 3.893$, $p = .056$, $\eta^2 = .100$), or an OUTCOME main effect ($F(1,35) = 2.943$, $p = .095$, $\eta^2 = .078$). There was a significant interaction ($F(1,35) = 4.637$, $p = .038$, $\eta^2 = .117$). Post hoc comparisons comparing ERPs for OUTCOME across REWARD TYPE failed to reveal a significant accuracy differences for point ($t(35) = .330$, $p = .743$, $d = .044$) but there was an OUTCOME accuracy difference for picture rewards (avoid > win: $t(35) = 2.390$, $p = .022$, $d = .310$). A one-samples *t*-test comparing accuracy for cues predicting pleasant pictures to a baseline of 50% revealed a non-significant effect ($t(35) = 1.453$, $p = .155$, $d = .242$), thus replicating the effects in Experiment 1 suggesting that the female participants in Experiment 2 did not learn from picture feedback (Figure 16a). I conducted a paired samples t-test comparing point > picture cue preference scores obtained during the testing phase of the RL task. This analysis revealed a significant effect ($t(35) = 2.1335$, $p = .040$, $d = .594$), whereby participants preferred cues which predicted winning points over cues predicting winning . Finally, accuracy scores were compared to the magnitude of BuyBack points, and similar to Experiment 1, a significant relationship was again revealed between the two measures (rho (37) = .368, $p = .027$).

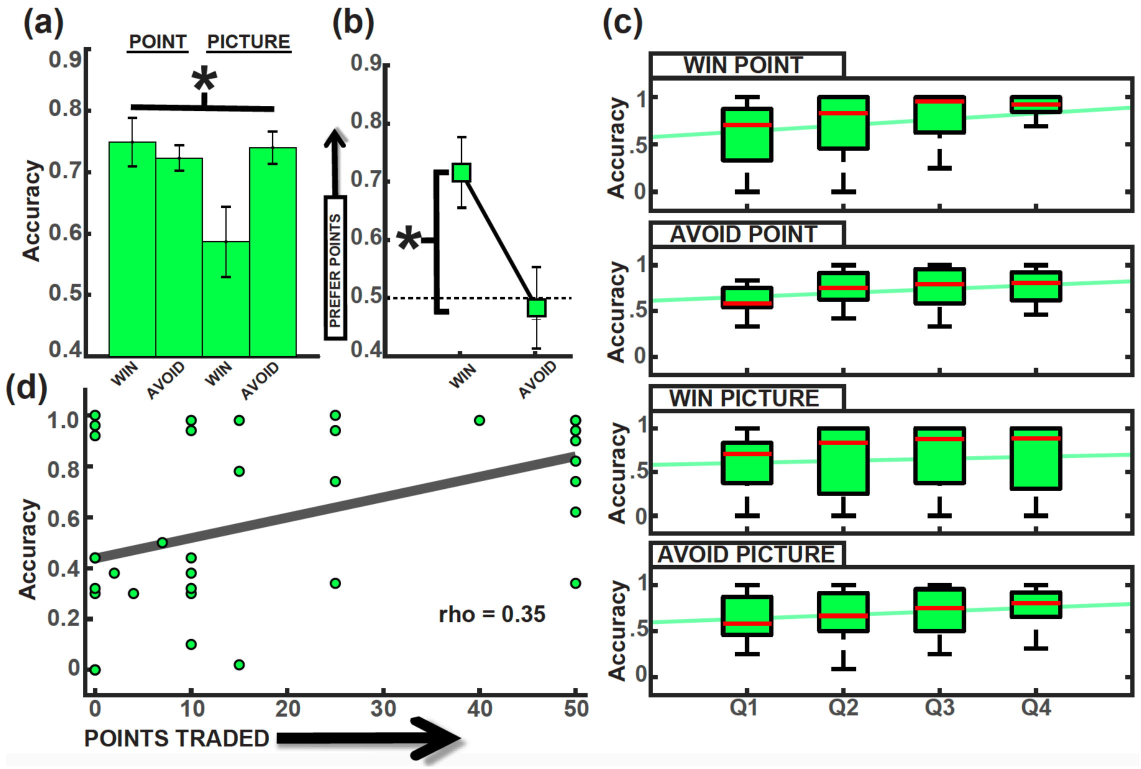

Figure 16. Behavioral results from Experiment 2. Accuracy scores (a) reveal that the participants learned the optimal response to all cues except those that predicted pleasant images. (b) Displays mean preference scores between cues during the testing phase of the RL task. There was a significance effect between point feedback preferences, whereby participants preferred cues predicting winning points over pleasant picture, yet preferred cues predicting neutral images over neutral points. The boxplots and lines of best fit in (b) further reveal that the accuracy for cues predicting picture rewards did not improve for female participants even after being told what the correct feedback for this cue was. (d) Depicts the scatterplot comparing points traded during the BuyBack task and response accuracy for cues predicting pleasant pictures. Replicating the effects form Experiment 1, there was a significant positive relationship. *p < .05

Figure 17 depicts ERPs and spectral decompositions plots to point and picture feedback. This 2(REWARD TYPE: point vs. picture) × 2(OUTCOME: win vs. avoid) ANOVA replicated the ERP effects seen in Experiment 1, whereby there was a REWARD TYPE main effect (point > picture; $F(1,35) = 80.789$, $p < .001$, $\eta^2 = .698$) and an OUTCOME main effect (avoid > win: $F(1,35) = 8.123$, $p = .007$, $\eta^2 = .188$). There was significant interaction ($F(1,35) = 86.790$, $p < .001$, $\eta^2 = .713$). In order to investigate this interaction, post-hoc t-test were conducted comparing ERPs for OUTCOME across

REWARD TYPE separately. These tests revealed a significant difference for point rewards (win > avoid: t(35) = 2.889, p = .007, d = .440), as well as a significant difference for picture reward (avoid > win: t(35) = 7.683, p < .001, d = 1.297).

The ANOVA for delta band power (300 – 1000ms) revealed a significant REWARD TYPE main effect (point > picture: $F(1,35) = 7.828$, $p = .008$, $\eta^2 = .698$). There was no OUTCOME main effect or interaction ($F(1,35) = .062$, $p = .804$, $\eta^2 = .002$; $F(1,35) = .414$, $p = .524$, $\eta^2 = .012$, respectively).

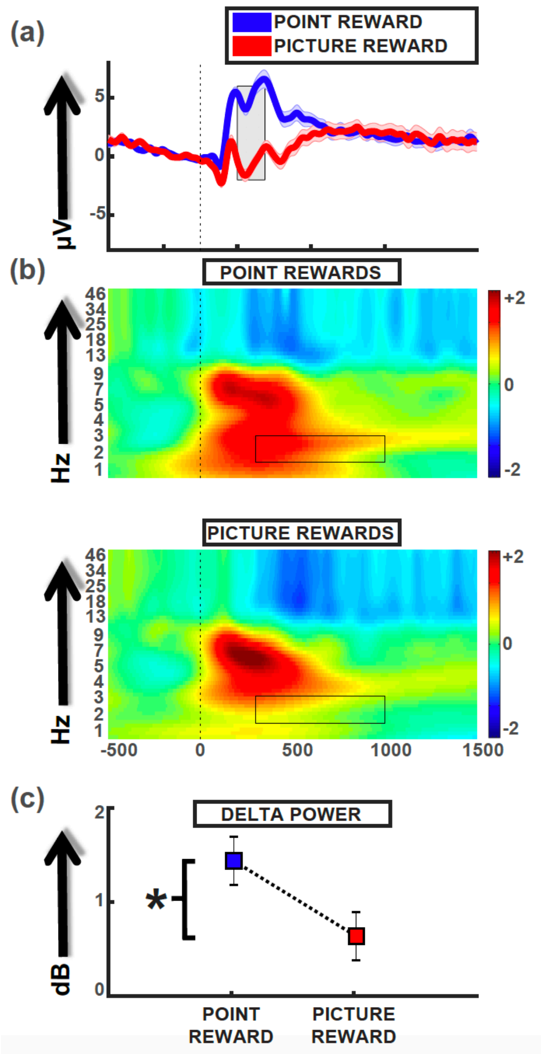

Figure 17. Experiment 2 EEG dynamics for point and picture rewards. ERPs (a) for reward outcomes reveal significantly smaller ERPs for picture rewards. Replicating the effects in Experiment 1, delta band power (b & c) mirrors accuracy rates, whereby there is larger point evoked delta band power than picture evoked delta band power. *$p < .05$

Discussion

In the current study, we examined learning and reward signals for differing types of rewards in a RL experimental paradigm. During the experiment, participants had to learn, through trial-and-error, to respond to shapes predicting different types of rewards (points or pictures) and different outcomes (win reward or avoid punishment). Additionally, motivated by the observed gender differences for emotional image processing, a variable of gender was added into the analyses. The results revealed that although there was no difference between the gender groups for learning from point rewards, male participants were better at learning from picture rewards than females. Results from the testing phase revealed that females found cues predicting winning points more rewarding than cues predicting winning pictures than males, suggesting that female participants were not as motivated by picture types of feedback as they were for points. This finding may explain why female participants did worse at learning the optimal response to pleasant picture rewards than males. To further support this interpretation, a correlation comparing accuracy and points traded during the BuyBack task revealed a significant positive relationship between the accuracy rates for winning pleasant pictures and points traded only in female participants. Additionally, a second experiment was conducted with only female participants. In this second study, participants were informed of the picture types they would potentially be viewing. This manipulation was conducted in order to assess that the participants knew the feedback they would be receiving so that reward types (positive picture and chairs) would not be confounded. The results of the second experiment replicated the effect from Experiment 1, whereby female participants failed to learn from pleasant images, however, the accuracy to learn from this type of reward was again correlated with the amount of points traded during the BuyBack task. Taken together, these behavioral results suggest that the motivation to obtain rewards in a trial-and-error learning paradigm relies heavily on an agent's hedonic appraisal of the reward.

The ERP results revealed gender differences as well. Although the ERPs differed between the types of reward (point > picture) for both gender groups, only in female participants was there a significant REWARD TYPE×OUTCOME interaction. Critically, in the unpacking of this interaction it was revealed that picture feedback for avoiding punishment (chairs) evoked significantly larger ERPs than picture feedback for winning rewards (pleasant pictures), suggesting that a gender-based difference in motivation played a role in the generation of this signal. Interestingly, delta power was larger for picture than for points rewards for males, yet smaller for pictures than for points for females. Although this effect was not statistically significant, the findings were in the expected direction.

The results of the current study suggest an important aspect of reward learning that is overlooked in human experiments examining these phenomena—an aspect of hedonic liking. To date, investigations of the RewP have mainly focused on information content of the generation of this reward signal, in that aspects of this signal relate to outcomes indicating some sort of win-over-lose signal. Although these studies have elucidated numerous factors which influence this signal, such as the size of reward magnitude (Bellebaum, Polezzi, & Daum, 2010; Wu & Zhou, 2009), the probability of reward occurrences (Bellebaum & Daum, 2008; Cohen & Ranganath, 2007), and the emotional state of the individual receiving the rewards (Angus et al., 2015), the over reliance on simple win-lose outcomes ignores a whole dimension of reward processing. In the current study it was revealed that emotional imagery has the potential to change a person's motivational states towards rewards, and thus their ability to learn from these outcomes. Indeed, the results of the current studies are a first step towards addressing how emotion influences hedonic aspects of reward and not just the probability of occurrence.

There were some limitations in the current studies. First, the ERPs evoked for picture feedback were visually different than those evoked for points, suggesting a separate process in their modulation. However, this effect was not surprising. In an earlier report, researchers attempted to utilize pleasant imagery in order to evoke the RewP (Brown & Cavanagh, 2018) and found that the ERPs evoked for this type of reward contained an enhanced N2 component. The authors suggested that this feature of the ERP was a product of novelty. This effect was demonstrated in a later report (Brown & Cavanagh, in prep) where the inclusion of novel shapes with rewarding feedback evoked an enhanced N2, thus blocking the RewP. Here, it was revealed that picture feedback, both pleasant and neutral images, evoked enhanced N2 components. However, spectral band decomposition was able to isolate the reward signal without interference from other aspects of the rewarding picture. Additionally, a future study may aim to work around this complication by leveraging principles of RPE and presenting a cue predicting the oncoming rewarding image. A well-observed characteristic of RPE signals is their migration to the first instance of reward (Angus et al., 2015; Schultz et al., 1997). Critically, a study employing this type of experimental design would be able to leverage measures common for emotion processing (i.e., valence and arousal ratings) for the reward stimuli. These measures, along with other measures of potential hedonic liking (e.g., BuyBack scores, preference scores, etc.) would allow researchers to isolate components of the RewP relating to prediction error from those relating to liking.

A second limitation relates to internal-external validity trade-offs due to the use of differing classes of rewarding imagery. Figure 12c illustrates how the class of image chosen for reward feedback varied between and within gender groups. Although I ultimately aim to utilize this line of experimentation in specific clinical populations (e.g., displaying alcohol cues to people with alcohol use disorder, or emotional imagery to people with major depressive disorder), the use of these varying types of imagery may have influenced the results. Indeed, when comparing the RewP across classes I saw significantly different RewP. Importantly, image classes not containing humans evoked smaller ERPs than those that did contain humans. It may be that aspects of these images evoke separate neural systems or motivational processes which in turn affect the RewP. This limitation was addressed in two ways. First, it has been demonstrated that images containing humans evoke higher rates of motivational significance than images not containing humans (Weinberg & Hajcak, 2010). To account for this potential motivational confound, I compared accuracy and BuyBack scores across the five picture classes and found no difference between the image groups. Second, in order to investigate if the presence of humans produced signals which interfered with the reward signals of interest, I investigated two ERPs shown to be sensitive to images containing humans. For one, a well-known ERP component called the N170 has been shown to be sensitive to face processing (Eimer & Holmes, 2002). The generation of this signal to images containing human faces (i.e., male models, nude women, and babies) may have had an influence on the reward signals for these image classes that was not present in image classes without human faces (i.e. puppies and scenery). Similarly, larger amplitudes of an ERP component related to an emotional process known as the late positive potential have been demonstrated in images containing humans (Weinberg & Hajcak, 2010). Again, the differences in the content of emotional image stimuli may have been influenced by these confounding signals. In order to address this, I tested the N170 component and the late positive potential and found no difference across the image classes. Taken together, there does not appear to be any differences in the motivational aspects, accuracy, or brain mechanisms between the image classes used. This potential issue can be rectified in a future study that would utilize a single class of emotional stimuli across all participants. However, this may cause additional motivational issues if the participant, as I saw in the current study, is not motivated to learn from the image reward.

In the current study, I investigated how differing reward types (points and pictures) can influence learning. I found important interactions between how well a person performs on a reinforcement learning paradigm and their motivation to receive rewards. These findings open up the possibility for new testable hypotheses relating to hedonic aspects of reward processing for intrinsic reward states. Future studies may aim to probe these underlying, understudied aspects of reward processing in order to gain a better understanding of aberrant reward processing conditions.

.

Chapter 4

CONCLUSION

Recent scientific endeavors have revealed an interactional relationship between emotion and reward processing (Murray, 2007; Pessoa, 2009; Phillips, Drevets, Rauch, & Lane, 2003). Indeed, the hedonic appraisal of some types of feedback can either inhibit (Foti & Hajcak, 2009) or boost (Brown & Cavanagh, 2018) a reward signal. However, EEG investigations of these complex reward-emotion signals have remained scarce. The goal of this dissertation was to begin addressing these shortcomings and to hopefully lay the foundation for numerous hypothesis-driven approaches to be tested. This dissertation accomplished these aims by addressing a major overlooked phenomenon in reward processing, providing approaches to address methodological issues, and lastly, providing a novel test of these intrinsic motivation signals.

The first study was designed to address a methodological issue revealed in a previous report (Brown & Cavanagh, 2018). In this past study, researchers aimed to investigate the RewP evoked for pleasant images. However, these signals appeared blocked or hidden beneath another ERP component, the N2. I hypothesized that this effect was a product of a signal of novelty that was evoked for the pleasant images. In the first experiment, participants were presented rewarding feedback paired with novel shapes. It was revealed that the inclusion of novel stimuli occluded the RewP with an enhanced N2. This finding suggests a major limitation for future studies aimed to investigate ecologically-valid types of feedback with EEG, whether that is complex images, food (Harmon-Jones, Gable, & Price, 2013), painful shocks (Talmi, Dayan, Kiebel, Frith, & Dolan, 2009), or genital stimulation (Prause, Siegle, Deblieck, Wu, & Iacoboni, 2016). Although this issue greatly limits the understanding of these reward processes, there are additional methodological strategies that may be utilized to successfully rectify this complication.

In the first experiment it was revealed that employing spectral decomposition was successful in separating the underlying frequencies of the novelty-contaminated RewP. Here, frequency bands relating to reward processing (delta) could be separated from the signal of novelty (theta). This effect was then replicated in the follow-up experiment. Additionally, a third experiment leveraged principles of RPE by presenting participants with a cue, which predicted the upcoming point or picture feedback. Here, it was revealed that the RewP and delta band activity migrated to the predictive cue. This is an important feature of this signal, as it appears to mimic the propagation of phasic midbrain dopaminergic projections (Schultz et al., 1997). A recent report revealed that people with Parkinson's disease display blunted RewP (Brown, Pirio-Richardson, & Cavanagh, in review). However, the blunted RewP normalized as the disease progressed, suggesting that this signal may be an indicator of early disease severity. A future study may aim to employ this experimental paradigm in order to investigate if Parkinsonian medication influences the migration of the RewP to cues predicting rewards in this clinical population. Due to the RewP's relationship to dopaminergic tone, and the ease to which it is elicited, future investigations could uncover important clinical applications for this signal as a candidate biomarker for early stage Parkinson's disease.

In a second series of studies, I used emotional imagery in order to manipulate learning rates in a RL task paradigm. The major aim of these studies was to investigate if intrinsically motivating stimuli (self-chosen pleasant images) would boost learning rates in the participants. Future application of this paradigm could investigate how clinical populations (e.g., individuals with substance use disorder or depression) learn from situational-specific rewards. However, the results revealed that female participants did not learn from rewarding images above chance unless they were willing to pay actual money to see them. In this study, female participants' accuracy to learn the optimal response to cues predicting pleasant picture rewards significantly correlated with how much money (or points) they were willing to trade in to see these positive pictures instead of negative pictures. This effect was then replicated in a second study where female participants were told the types of rewards they would expect to see. In the first RL experiment, participants were not informed of the picture content they would view, the theory here being that a participant would intrinsically place value on the image (since it was chosen in the preceding task), and that would motivate them to learn the optimal response. By informing the participants of the type of pictures they would see in the second RL experiment, I theoretically limited the degree to which intrinsic motivation would influence learning rate, thus making the self-selected pleasant picture "conditioned stimuli" (similar to the point rewards). Surprisingly, even after employing this manipulation, female participants did not learn the optimal responses for these types of reward. Again, this experiment revealed a strong relationship between a participant's accuracy scores and their willingness to trade actual money to see pleasant images. This finding reveals an important aspect of reward learning: that an agent must be motivated by the reward in order to learn to receive it. A future study could further unpack this effect by using positive imagery as rewards that have differing degrees of reward salience to an individual.

The major aim of this series of studies was to investigate the modulation of the RewP to differing types of rewarding feedback. Although employing complex emotional images as trial feedback in Experiments 4 and 5 resulted in a novel motivation-dependent learning effect, I was not able to conclusively suggest this idiosyncratic reward type had any modulatory effects on reward related aspects of the RewP. This limitation is in part due to a theorized novelty enhanced N2 component that occluded the RewP generation. Although in the current series of studies there was no evidence for this modulation, future investigations employing the techniques presented in Experiment 3 (presentation of a predictive cue preceding reward feedback) may allow this effect to come through.

Future EEG investigations may move toward employing complex rewarding stimuli in order to assess how hedonic appraisal, or liking, interacts with the RewP. Indeed, a recent report proposed that the blunted reward signals seen in major depressive disorder may not relate to dysfunctional reward processes per se, but instead could be a product of aberrant hedonic appraisal (Cavanagh et al., 2019). This suggests that some sort of emotion-based dynamics is contained in the RewP. Indeed, the progression of RPE signals in RL tasks faces a type of tug-of-war between expectation and hedonic appraisal, whereby as an agent learns how to optimally operate in an environment, the RPE signal is reduced because the rewarding outcomes are less surprising. However, recent studies have shown that the RewP can be enhanced through positive emotional modulation (Brown & Cavanagh, 2018; Threadgill & Gable, 2016), suggesting that this signal may also be manipulated by the magnitude of reward liking. A future study may aim to disentangle these competing systems. Furthermore, the dissemination of prediction error from hedonic appraisal in the RewP may offer important insights for substance use disorder investigations. This line of work may coincide with the *Incentive-Sensitization Theory* of addiction (Robinson & Berridge, 2008) that posits that neural systems of wanting and liking change throughout addiction. A future investigation may opt to use situational-specific feedback cues (e.g., pictures related to drug craving) in order to investigate how hedonic aspects of this signal change throughout addiction. Due to its sensitivity to reward expectation and emotional modulation, the RewP may make a candidate biological signal worth investigating for this and many other clinical populations.

References

Ait Oumeziane, B., & Foti, D. (2016). Reward-related neural dysfunction across depression and impulsivity: A dimensional approach. *Psychophysiology*, *53*(8), 1174–1184. http://doi.org/10.1111/psyp.12672

Angus, D. J., Kemkes, K., Schutter, D. J. L. G., & Harmon-Jones, E. (2015). Anger is associated with reward-related electrocortical activity: Evidence from the reward positivity. *Psychophysiology*, *52*(10), 1271–1280. http://doi.org/10.1111/psyp.12460

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. http://doi.org/10.1016/j.biopsycho.2011.01.010

Baker, T. E., Stockwell, T., Barnes, G., Haesevoets, R., & Holroyd, C. B. (2016). Reward Sensitivity of ACC as an Intermediate Phenotype between DRD4-521T and Substance Misuse. *Journal of Cognitive Neuroscience*, *28*(3), 460–471. http://doi.org/10.1162/jocn

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or Surprise? *Frontiers in Psychology*, *4*(DEC), 1–15. http://doi.org/10.3389/fpsyg.2013.00907

Beck, S. M., Locke, H. S., Savine, A. C., Jimura, K., & Braver, T. S. (2010). Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PLoS ONE*, *5*(2). http://doi.org/10.1371/journal.pone.0009251

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2014). A Single-Trial Estimation of the Feedback-Related Negativity and Its Relation to BOLD Responses in a Time-Estimation Task. *Journal of Neuroscience*, *34*(8), 3005–3012. http://doi.org/10.1523/JNEUROSCI.3684-13.2014

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

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

Benke, T., Bosch, S., & Andree, B. (1998). A Study of Emotional Processing in Parkinson' s Disease. *Brain and Cognition*, *52*(38), 36–52.

Bernat, E. M., Nelson, L. D., & Baskin-Sommers, A. R. (2016). Time-Frequency Theta and Delta Measures Index Separable Components of Feedback Processing in a Gambling Task. *Psychophysiology*, *52*(5), 626–637. http://doi.org/10.1111/psyp.12390

Bernat, E. M., Nelson, L. D., Holroyd, C. B., Gehring, W. J., & Patrick, C. J. (2008). Separating cognitive processes with principal components analysis of EEG time-frequency distributions. In *Advanced Signal Processing Algorithms, Architectures, and Implementations XVIII* (Vol. 7074, p. 70740S). International Society for Optics and Photonics.

Berridge, K. C. (1996). Food reward: brain substrates of wanting and liking. *Neuroscience and Biobehavioral Reviews*, *20*(1), 1–25. http://doi.org/10.1016/0149-7634(95)00033-B

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*(3), 309–369. http://doi.org/10.1016/S0165-0173(98)00019-8

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*, 1–13. http://doi.org/10.1186/1744-9081-3-7

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and Motivation II: Sex Differences in Picture Processing. *Emotion*, *1*(3), 300–319. http://doi.org/10.1037/1528-3542.1.3.300

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. http://doi.org/10.1163/156856897X00357

Bress, J. N., Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2013). Blunted neural response to rewards prospectively predicts depression in adolescent girls. *Psychophysiology*, *50*(1), 74–81. http://doi.org/10.1111/j.1469-8986.2012.01485.x

Brown, D. R., & Cavanagh, J. F. (2017). The sound and the fury: Late positive potential is sensitive to sound affect. *Psychophysiology*, *54*(12), 1812–1825. http://doi.org/10.1111/psyp.12959

Brown, D. R., & Cavanagh, J. F. (2018). Rewarding images do not invoke the reward positivity: They inflate it. *International Journal of Psychophysiology*, *132*, 226–235. http://doi.org/10.1016/j.ijpsycho.2018.02.012

Bunzeck, N., Dayan, P., Dolan, R. J., & Duzel, E. (2010). A common mechanism for adaptive scaling of reward and novelty. *Human Brain Mapping*, *31*(9), 1380–1394. http://doi.org/10.1002/hbm.20939

Bunzeck, N., Doeller, C. F., Dolan, R. J., & Duzel, E. (2012). Contextual interaction between novelty and reward processing within the mesolimbic system. *Human Brain Mapping*, *33*(6), 1309–1324. http://doi.org/10.1002/hbm.21288

Bunzeck, N., & Düzel, E. (2006). Absolute Coding of Stimulus Novelty in the Human

Substantia Nigra/VTA. *Neuron*, *51*(3), 369–379. http://doi.org/10.1016/j.neuron.2006.06.021

Calvo, M. G., & Avero, P. (2009). Reaction time normative data for the IAPS as a function of display time, gender, and picture content. *Behavior Research Methods*, *41*(1), 184–191. http://doi.org/10.3758/BRM.41.1.184

Caplin, A., & Dean, M. (2008). Axiomatic methods, dopamine and reward prediction error. *Current Opinion in Neurobiology*, *18*(2), 197–202. http://doi.org/10.1016/j.conb.2008.07.007

Carlson, J. M., Foti, D., Harmon-Jones, E., & Proudfit, G. H. (2015). Midbrain volume predicts fMRI and ERP measures of reward reactivity. *Brain Structure and Function*, *220*(3), 1861–1866. http://doi.org/10.1007/s00429-014-0725-9

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

Carvalhaes, C., & De Barros, J. A. (2015). The surface Laplacian technique in EEG: Theory and methods. *International Journal of Psychophysiology*, *97*(3), 174–188. http://doi.org/10.1016/j.ijpsycho.2015.04.023

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage*, *110*, 205–216. http://doi.org/10.1016/j.neuroimage.2015.02.007

Cavanagh, J. F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2019). Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed EEG. *Computational*

Psychiatry, *3*, 1–17. http://doi.org/10.1162/CPSY

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Synchrony Reveals Cognitive Control Dynamics. *Journal of Neuroscience*, *29*(1), 98–105. http://doi.org/10.1523/JNEUROSCI.4137-08.2009.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. http://doi.org/10.1016/j.tics.2014.04.012

Cockburn, J., & Holroyd, C. B. (2018). Feedback information and the reward positivity. *International Journal of Psychophysiology*, *132*(November 2017), 243–251. http://doi.org/10.1016/j.ijpsycho.2017.11.017

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

Cohen, M. X., & Ranganath, C. (2007). Reinforcement Learning Signals Predict Future Decisions. *Journal of Neuroscience*, *27*(2), 371–378. http://doi.org/10.1523/JNEUROSCI.4421-06.2007

David, S. P., Munafò, M. R., Johansen-berg, H., Smith, S. M., Rogers, R. D., Matthews, P. M., & Walton, R. T. (2005). Ventral Striatum / Nucleus Accumbens Activation to Smoking-Related Pictorial Cues in Smokers and Nonsmokers : A Functional Magnetic Resonance Imaging Study. *Addiction Biology*, *17*(4), 807–816. http://doi.org/10.1016/j.biopsych.2005.04.028

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

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

http://doi.org/10.1196/annals.1390.002

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *13*, 9–21. http://doi.org/http://dx.doi.org/10.1016/j.jneumeth.2003.10.009

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. *NeuroReport*, *13*(4), 427–431. http://doi.org/10.1097/00001756-200203250-00013

Ferrari, V., Bradley, M. M., Codispoti, M., & Lang, P. J. (2010). Detecting Novelty and Significance. *Journal of Cognitive Neuroscience*, *22*(2), 404–411. http://doi.org/10.1162/jocn.2009.21244.Detecting

Finlayson, G., King, N., & Blundell, J. E. (2007). Is it possible to dissociate "liking" and "wanting" for foods in humans? A novel experimental procedure. *Physiology and Behavior*, *90*(1), 36–42. http://doi.org/10.1016/j.physbeh.2006.08.020

Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biological Psychology*, *81*(1), 1–8. http://doi.org/10.1016/j.biopsycho.2008.12.004

Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2011). Abnormal neural sensitivity to monetary gains versus losses among adolescents at risk for depression. *Journal of Abnormal Child Psychology*, *39*(7), 913–924. http://doi.org/10.1007/s10802-011-9503-9

Fries, P. (2009). Neuronal Gamma-Band Synchronization as a Fundamental Process in Cortical Computation. *Annual Review of Neuroscience*, *32*(1), 209–224. http://doi.org/10.1146/annurev.neuro.051508.135603

Gable, P. A., & Harmon-Jones, E. (2010). Late Positive Potential to Appetitive Stimuli and Local Attentional Bias. *Emotion*, *10*(3), 441–446. http://doi.org/10.1037/a0018425

Gehring, W., & Willoughby, A. (2002). THe medial frontal cortex and rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282.

George, M. S., Anton, R. F., Bloomer, C., Teneback, C., Drobes, D. J., Lorberbaum, J. P., … Vincent, D. J. (2001). Activation of prefrontal cortex and anterior thalamus in alcoholic subjects on exposure to alcohol-specific cues. *Archives of General Psychiatry*, *58*(4), 345–352. Retrieved from http://0-ovidsp.ovid.com.lib.exeter.ac.uk/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=32299531

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: A deficit in the representation of value. *Schizophrenia Bulletin*, *34*(5), 835–847. http://doi.org/10.1093/schbul/sbn068

Hajcak, G., Macnamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*(2), 129–155. http://doi.org/10.1080/87565640903526504

Hajcak, G., & Olvet, D. M. (2008). The Persistence of Attention to Emotion : Brain Potentials During and After Picture Presentation. *Emotion*, *8*(2), 250–255. http://doi.org/10.1037/1528-3542.8.2.250

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology*, *14*(2), 233–238. http://doi.org/10.1016/j.conb.2004.03.010

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2013). Does Negative Affect Always Narrow and Positive Affect Always Broaden the Mind? Considering the Influence of Motivational Intensity on Cognitive Scope. *Current Directions in Psychological*

Science, *22*(4), 301–307. http://doi.org/10.1177/0963721413481353

Heydari, S., & Holroyd, C. B. (2016). Reward positivity: Reward prediction error or salience prediction error? *Psychophysiology*, *53*(8), 1185–1192. http://doi.org/10.1111/psyp.12673

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

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by predictive cues. *NeuroReport*, *22*(5), 249–252. http://doi.org/10.1097/WNR.0b013e328345441d

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. http://doi.org/10.1111/j.1469-8986.2008.00668.x

Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience and Biobehavioral Reviews*, *71*, 418–443. http://doi.org/10.1016/j.neubiorev.2016.09.021

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., … Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am Psychiatric Assoc.

Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, *15*(4–6), 549–559. http://doi.org/10.1016/S0893-6080(02)00048-5

Kemp, A. H., Silberstein, R. B., Armstrong, S. M., & Nathan, P. J. (2004). Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *NeuroImage*, *21*(2), 632–646. http://doi.org/10.1016/j.neuroimage.2003.09.055

Krigolson, O. E. (2017). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, (November), 1–9. http://doi.org/10.1016/j.ijpsycho.2017.11.007

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

Kujawaa, A., Proudfit, G. H., Laptook, R., & Klein, D. N. (2016). Early Parenting Moderates the Association between Parental Depression and Neural Reactivity to Rewards and Losses in Offspring Autumn, *35*(14), 1252–1260. http://doi.org/10.1177/0333102415576222.Is

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., & Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, *131*(8), 2084–2093. http://doi.org/10.1093/brain/awn136

Lang, P. ., Bradley, M. ., & Cuthbert, B. . (1997). International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *NIMH Center for the Study of Emotion and Attention*. http://doi.org/10.1027/0269-8803/a000147

Levinson, A. R., Speed, B. C., Nelson, B., Bress, J. N., & Hajcak, G. (2017). Authoritarian parenting predicts reduced electrocortical response to observed adolescent offspring rewards. *Social Cognitive and Affective Neuroscience*, *12*(3), 363–371. http://doi.org/10.1093/scan/nsw130

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., … Mather, M. (2012). Gender differences in reward-related decision processing under stress. *Social Cognitive and Affective Neuroscience*, *7*(4), 476–484. http://doi.org/10.1093/scan/nsr026

Lukie, C. N., Montazer-Hojat, S., & Holroyd, C. B. (2014). Developmental changes in the reward positivity: An electrophysiological trajectory of reward processing. *Developmental Cognitive Neuroscience*, *9*, 191–199. http://doi.org/10.1016/j.dcn.2014.04.003

Luque, D., López, F. J., & Marco-pallares, J. (2012). Feedback-related Brain Potential Activity Complies with Basic Assumptions of Associative Learning Theory. *Journal of Cognitive Neuroscience*, *24*(4), 794–808. http://doi.org/10.1162/jocn_a_00145

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*, *8*, 145–151. http://doi.org/10.1109/ICOSP.2002.1180091

Mcclure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, M., & Montague, P. R. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks, *44*(1), 379–387. http://doi.org/10.1016/j.neuron.2004.09.019

Meadows, C. C., Gable, P. A., Lohse, K. R., & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology*, *118*, 154–160. http://doi.org/10.1016/j.biopsycho.2016.06.002

Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1696), 2895–2904. http://doi.org/10.1098/rspb.2010.0303

Morris, S. E., Heerey, E. A., Gold, J. M., & Holroyd, C. B. (2008). Learning-related changes in brain activity following errors and performance feedback in schizophrenia. *Schizophrenia Research*, *99*(1–3), 274–285. http://doi.org/10.1016/j.schres.2007.08.027

Murray, E. A. (2007). The amygdala, reward and emotion. *Trends in Cognitive Sciences*, *11*(11), 489–497.

National Institute of Mental Health, 2015. Research Domain Criteria (RdoC) (Retrieved from) http://www.nimh.nih.gov/research-priorities/rdoc/index. shtml.

Nelson, B. D., Perlman, G., Klein, D. N., Kotov, R., & Hajcak, G. (2016). Blunted neural response to rewards as a prospective predictor of the development of depression in adolescent girls. *American Journal of Psychiatry*, *173*(12), 1223–1230. http://doi.org/10.1176/appi.ajp.2016.15121524

Nestor, P. G., Choate, V., Niznikiewicz, M., Levitt, J. J., Shenton, M. E., & McCarley, R. W. (2014). Neuropsychology of reward learning and negative symptoms in schizophrenia. *Schizophrenia Research*, *159*(2–3), 506–508. http://doi.org/10.1016/j.schres.2014.08.028

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162.

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

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776. http://doi.org/10.1016/j.conb.2004.10.016

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviation and the Representation of Action–Outcome Associations. *Journal of Cognitive Neuroscience*, *19*(12), 1994–2004. http://doi.org/10.1162/jocn.2007.19.12.1994.

Pasquereau, B., Tremblay, L., & Turner, R. S. (2019). Local Field Potentials Reflect Dopaminergic and Non-Dopaminergic Activities within the Primate Midbrain. *Neuroscience*, *399*, 167–183. http://doi.org/10.1016/j.neuroscience.2018.12.016

Pavlov, P. I. (2010). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, *17*(3), 136–141. http://doi.org/10.5214/ans.0972-7531.1017309

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, *13*(4), 160–166.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: the neural basis of normal emotion perception. *Biological Psychiatry*, *54*(5), 504–514.

Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: The medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, *18*(7), 1112–1119. http://doi.org/10.1162/jocn.2006.18.7.1112

Prause, N., Siegle, G. J., Deblieck, C., Wu, A., & Iacoboni, M. (2016). EEG to primary rewards: Predictive utility and malleability by brain stimulation. *PLoS ONE*, *11*(11). http://doi.org/10.1371/journal.pone.0165646

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

http://doi.org/10.1111/psyp.12370

Proudfit, G. H., Bress, J. N., Foti, D., Kujawa, A., & Klein, D. N. (2015). Depression and Event-related Potentials: Emotional disengagement and reward insensitivity. *Current Opinion in Psychology*, *4*, 110–113. http://doi.org/10.1111/ajt.13087.TIM-1

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

Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: Some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1507), 3137–3146. http://doi.org/10.1098/rstb.2008.0093

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects. *NeuroReport*, *11*(1), 69–72. http://doi.org/10.1097/00001756-200001170-00014

Sabatinelli, D., Bradley, M. M., Lang, P. J., Costa, V. D., & Versace, F. (2007). Pleasure Rather Than Salience Activates Human Nucleus Accumbens and Medial Prefrontal Cortex. *Journal of Neurophysiology*, *98*(3), 1374–1379. http://doi.org/10.1152/jn.00230.2007

Sabatino, A., Richey, J. A., Bodfish, J. W., Dichter, G. S., & Rittenberg, A. M. (2011). Reward Circuitry Function in Autism During Face Anticipation and Outcomes. *Journal of Autism and Developmental Disorders*, *42*(2), 147–160. http://doi.org/10.1007/s10803-011-1221-1

Sambrook, T. D., & Goslin, J. (2015). A Neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. http://doi.org/10.1037/bul0000006

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping*, *30*(7), 1963–1976. http://doi.org/10.1002/hbm.20642

Schacht, J., Anton, R., & Myrick, H. (2014). Functional neuroimaging studies of alcohol cue reactivity: A quantitative meta-analysis and systematic review. *Addiction Biology*, *18*(1), 121–133. http://doi.org/10.1111/j.1369-1600.2012.00464.x.Functional

Schneider, F., Gur, R. C., Gur, R. E., & Shtasel, D. L. (1995). Emotional processing in schizophrenia: Neurobehavioral probes in relation to psychopathology. *Schizophrenia Research*, *17*(1), 67–75. http://doi.org/10.1016/0920-9964(95)00031-G

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

Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*, *95*(3), 853–951. http://doi.org/10.1152/physrev.00023.2014

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

Seeman, P., & Kapur, S. (2000). Schizophrenia : More dopamine , more D 2 receptors. *Proceedings of the National Academy of Sciences*, *97*(14), 7673–7675.

Smith, B. W., Mitchell, D. G. V., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J. R.,

& Ernst, M. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *NeuroImage*, *44*(2), 600–609. http://doi.org/10.1016/j.neuroimage.2008.08.016

Smith, K. S., & Berridge, K. C. (2005). The Ventral Pallidum and Hedonic Reward: Neurochemical Maps of Sucrose "Liking" and Food Intake. *Journal of Neuroscience*, *25*(38), 8637–8649. http://doi.org/10.1523/jneurosci.1902-05.2005

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The Feedback-Related Negativity Signals Salience Prediction Errors, Not Reward Prediction Errors. *Journal of Neuroscience*, *33*(19), 8264–8269. http://doi.org/10.1523/JNEUROSCI.5695-12.2013

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How Humans Integrate the Prospects of Pain and Reward during Choice. *Journal of Neuroscience*, *29*(46), 14617–14626. http://doi.org/10.1523/JNEUROSCI.2026-09.2009

Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., & Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *NeuroImage*, *59*(1), 635–645. http://doi.org/10.1016/j.neuroimage.2011.06.051

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review*, *2*(4), 1–107. http://doi.org/10.1097/00005053-190001000-00013

Threadgill, A. H., & Gable, P. A. (2016). Approach-motivated pregoal states enhance the reward positivity. *Psychophysiology*, *00*, n/a-n/a. http://doi.org/10.1111/psyp.12611

Tremblay, L. K., Naranjo, C. A., Graham, S. J., Herrmann, N., Mayberg, H. S., Hevenor, S., & Busto, U. E. (2005). Functional Neuroanatomical Substrates of Altered Reward Processing in Major Depressive Disorder Revealed by a Dopaminergic Probe. *Archives of General Psychiatry*, *62*(11), 1228–1236.

Verdejo-Garcia, A., Perez-Garcia, M., & Bechara, A. (2005). Emotion, Decision-Making and Substance Dependence: A Somatic-Marker Model of Addiction. *Current Neuropharmacology*, *4*(1), 17–31. http://doi.org/10.2174/157015906775203057

Vezoli, J., Procyk, E., Cell, S., & Lyon, D. (2009). Frontal Feedback-Related Potentials in Nonhuman Primates : Modulation during Learning and under Haloperidol. *Journal of Neuroscience*, *29*(50), 15675–15683. http://doi.org/10.1523/JNEUROSCI.4943-09.2009

Weinberg, A., & Hajcak, G. (2010). Beyond Good and Evil: The Time-Course of Neural Activity Elicited by Specific Picture Content. *Emotion*, *10*(6), 767–782. http://doi.org/10.1037/a0020242

Wrase, J., Klein, S., Gruesser, S. M., Hermann, D., Flor, H., Mann, K., … Heinz, A. (2003). Gender differences in the processing of standardized emotional visual stimuli in humans: a functional magnetic resonance imaging study. *Neuroscience Letters*, *348*(1), 41–45. http://doi.org/10.1016/s0304-3940(03)00565-2

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

Xue, G., Lu, Z., Levin, I. P., Weller, J. A., Li, X., & Bechara, A. (2008). Functional Dissociations of Risk and Reward Processing in the Medial Prefrontal Cortex. *Cerebral Cortex*, *19*(5), 1019–1027. http://doi.org/10.1093/cercor/bhn147

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and

Valence in the Human Brain. *Journal of Neuroscience*, *24*(28), 6258–6264. http://doi.org/10.1523/JNEUROSCI.4537-03.2004

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., Mcgill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human Substantia Nigra Neurons Encode Unexpected Financial Rewards. *Science*, *323*(5920).