ABSTRACT

Title of Thesis:	EFFECTS OF ACTION-OUTCOME AGENCY ON FEEDBACK PROCESSING
	Anne Tootell, Master of Science, 2017
Thesis Directed By:	Dr. Edward Bernat, Associate Professor Department of Psychology

The current study investigated the effect of action-outcome agency, or one's ability to guide behavior during reinforcement learning, on reward and loss processing in a gambling task. Thirty undergraduates (13 females; M age = 19.57, SD = 2.18) completed two computer gambling tasks, one designed to exhibit high levels of action-outcome agency and one with low, while attached to a 128-channel EEG system. Time-frequency event-related potential (TF- ERP) analysis was conducted on the acquired EEG data. ERP components associated with reward and loss processing were significantly dampened in the low action-outcome agency task relative to the high action-outcome agency task. Interestingly, TF-ERP analysis demonstrated a significant effect of action-outcome agency on gain-loss differences in theta but not delta frequencies, suggesting a more central role of loss processing in guidance of goal-directed behavior. These results challenge components of the well-established predicted response-outcome (PRO) model of reinforcement learning.

EFFECTS OF ACTION-OUTCOME AGENCY ON
FEEDBACK PROCESSING

by

Anne Tootell

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2017

Advisory Committee:
Professor Edward Bernat, Chair
Professor Andres De Los Reyes
Professor Michael Dougherty

Acknowledgements

I would like to deeply thank my mentor Dr. Edward Bernat for his expertise and support both on this project and my continued research training. I would also like to thank my Master's Thesis committee, Dr. De Los Reyes and Dr. Dougherty, whom provided integral guidance for the development and successful completion of this project. Finally, I am thankful for my research colleagues Adreanna Massey, Jessica Ellis, and Matthew Bachman for their support throughout the process.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The current study investigated the effect action-outcome agency had on theta and delta processing in the FN component. Action-outcome agency, or control over the ability to enact an action that elicits a desired outcome, is a core process underlying task learning and goal-directed behavior modulation. Within ERP literature, the feedback negativity (FN) component has been conceptualized as a marker of this adaptive learning process due to the prevalence of this component during processing of task feedback, i.e. errors, gains, losses. Conflicting findings have tied the FN component to both positively and negatively-valenced stimuli, suggesting the relative importance of positive and negative stimuli processing to the learning process. A recent integrative model hypothesized that the FN is modulated by surprising outcomes, regardless of valence. This predicted response-outcome (PRO) model conceptualizes activity in the anterior cingulate cortex (ACC) as a constantly updating outcome predictor, which perceives any surprising stimuli that deviates from the expected outcome as particularly salient (Alexander & Brown, 2011). Within the context of this model, we hypothesized that both negatively and positively-valenced feedback would modulate the FN component when the task allowed participants to have control over the ability to garner the most desirable outcomes. Methodologically, previous attempts to investigate the effect of action-outcome agency on adaptive learning have focused on conventional time-domain measures of FN activity. But time-domain approaches have not been effective for disentangling feedback processing related to positively and negatively-valenced stimuli. Newer time-frequency ERP analysis approaches, on the other hand, have distinctly linked

theta and delta frequency activity to processing of negative and positive stimuli, respectively, allowing the measures in the current study to isolate negative and positive feedback processes that could be modulated differently with and without action-outcome agency.

Action-outcome agency

There is a growing body of ERP research that focuses on the neural mechanisms underlying adaptive behavior. In laboratory experiments, learning adaptive behavior has been effectively operationalized as task optimization through reinforcement learning. A necessary element of reinforcement learning is the ability to create an association between actions and outcomes (Holroyd & Coles, 2002, 2008). Action-outcome agency refers to a person's capacity to control the association between action and outcome in order to learn the most adaptive action-outcome pairing and optimize performance. To achieve optimal performance on a specific task, task stimuli must be evaluated for relative motivational value, valence, and expectancy. Optimal performance can represent everything from an elimination of errors in a learning task to a maximization of monetary gain in a generic gambling task. In either case an optimal performance can be achieved through the continuous evaluation of action-outcome pairs. The participant must learn which action yields the most desired outcome; without this knowledge optimization cannot be achieved. A common assumption underlying learning optimal goal-directed behavior states that for learning to take place the participant must have agency over the action-outcome pairing (Yeung et al., 2005). Behavior can only be perceived as adaptive if one believes their behavior produces a desired outcome. Thus, and central to the current

project, optimization of a task is thought to be fundamentally blunted when one's actions are not tied to the resultant outcomes.

ERP indices of reinforcement learning processes: ERN, FN, and RewP

Many EEG studies have investigated the ERP components that are elicited when participants are provided with feedback during reinforcement learning tasks. Reinforcement learning refers to the process utilized within an uncertain environment to avoid negative outcomes and obtain positive ones through trial and error. On each trial of the undefined task, one analyzes and internalizes the difference between the expected outcome and the presented outcome to influence the action on the next trial (Sutton & Barto, 1998). Early work isolated a negative deflection in the ERP waveform that was elicited approximately 80-100ms after an incorrect response. This error-related negativity (ERN) was elicited by speed-related tasks after a participant's response to a trial was incorrect, which deviated from the participant's expectation of delivering a correct response (Coles et al., 1995; Falkenstein et al., 1995; Carter et al., 1998). Additionally, Gehring and Fencsik (2001) demonstrated that errors that more closely resembled the correct response elicited comparatively smaller ERNs than errors that were more dissimilar to the correct response. Further work with reinforcement learning tasks focused on a similar negative deflection that was associated with processing of feedback that denoted the participant's response was not optimal (Gehring and Willoghby, 2002; Milter et al., 1997). This negative deflection linked with processing of positive or negative feedback was labeled the feedback-ERN or feedback negativity (FN). The FN component is observed as a negative deflection at medial-frontal electrodes, which occurs in the range of 180-350

ms post-stimulus (Miltner et al., 1997; Gehring and Willoughby, 2002). Source-localization EEG studies and functional magnetic imaging (fMRI) work have provided evidence for the localization of both the ERN and FN components to the ACC, (Dehaene et al., 1994; Gehring and Willoughby, 2002) which has been widely implicated in reinforcement learning processes (Crino et al., 1993; Ito et al., 2003; Monchi et al., 2001; Bush et al., 2002; Delgado et al., 2003).

While the ERN and FN play functionally similar roles in the monitoring and processing of errors, evidence suggests that the FN is also sensitive to reward-related processing. Broadly, previous research has posited that the FN demonstrated differential activity to negatively and positively valenced outcomes, with the FN being relatively diminished or non-existent to positive outcomes (Miltner et al., 1997; Gehring and Willoughby, 2002). Incorporating theories of reinforcement learning in addition to evaluative processing, several studies have found that the FN is elicited by unexpected negative but not unexpected positive outcomes, providing support for differential learning processes for feedback with opposite valence (Holroyd & Krigolson, 2007; Miltner et al., 1997). Others posit that modulation of the FN by unexpected positive or negative outcomes is dependent on the parameters of the task and the relative purpose of a participant's goal-directed behavior (Holroyd et al., 2002; Holroyd et al., 2004; Nieuwenhuis et al., 2004). Thus, there are reasons to think that both positively and negatively-valenced feedback may be integral to the reinforcement learning process.

Recently, a focus on the role of rewards in reinforcement learning has begun to emerge. Instead of further emphasis on the modulation of the FN component by

unexpected negatively-valenced feedback, researchers have investigated the lack of negative deflection present on gain or correct trials. These studies argue that the core processing during this time window (around 250 ms post-stimulus) is characterized by a positive deflection in response to rewarding stimuli: the reward positivity (RewP) ERP component (Holroyd et al., 2008; Potts et al., 2006; Holroyd et al. 2011; Proudfit, 2012). Holroyd et al. (2011) interprets the role of the RewP in a learning context to be similar to that of the FN -- to evaluate whether the desired outcome was achieved. He describes this as a reward prediction error (RPE) signal, which is enhanced for more unexpected positive outcomes (Holroyd et al., 2011; Holroyd & Coles, 2002). From this perspective, it is not the error stimuli that are primarily guiding task learning behavior, but the processing of unexpected rewards. Both FN and RewP components are understood to index processing of unexpected outcomes in line with a reinforcement learning conceptualization. This work now provides the opportunity to assess the relative roles of positive or negative feedback valence is in question – e.g. is one's trial and error learning of appropriate actions influenced more by unexpected positively or negatively-valenced outcome feedback?

PRO model

Debate over the exact role of the FN and RewP components within the reinforcement learning process contributed to the development of the predicted response-outcome (PRO) model, which integrates evidence from previous research in the reinforcement-learning field. This model posits that activity in the mPFC and ACC (which previous work has linked to activity in the FN) creates representations of the likelihood of specific outcomes, regardless of valence (Alexander & Brown,

2011). These regions are proposed to house representations of all the possible

outcomes of an action and frequently update the likelihood of each outcome with

each new feedback presentation. Each representation integrates both positively and

negatively valenced information to form a cost-benefit analysis for possible actions,

and the action with the most favorable cost-benefit difference will be the action

chosen by the participant. In order to learn which outcome is the most beneficial

option, there needs to be a mechanism to detect differences in the presented feedback

and the current mental representation of that action-outcome pair. Alexander and

Brown suggest that the functional activity in the ACC is a central mechanism for

detecting discrepancies between predicted and presented outcomes, and signals for

modification of the mental representation. Thus, the PRO model supports the

importance of unexpected negative and positive feedback indexed in the FN and

RewP components.

Time-Frequency analysis

Review of the conflicting findings on the role of the FN and RewP

components poses a key question relevant to feedback learning -- do the core

processes underlying these observed feedback component more directly reflect the

influence of rewarding and positive outcomes, errors and negative outcomes, or some

combination of the two? Methodologically it has been difficult to disentangle the

effects of these processes because they overlap in time. Therefore, it is useful to

separate unique processing within the time-domain through time-frequency analysis.

Past work from our lab and others has demonstrated that the FN component

can be separated into independent processing in theta and delta frequencies (Bernat et

al., 2008; Bernat et al., 2011; Bernat et al., 2015). In a gambling task adapted from Gehring and Willoughby (2002) that is similar to the gambling tasks utilized in this current study, these studies found that theta activity was modulated by negative salient feedback attributes (losses) while delta was modulated by positive primary feedback (gains). Additionally, both theta and delta activity explain unique variance when considered in a regression model predicting the time-domain FN. Thus, for activity within the time-domain window associated with feedback processing, time-frequency analyses can index modulations in separable brain systems engaged differentially by positive and negative feedback, suggesting that theta and delta activity are a better fit as indices for the proposed functional roles of FN and RewP components, respectively. Using this approach, feedback processing associated with the FN and RewP components can then be better isolated and analyzed separately, providing an important approach to assessing processing underlying reinforcement learning.

Efforts to manipulate action-outcome contingencies

Previous work has manipulated action-outcome contingencies and assessed FN and RewP activity, in the time-domain, and this work provides an important starting point for this proposed project. Yeung et al. (2005) hypothesized that if the FN were a result of reinforcement learning processes, and not simply due to stimulus evaluation, then the FN would only be generated when participants had agency or control over the selection of their outcome. To assess this hypothesis, Yeung et al. developed a "Choice" and "No-choice task". The Choice task allowed the participant to choose between two options, while the No-choice task presented only one option

that the participant was forced to choose. Yeung and his colleagues found that both tasks generated a FN, challenging previous conceptualization of the FN as an indicator of action-outcome learning. However, this study did find that the No-choice task elicited a smaller FN relative to the Choice task on loss-gain trial differences. Because of the overlap between FN and RewP components, the time-frequency approaches proposed in the current work may help isolate the relevant processes.

Current Study

In the current study, through utilization of time-frequency approaches, we will assess the effect manipulating participant control over task action-outcome contingences has on loss and gain processing. In order to manipulate control, we designed two gambling tasks, in which the feedback (i.e. monetary gain or loss) was either causally linked to participant response choice or not. Within these tasks, designated "high-control" and "low-control" respectively, participants were able to choose between two options that resulted in different monetary feedback (high-control) or "chose" between two identical options with identical outcomes (low-control). ERPs associated with gain and loss feedback were computed and assessed in theta and delta frequency bands. This involves three central hypotheses:

1. We predicted a diminution in ERP amplitude in both delta and theta in the low-control task, due to anticipated decreased attention and motivation concerning a task with which a participant has no control.

2. We predicted a greater decrease for loss trials (relative to gains) in the low-control condition compared to high-control, due to participant's decreased

8

control over action-outcome contingencies in the low-control condition. We

expected to see this effect in theta because theta processing has been shown to

be more sensitive to loss feedback.

We predicted a greater decrease for gain trials (relative to losses) in the low-

control condition relative to the high-control, due to participant's decreased control

over action-outcome contingencies in the low-control condition. We expected to see

this effect in delta because delta processing has been shown to be more sensitive to

gain feedback.

Chapter 2: Methods

Participants

A total of 35 undergraduate students were recruited from a large state university in the southeast. Four of these subjects were excluded due to excessive artifacts (>50% of trials rejected using methods described below) and one was excluded due to data collection error. A total of 30 participants remained for analysis (13 females; M age = 19.57, SD = 2.18). All participants were 18 years of age or older and were screened for neurological conditions, visual impairments, and/or traumatic brain injuries. Before starting the study, students provided informed consent and were given the option of course credit or monetary compensation ($10/hr) for their participation.

Procedure

Data collection was conducted in a dimly lit, sound-attenuated room. Experimental stimuli were displayed in the center of a 21-inch Dell high-definition CRT color monitor placed 100cm from the seated participant. Stimuli were presented with E-Prime 1.1. Participants provided behavioral responses to the task through the use of a PST Serial Response box (Psychology Software Tools, Inc.).

Subjects performed two tasks that were modified versions of a two-choice gambling task developed by Gehring and Willoughby (2002). In the first task, participants were instructed to choose between two presented squares, each of which contained a monetary amount corresponding to 5 cents or 25 cents. Once the subject chose one of the two squares, each square would turn red or green signaling whether the participant's choice resulted in a win or a loss of the money amount. The winning

color was counterbalanced between subjects; for half of the participants green designated a win and red designated a loss, and the opposite was true for the remaining participants. Participants were told they were given 20 cents at the beginning of the task. Prior to task administration, subjects completed a brief set of practice trials, during which data were not collected.

The second task, much like the first, consisted of two squares containing money amounts corresponding to 5 or 25 cents. Similarly, participants received feedback (i.e. whether they won or loss) when the two squares turned red or green. However, unlike the first task, both squares contained the same monetary amount and lead to the same outcome. For example, if two squares were presented with 5 cents displayed within each square, then the feedback would consist of both of the squares signifying a loss. Thus, whichever square the participant chose would lead to the same monetary gain or monetary loss. As with the first task, participants completed a brief set of practice trials prior to task administration. The first task was referred to as the "high-control task" while the second task was deemed the "low-control task". Participants were not aware of this distinction, but were explicitly told that the two squares in the low-control task "will always display the same point amount and the outcome will be the same on both sides."

Psychophysiological Data Acquisition

Electroencephalographic data were collected using a 128-channel Synamps RT amplifier (Neuroscan, Inc.) and Neuroscan 128-channel Quik-Caps (sintered Ag-Ag/Cl; non-standard layout). Ten channels around the ears were removed due to inadequate scalp connection, leaving 113 channels available for analysis. Bipolar

horizontal channels were placed on the outer canthus of both eyes, and bipolar vertical channels were placed above and below the left eye. Impedances in all electrodes were below 10 kΩ. During recording EEG signals were referenced to the vertex electrode (between Cz and CPz) and rereferenced to the averaged mastoid signals offline. EEG signals were collected through an analog 0.05 to 200 Hz bandpass filter and digitized at 1000 Hz.

Data Preprocessing

Epochs three seconds in length were extracted from the continuous data from 1000ms pre-stimulus to 2000ms post-stimulus with a 150ms pre-stimulus baseline. Data were corrected for ocular artifacts using an algorithm developed by Semilitsch and colleagues (1986), and implemented in the Neuroscan Edit 4.5 software (Neuroscan, Inc.). Data were downsampled to 128 Hz using the Matlab resample function (Mathworks, Inc.), which applies an anti-aliasing filter during resampling. Trial-level artifact rejection was performed in a two-step process. Whole trials were rejected if activity at F3 and F4 exceeded ±100 µV in either the pre- (-1000 to -1ms) or post-stimulus (1 to 2000ms) time windows. Within-trial individual electrodes were rejected if activity exceeded ±100 µV within the same pre- and post-stimulus window. This removed 11% of all trials from analysis. Through visual analysis of the averaged waveforms, 37 electrodes out of 3,051 total electrodes were identified as disconnected during recording and were removed from analysis. After preprocessing, the data were averaged according to feedback type (i.e. wins or losses).

Time-Domain FN Amplitude Identification

Time-domain (TD) measures of evoked amplitude were extracted for the FN component. The FN component was defined as a negative deflection ranging from 180 to 350 milliseconds post stimulus. All time ranges were converted to bins, where each 1000 millisecond epoch consisted of 128 bins.

Time-Frequency Evoked Power

Trial-averaged ERPs were pre-filtered using 3rd order Butterworth filters to isolate theta (3Hz highpass filter, 8Hz lowpass filter) and delta (4 Hz lowpass filter) activity. We have employed this filter approach to isolate theta and delta activity in past studies (Bernat et al., 2011; Nelson et al., 2011; Harper et al., 2014; Bernat et al., 2015). Filter cutoffs were chosen based on visual inspection of the unfiltered grand average time-frequency representation. Filtered signals were transformed to time-frequency (TF) representations using the binomial reduced interference distribution (RID) variant of Cohen's class of transforms, with 32 time bins per second and 2 frequency bins per Hz. The RID transform was appropriate for these analyses because of the transform's ability to better represent lower frequency activity.

Principal component analysis (based on the covariance matrix with Varimax rotation; Bernat et al., 2005) was applied independently to the theta and delta filtered decompositions within a 0-14 Hz frequency window and 0-1000 millisecond post-stimulus time window. This principal component analysis (PCA) as applied to the time-frequency domain is equivalent to its application the frequency or time domain. The covariance data matrix consists of time-frequency points as vectors and subject/electrodes/trial-averaged scores as rows (see Bernat et al., 2005 for a detailed explanation of this methodology).

The grand average TF-PCA decomposition is displayed in Figure 2. One principal component (PC) was extracted independently for the theta and delta decompositions. For the theta decomposition, the PC explained 40.56% of the total variance. PC1 reflected medial frontal theta activity during the FN component (approximately 250-450ms). For the delta decomposition, the PC explained 68.14% of the total variance and reflected centro-parietal activity within an approximate 200-500 ms time window. Nine-electrode clusters depicting the mean PC-weighted TF evoked energy for theta and delta PCs were used for statistical analyses (see Figure 3).

Inter-trial Phase Synchrony

Inter-trial Phase Synchrony (ITPS) was computed for each type of trial (i.e. monetary gains and losses) in both the theta and delta frequency ranges. ITPS measured the consistency of responses on different trials by analyzing the similarity or synchronicity between oscillations in the ERP waveform. More consistent ERP responses may be interpreted as greater cognitive regularity or efficiency during feedback processing.

Inter-channel Phase Synchrony

Similarly, inter-channel phase synchrony (ICPS) measures computed the synchrony between signals present in different EEG channels, which correspond to various brain regions. ICPS was calculated through time-varying, frequency-locked phase synchrony computations based on Cohen's class of time frequency distributions. (Aviyente et al., 2011). Channels associated with brain regions of particular interest in this study were extracted for further analysis (see figure 4).

Specifically, ICPS between medial frontal channels, which correspond with theta activity in the anterior cingulate cortex, and lateral frontal channels have been used to denote engagement of the cognitive control network (Aviyente & Multu, 2011; Aviyente et al., in press). The cognitive control network involves lateral prefrontal cortex, orbitofrontal cortex, and the inferior frontal gyrus and is associated with mental representations of expected outcomes that are used to guide behavior (Miller & Cohen, 2001). Previously, it has been found that engagement of the lateral prefrontal cortex, or cognitive control network, is greater during trials that require more cognitive resources, such as response inhibition during a no-go trial or response conflict during incongruent Flanker trials (Aviyente et al., in press; Moran et al., 2015; Bolanos et al., 2013)

Data Analytic Plan

To evaluate the effect of task type on theta and delta processing, 2X2 repeated measures ANOVAs were computed with task type and feedback type (wins and losses) as within-subject factors. These ANOVAs were computed separately for theta and delta PCs for the three measures included in the analyses: amplitude, ITPS, and ICPS. To further investigate the relationship between task type and feedback processing in theta and delta, paired t-tests for all measures were computed for gain and loss trials separately within the two gambling tasks. Additionally, regression analyses including time-frequency amplitude, ITPS, and ICPS measures were computed to investigate whether any of these time-frequency measures explained unique variance within feedback processing in each task type. Finally, JZS Bayes factor repeated measures ANOVAs were conducted to compare alternative models to

the null and compare main effect models and interaction models in amplitude, ITPS, and ICPS measures.

Design Considerations

Extensive considerations were made when developing the high and low-control tasks. Previously, Yeung and colleagues implemented their choice and no-choice tasks to investigate action-outcome processing. To manipulate action-outcome processing, the choice task allowed the participant to choose one of four stimuli (i.e. four possible buttons) that would either show a monetary gain or loss. In the no-choice task the participants were told the task was similar to roulette. Participants were only able to press one button, which would start the spinning wheel. Then almost 2 seconds later the wheel would stop on one of the four possible options and gain or loss feedback would be presented.

The design of the high-control task was very similar to that of Yeung's choice task. In the high-control task the participant could choose between two stimuli (i.e. two button options) and would be shown the gain or loss feedback for both options. In comparison, the design of the low-control task was significantly different from that of Yeung's no-choice task to make the action-outcome agency manipulation more explicit. Similar to the high-control task, the low-control task contained two stimuli options and feedback would be displayed for the chosen and unchosen option. Participants were explicitly told that the stimuli and outcome for each trial would be the same before they began the task. Unlike Yeung's no-choice task where participants still had control over the timing of the task and when they pressed the button to start the spinning wheel, the low-control trials would result in the same

16

outcome, regardless of when the participant pressed a button. Additionally, Yeung and colleagues informed the participants that the task was like roulette, but did not explicitly say that the timing of their button press had no effect on the outcome. Thus, the low-control task, more so than the no-choice task, disallowed participants to be under any illusion that their choice would affect the resulting outcome.

The two tasks described above were administered within a protocol containing five total gambling tasks. In order to minimize the introduction of additional noise, the five tasks were administered in the same order for each participant. Because of this design decision, the high-control task was always completed before the low-control task. This experimental consideration could lead to habituation effects; the ERPs elicited from the low-control task would be smaller than those elicited by the high-control task simply because the participant was more familiar with the task stimuli. Without knowing whether differences in ERP amplitude between the high and low-control tasks were due to habituation effects or valid differences in the processing of task stimuli, no interpretations concerning feedback processing can be made.

Two elucidate this potential confound, average ERP waveforms from the five tasks were analyzed. Delta activity for gain trials and theta activity for loss trials from the high and low control tasks were compared with the final gambling task administered in the protocol (see Figure 5). The average amplitude for delta on gain trials in the high-control task was 20 microvolts, while the average amplitude for the low-control task was approximately 10 microvolts. However, the average delta amplitude on gain trials in the final gambling task of the protocol was again 20

microvolts. Thus, even though the gambling tasks involved similar stimuli, there was

no evidence of habituation because the final task in the protocol elicited ERP

amplitude similar to that of the first experimental task (i.e. the high-control task).

Chapter 3: Results

Effect of task type on theta and delta amplitude processing

Figures 6 and 8 display the results for theta amplitude gain and loss processing in the high-control and low-control tasks. To assess the effects of task type and feedback type in theta, a repeated measure 2X2 ANOVA was computed. The interaction between task and feedback was significant ($F(1,29)=13.256$, $p=0.001$). Paired t-tests were used to better understand the nature of the significant interaction. During the high-control task, theta amplitude was significantly enhanced for loss trials compared to gain trials (t=3.448, p=0.002). However, during the low-control task there was no significant difference between gain and loss trials (t=1.626, p=0.115).

The results of delta amplitude processing of gain and loss trials in both tasks can be seen in figures 7 and 8. To test the effect of task type and feedback type in delta amplitude processing another repeated measure 2X2 ANOVA was conducted. The interaction between task and feedback was not significant ($F(1,29)=0.158$, $p=0.694$). Main effects for task type and feedback type were significant ($F(1,29)=51.621$, $p<0.001$, $F(1,29)=5.373$, $p=0.028$, respectively). Paired t-tests for both high and low-control tasks displayed enhanced amplitude processing for gain trials compared to loss trials. A test of simple effects between gain and loss trials was significant in the low-control task and the difference was trend level in the high-control task (high-control task: t=1.703, p=0.099; low-control task: t=2.330, p=0.027).

Finally, an overall 2X2X2 ANOVA was performed to confirm that there was a difference in the task by feedback interaction in theta and delta frequencies. The three-way interaction between frequency, task, and feedback was significant ($F(1,29)=5.715, p=0.024$).

Effect of task type on theta and delta ITPS processing

Figures 6 and 8 display the results for theta ITPS gain and loss processing in the high-control and low-control tasks. To assess the effects of task type and feedback type in the theta ITPS measure, a repeated measure 2X2 ANOVA was computed. The interaction between task and feedback was significant ($F(1,29)=7.195, p=0.012$). Similarly, paired t-tests were used to better understand the nature of the significant interaction. During the high-control task, theta ITPS was significantly enhanced for loss trials compared to gain trials (t=3.299, p=0.003). While no significant difference between gain and loss trials was found in the low-control task (t=1.717, p=0.097).

The results of delta ITPS processing of gain and loss trials in both tasks can be seen in figures 7 and 8. To test the effect of task type and feedback type in delta ITPS processing a repeated measure 2X2 ANOVA was conducted. The interaction between task and feedback was not significant ($F(1,29)=0.086, p=0.772$.) Main effects for task type and feedback type were significant ($F(1,29)=66.890, p<0.001, F(1,29)=9.697, p=0.004$, respectively). A test of simple effects between gain and loss trials demonstrated significantly enhanced ITPS for gain trials for both tasks (high-control task: t=2.427, p=0.022; low-control task: t=2.585, p=0.015).

Again, an overall 2X2X2 ANOVA was performed. The three-way interaction between frequency, task, and feedback was significant ($F(1,29)=5.993, p=0.021$).

Effect of task type on theta ICPS processing

Results for ICPS between medial-frontal and lateral prefrontal electrodes in the theta frequency are displayed in figures 6 and 8. A repeated measure 2X2 ANOVA demonstrated a significant interaction between task and feedback type ($F(1,29)=12.64$, $p=0.001$). Paired t-tests computing the difference between gain and loss feedback showed significantly enhanced ICPS on loss trials in the high-control task and not in the low-control task (high-control task: t=4.829, p<0.001; low-control task: t=1.459, p=0.155).

Unique variance explained by the different measures

Model residuals were extracted from regression analyses that included all time-frequency measures of gain/loss feedback differences in the theta frequency (amplitude, ITPS, and ICPS) and delta frequency (amplitude and ITPS). In total 5 regressions were conducted with each of the five possible measures input as the dependent variable of the regression. The extracted residuals of each regression, with the constant removed, were then compared to zero in a one-sample t-test. Residual values significantly different from zero represented some additional unique variance explained by the dependent variable in the regression above and beyond the independent variable measures. For theta, only the ICPS measure was found to uniquely explain differences in gain and loss processing for the high-control task (t=2.308, p=0.028). In the low-control task, no measure in the theta frequency explained variance in gain/loss processing above and beyond the other measures. In both the high and low-control tasks, delta amplitude and delta ITPS did not provide any significantly unique contribution to the gain/loss differences.

Bayesian model comparisons

   All Bayesian analyses consisted of a model comparison approach using Bayes Factors within the repeated measures ANOVA framework (Rouder, Morey, Speckman & Province, 2012). This approach allowed us to investigate whether the data support one particular model above all others. Bayes Factor values greater than 1 represent greater support for the alternative hypothesis, while Bayes Factors less than 1 represent greater support for the null hypothesis. Table 1 displays the relative strengths of Bayes Factors with support for the alternative or null models (Jeffreys, 1961). A JZS Bayes repeated measures ANOVA for theta amplitude using the default prior scales (r=0.707) revealed that the main effects model with task type and feedback type was preferred over the null model with $BF_{10} = 1.46e7$. These data indicate decisive evidence for the significant effects of both task type and feedback type on theta amplitude. However, comparison of the main effects model to the interaction model (i.e. task type X feedback type) provided substantial support for the interaction model as the more compelling model compared to the null ($BF_{21} = 5.49$). For delta amplitude there was again decisive evidence for the main effects model with task and feedback terms over the null ($BF_{10} = 3.88e15$). Interestingly, when comparing the main effects model to the model including the interaction, there was substantial support for the main effects model over the interaction model ($BF_{12} = 3.72$). These findings are in line with the non-Bayesian analyses, which demonstrate that the interaction of task type and feedback type was significant in the theta but not delta frequency.

This pattern of results was also found in the ITPS and ICPS measures. For all frequencies and measures the main effects models were extremely compelling when compared to the null (BF_{10} ranging from 2.56e6 to 1.28e16). For the theta measures, ITPS and ICPS, the interaction model was supported over the main effects model. In theta ITPS there was weak support for the inclusion of the interaction term (BF_{21} = 2.90), while in theta ICPS there was substantial support for the inclusion of the interaction term compared to the main effects model (BF_{21} = 7.15). Instead, in delta ITPS there was substantial evidence for the main effects model being more compelling than the interaction model (BF_{12} = 4.08). Again, these findings provide further support for the non-Bayesian analyses.

Chapter 4: Discussion

The present study used time-frequency methodology to investigate the effect of action-outcome agency on gain and loss processing within our developed gambling tasks, designated high-control and low-control. As hypothesized, there was a global diminution in activity for the low-control task regardless of trial type and frequency. Because participants in the low-control task were instructed that their choices had no bearing on the subsequent monetary outcome, this finding could reflect an overall decrease in attention or motivation in this task relative to the high-control task. This result was seen in all three measurements analyzed in this study (i.e. ERP TF-amplitude, TF-ITPS, and TF-ICPS), which supports our assertion that the low-control task sufficiently primed participants to internalize that they had no control over the low-control task's action-outcome contingencies.

Since the high-control task was very similar to gambling tasks used to elicit the FN and RewP components in past research, there were significant differences between gain and loss processing as predicted. Consistent with previous findings in this lab, in both gambling tasks theta amplitude activity was more sensitive to loss trials while delta amplitude was sensitive to gain trials (Bernat et al., 2011; Bernat et al., 2015). Interestingly, there were differences between theta and delta amplitude in the low-control task. Delta sensitivity to gains relative to losses was very similar (e.g. not significantly different) in both the high and low-control tasks. In fact, delta amplitude differences between gain and loss trials were actually nominally greater in the low-control task. In comparison, theta's sensitivity to losses compared to gains was significant in the high-control task and relatively nonexistent in the low-control

24

task. We had hypothesized that in the absence of perceived control over action-outcome contingencies there would be no difference between gain and loss processing in theta and delta, yet this effect was only seen in theta.

Similarly, TF-ITPS measures in both theta and delta frequencies displayed the same pattern of results as that of the amplitude measures. Delta ITPS was significantly enhanced for gain trials compared to loss trials in both the high and low-control tasks, while theta ITPS was significantly sensitive to loss trials only in the high-control task. Therefore, delta gain trials in both tasks reflected more consistent and efficient processing than delta loss trials, while theta loss trials reflected more consistency in ERP response solely in the low-control task. Our TF-ICPS measure was utilized as an indicator of functional engagement of lateral prefrontal areas associated with cognitive control processes. In the theta frequency, functional connectivity between medial frontal regions (representative of theta amplitude activity) and lateral prefrontal regions was significantly increased for loss trials in the high-control task but not the low-control task. Thus, loss trials, which signaled a discrepancy from the desired outcome, triggered the recruitment of more cognitive resources to avoid future loss. This enhancement of cognitive control processes on loss trials was not seen in the low-control task, consistent with the idea that there was no need for further processing to modulate behavior on subsequent trials because there was no relationship between behavior and outcome. Additionally, theta ICPS was able to uniquely explain differences in feedback processing above and beyond theta amplitude and ITPS measures in the high-control task but not in the low-control task. The theta-ICPS measure, a signifier of engagement of cognitive control

25

processing, explained unique variance in feedback processing only when participants'

had control over their action-outcome contingencies.

Through analysis and interpretation of these results, one question remains:

why is theta but not delta processing modulated by action-outcome agency? The PRO

model proposed that discrepancies between actual and expected outcomes are

monitored by activity originating in the ACC, regardless of outcome valence. If this

were true, we would expect to see diminutions in gain/loss differences in the low-

control task in both theta (loss) and delta (gain) processing. The results of this study

suggest that only processing of losses is central to updating mental representations of

action-outcome contingencies and thus modulating behavior to attain optimal task

performance. This interpretation challenges the position underlying the PRO model

that discrepant outcomes, regardless of valence, affect the learning process. Instead,

reward processing seems to be independent from the learning process and remains

intact when the ability to modulate behavior is removed.

Limitations and future directions

These results suggest that the reward processing underlying the FN is

separable from the action-outcome processing thought to be characteristic of this

component. In order to further elucidate existing literature surrounding the RewP

component and challenge the main assumption within the PRO model, replication of

these effects in larger samples is necessary. Also, since there is no gold standard task

that manipulates action-outcome agency, future studies should continue to investigate

the role of participant agency in action-outcome learning in a variety of learning

tasks. Finally, this study demonstrated that the processing of negatively and

positively-valenced feedback within the FN time-domain component can be independently analyzed and be modulated differently. Utilizing the time-frequency approach to ERP analyses should be an integral addition to all future FN research to better extract elements that affect loss and gain processing separately. This study used new analytic methods to explore the relationship between learning and reward processing in the FN time component and found that losses and not gains were sensitive to manipulation of action-outcome agency, a vital component of task learning.

Tables

Table 1.

Bayes Factor interpretations for all possible Bayes Factor values. Potential values
for all Bayes Factors that favor the null model (left-hand side) and values that favor
an alternative model (right-hand side) are included.

	For Null	For Alternative
Weak Support (Uninformative)	.33 – 1	1-3
Substantial Evidence	.10 - .33	3-10
Strong Evidence	.03 - .1	10 - 30
Decisive Evidence	<.01	>100

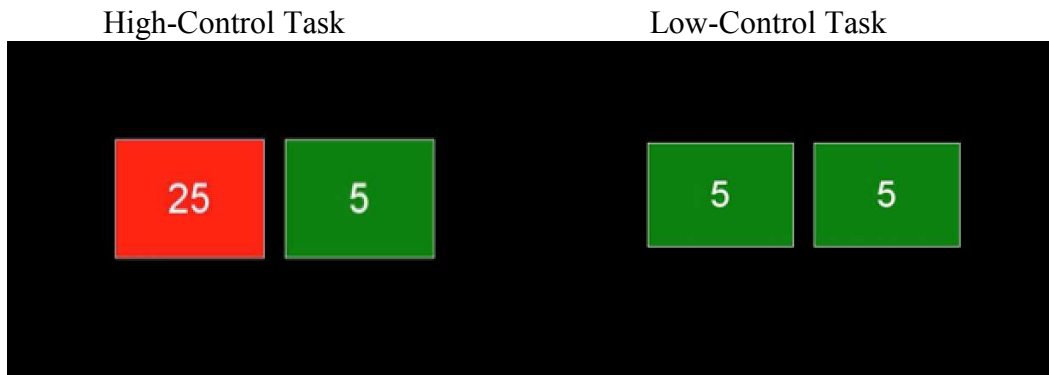High-Control Task   Low-Control Task

Figure 1. Example trials from the high-control task (left) and low-control task (right).

In the high-control task both the point amounts (25 vs. 5) and outcomes (red vs.

green) are different. In the low-control task both squares show the same point
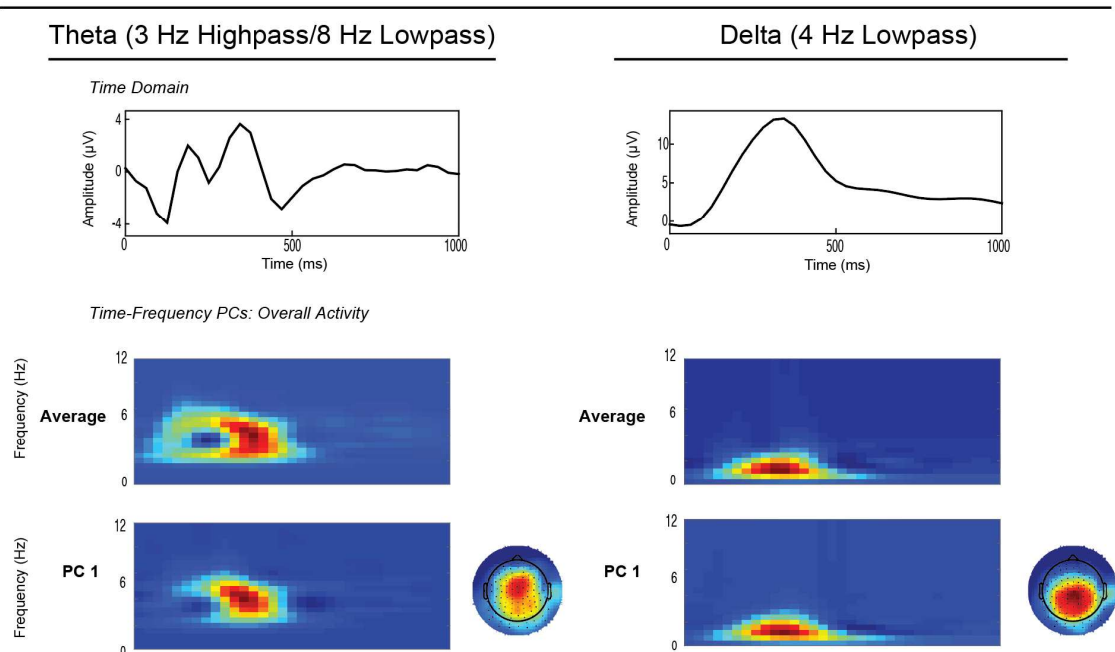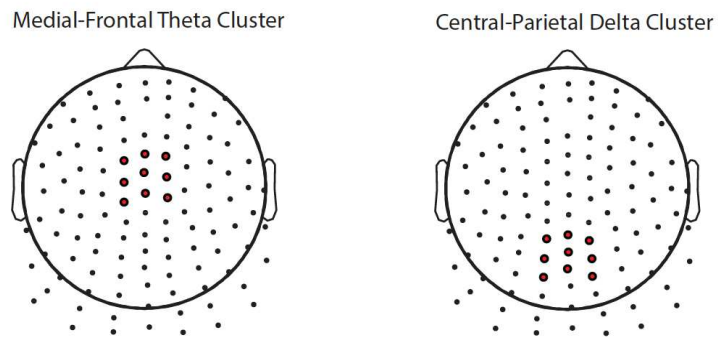
amounts and outcome.

Figure 2. Grand average TF-PCA solutions for theta (left) and delta (right) activity.

One PC solutions were chosen for both theta and delta to best characterize the activity

within the FN time window. As predicted theta activity was maximal over medial-frontal regions and delta activity was maximal over centro-parietal regions.


Figure 3. Nine-electrode clusters were extracted for further statistical analysis based on the maximal theta and delta activity.
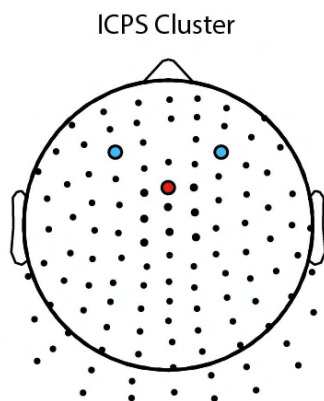
Figure 4. Inter-channel phase synchrony (ICPS) was computed between medial-frontal electrode 61 (red) and two lateral prefrontal electrodes 33 and 88 (blue).
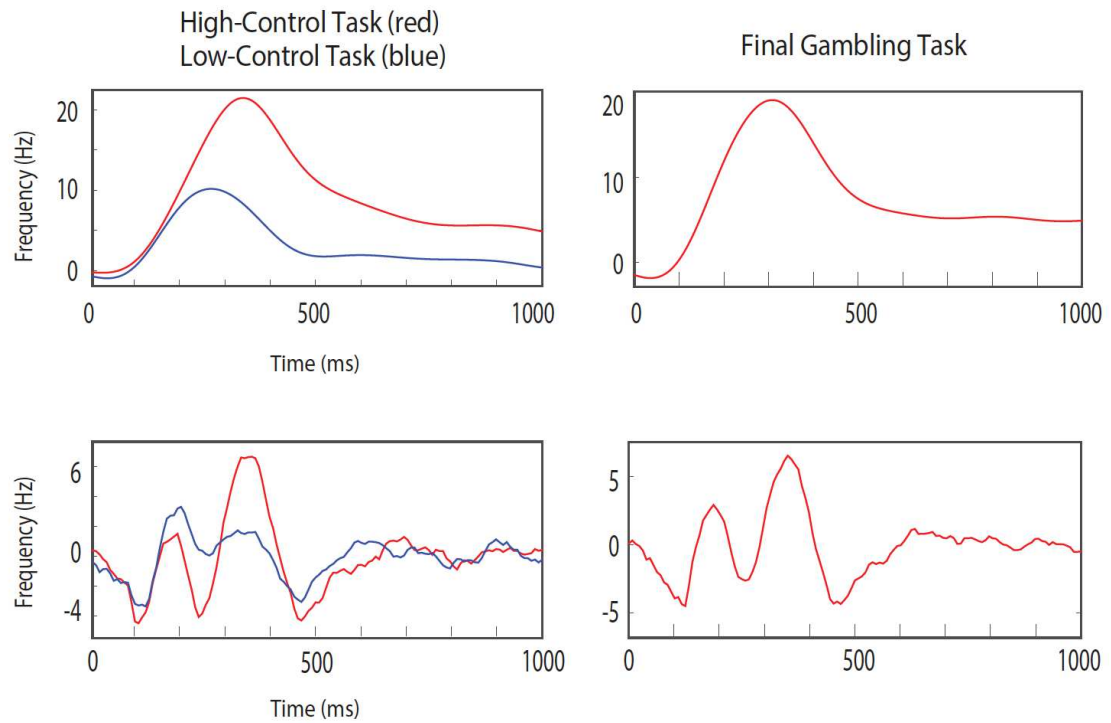
Figure 5. Average delta amplitude on gain trials for the high and low-control tasks and the final gambling task administered (top row) and average theta amplitude on loss trials for the high and low-control tasks and the final gambling task (bottom row).
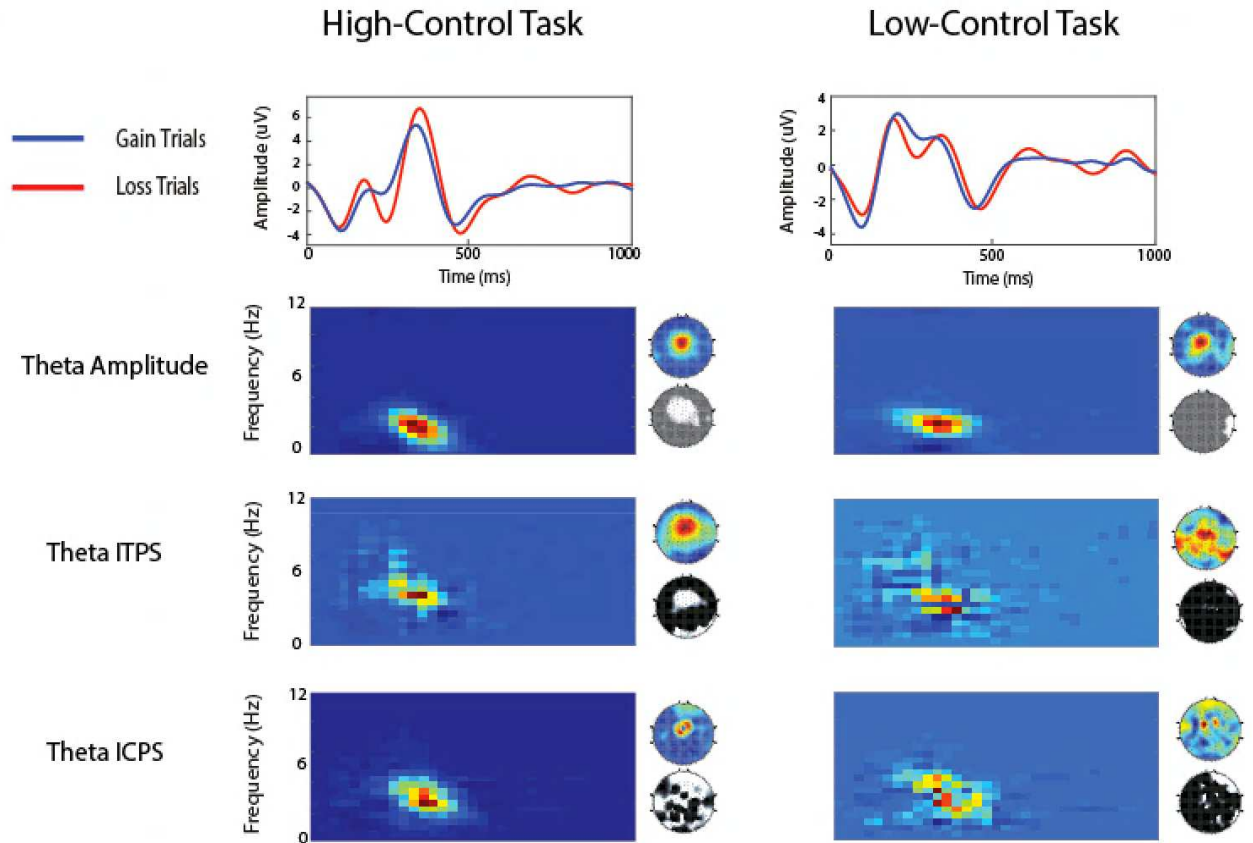
Figure 6. Time-domain and time-frequency plots for theta amplitude, ITPS, and ICPS

measures. Plots display gain-loss differences for the high-control task (left) and the

low-control task (right). Topographs display the distribution of the activity associated

with the difference between loss (red) and gain (blue) trials and the significance of

that difference. White in the significance topographs signifies a significance level <

.05.
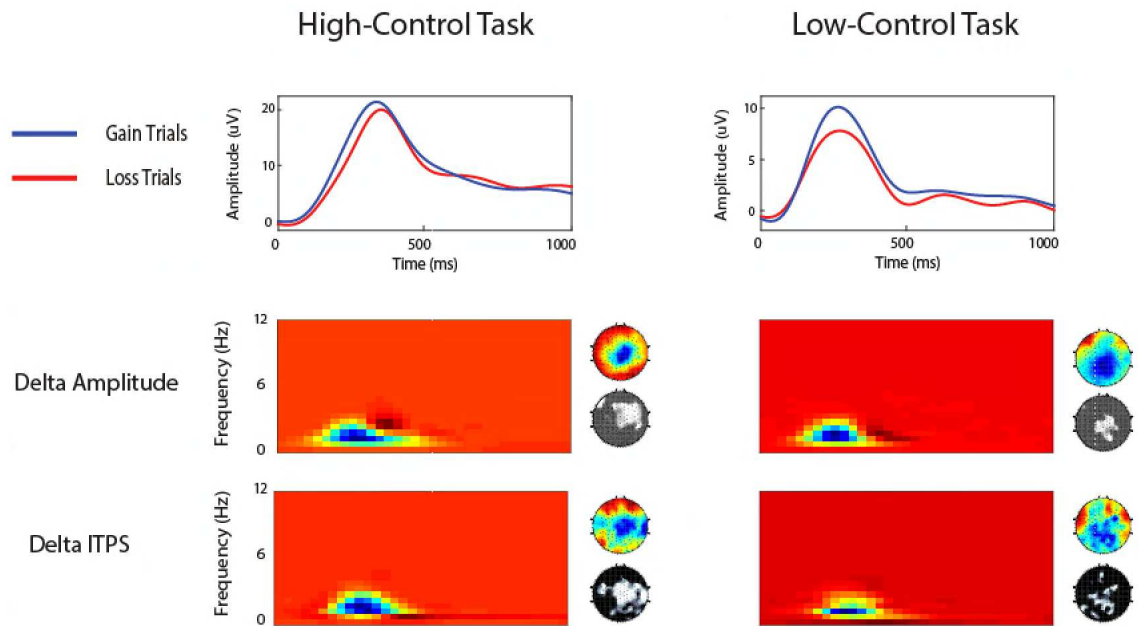
Figure 7. Time-domain and time-frequency plots for delta amplitude, ITPS, and ICPS measures. Plots display gain-loss differences for the high-control task (left) and the low-control task (right). Topographs display the distribution of the activity associated with the difference between loss (red) and gain (blue) trials and the significance of that difference. White in the significance topographs signifies a significance level < .05.
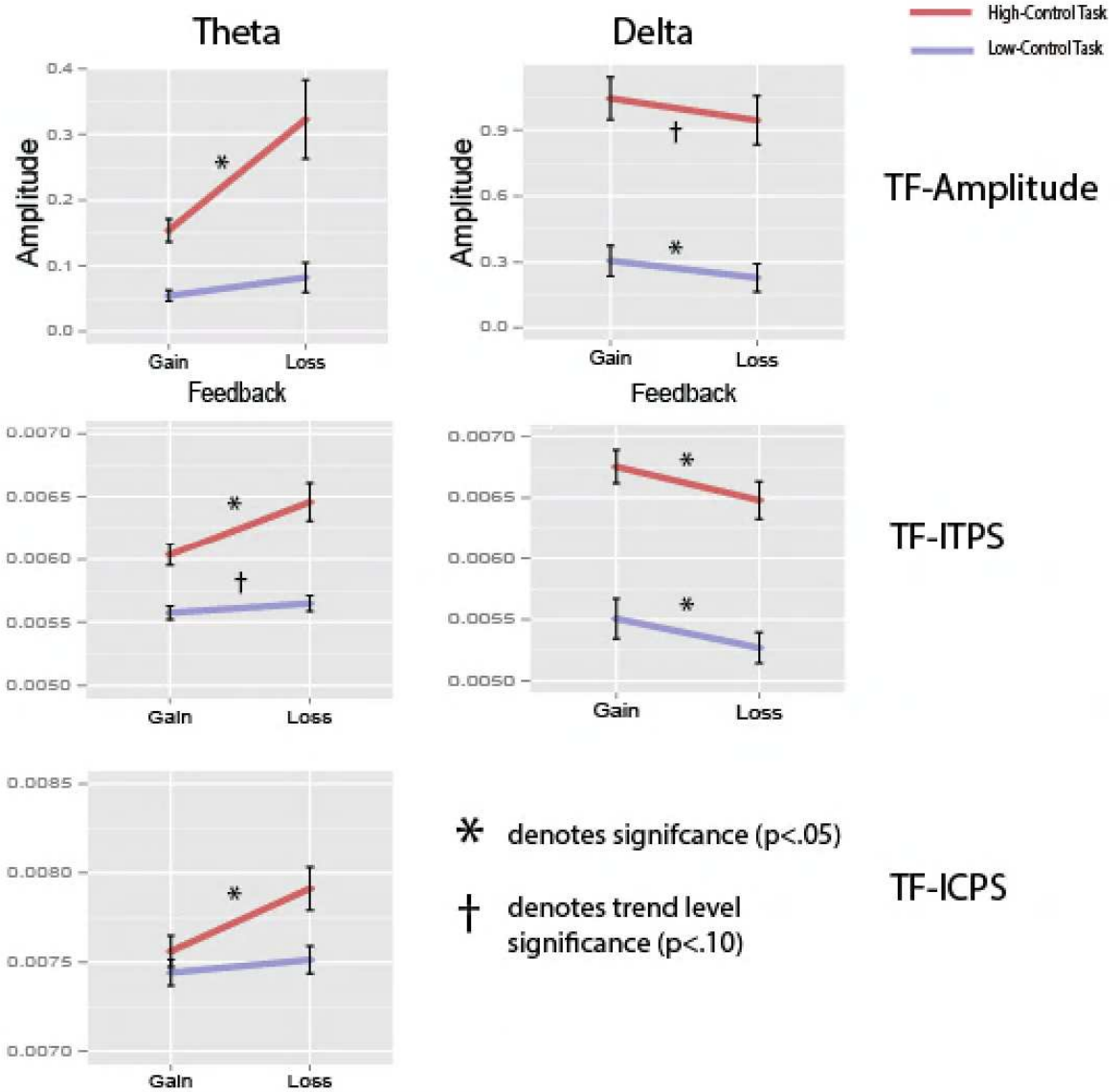
Figure 8. Average line plots for TF-Amplitude (top row), TF-ITPS (middle row), and

TF-ICPS (bottom row) measures on both gain and loss trials. Mean values for the

high-control task are represented in blue and values for the low-control task are

represented in green.

References

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344.

Aston-Jones, G., & Cohen, J. D. (2005). AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Aviyente, S., Bernat, E. M., Evans, W. S., & Sponheim, S. R. (2011). A phase synchrony measure for quantifying dynamic functional integration in the brain. *Human brain mapping*, *32*(1), 80-93.

Aviyente, S., & Mutlu, A. Y. (2011). A time-frequency-based approach to phase and phase synchrony estimation. *IEEE Transactions on Signal Processing*, *59*(7), 3086-3098.

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.

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. *The Journal of Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/JNEUROSCI.3684-13.2014

Bernat, E. M., Cadwallader, M., Seo, D., Vizueta, N., & Patrick, C. J. (2011). Effects of Instructed Emotion Regulation on Valence, Arousal, and Attentional Measures of Affective Processing. *Developmental Neuropsychology*, *36*(4),

493–518. https://doi.org/10.1080/87565641.2010.549881

Bernat, E. M., Nelson, L. D., & Baskin-Sommers, A. R. (2015). Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*, *52*(5), 626–637. https://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 *Optical Engineering+ Applications* (p. 70740S–70740S). International Society for Optics and Photonics. Retrieved from http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1327214

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences*, *99*(1), 523-528.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *The Journal of Neuroscience*, *29*(1), 98–105.

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

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, *35*(2), 968–978.

Crino, P. B., Morrison, J. H., & Hof, P. R. (1993). Monoaminergic innervation of

cingulate cortex. In *Neurobiology of cingulate cortex and limbic thalamus* (pp. 285-310). Birkhäuser Boston.

Delgado, M. R., 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.

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

Foti, D., Weinberg, A., Bernat, E. M., & Proudfit, G. H. (2014). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology*. Retrieved from http://www.sciencedirect.com/science/article/pii/S1388245714005148

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

Gehring, W. J., & Fencsik, D. E. (2001). Functions of the Medial Frontal Cortex in the Processing of Conflict and Errors. *The Journal of Neuroscience*, *21*(23), 9430–9437.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385–390.

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

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912.

Hall, J. R., Bernat, E. M., & Patrick, C. J. (2007). Externalizing Psychopathology and the Error-Related Negativity. *Psychological Science*, *18*(4), 326–333. https://doi.org/10.1111/j.1467-9280.2007.01899.x

Holroyd, C. B., & Coles, M. G. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex*, *44*(5), 548-559.

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

Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain research*, *1105*(1), 93-101.

Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917.

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 59–70.

Holroyd, C. B., Krigolson, O. E., & Lee, S. (2011). Reward positivity elicited by

    predictive cues. *Neuroreport*, *22*(5), 249–252.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the

    event-related brain potential associated with reward and punishment.

    Psychophysiology, *41*(2), 245–253. https://doi.org/10.1111/j.1469-

    8986.2004.00152.x

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G.,

    … others. (2004). Dorsal anterior cingulate cortex shows fMRI response to

    internal and external error signals. *Nature Neuroscience*, *7*(5), 497–498.

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.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring

    by the anterior cingulate cortex during saccade countermanding. *Science*,

    302(5642), 120-122.

Jeffries, H. (1961). Theory of probability.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin

    Card Sorting revisited: distinct neural circuits participating in different stages

    of the task identified by event-related functional magnetic resonance imaging.

    The Journal of Neuroscience, *21*(19), 7733-7741.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004).

    Sensitivity of electrophysiological activity from medial frontal cortex to

    utilitarian and performance feedback. *Cerebral Cortex*, *14*(7), 741-747.

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of

> feedback expectancy and valence induces negative and positive reward

> prediction error signals manifest in event-related brain potentials.

> *Psychophysiology*, *48*(5), 656–664.

Potts, G., 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. *Cognitive Neuroscience, Journal of*, *18*(7), 1112–1119.

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a

> biomarker for depression. *Psychophysiology*, *52*(4), 449–459.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default

> Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*,

> *56*(5), 356-374.

San Martín, R. (2012). Event-related potential studies of outcome processing and

> feedback-guided learning. *Frontiers in Human Neuroscience*, *6*.

> https://doi.org/10.3389/fnhum.2012.00304

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol.

> 135). Cambridge: MIT Press.

van Veen, V., & Carter, C. S. (2002). The Timing of Action-Monitoring Processes in

> the Anterior Cingulate Cortex. *Journal of Cognitive Neuroscience*, *14*(4),

> 593–602. https://doi.org/10.1162/08989290260045837

Vassena, E., Krebs, R. M., Silvetti, M., Fias, W., & Verguts, T. (2014). Dissociating

> contributions of ACC and vmPFC in reward prediction, outcome, and choice.

> *Neuropsychologia*, *59*, 112–123.

https://doi.org/10.1016/j.neuropsychologia.2014.04.019

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

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*(5), 535–544.