FACTORS AFFECTING THE AMPLITUDE OF THE FEEDBACK-RELATED NEGATIVITY DURING THE BALLOON ANALOGUE RISK TASK

by

ANTHONY WILLIAM MCCOY

B.A., Albion College 2011

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Michael E. Young

Abstract

When making decisions, the probability and magnitude of errors can play a major role in changing preferences. Electroencephalography (EEG) research examining the error-related negativity (ERN) and the associated feedback-related negativity (FRN) has indicated that the amplitude of each component may predict subsequent behavioral change. The current study used a version of the Balloon Analogue Risk Task (BART) that involves outcomes that are dynamically changing over time. As the balloon grows, more points are available but the probability of the balloon popping (netting zero points) is higher; the participant decides when to stop the balloon's expansion to maximize points. The BART was adapted to facilitate the study of the FRN in dynamic environments. The purpose of Experiment 1 was to determine the effect of error magnitude on FRN amplitude during popped (incorrect) trials, whereas Experiment 2 was aimed at determining the effect of error magnitude on FRN amplitude during cashed-in (correct) trials. It was hypothesized that larger errors (i.e., the balloon popping after waiting a long time to cash-in) would result in a larger FRN than smaller errors. In Experiment 1, error magnitude did not contribute to the amplitude of the FRN. In Experiment 2, the masked points possible condition was a replication of Experiment 1. In the unmasked points possible condition, the number of points that could have been earned for each balloon was presented before participants found out how many points were earned. It was expected that there would be a larger FRN magnitude after cashed-in trials in the unmasked points possible condition compared to the masked points possible condition based on the magnitude of the error. In Experiment 2, the amplitude of the FRN was affected by the magnitude of the error on cashed-in trials in the unmasked condition, but not the masked condition. These results are seemingly at odds, and cannot be assimilated into any currently extant model of the FRN. An explanation relying on the motivational importance of errors is discussed.

Table of Contents

List of Figures

Acknowledgements

I could not have completed this major milestone without the care and support of numerous people. My parents have supported me in my academic and scientific endeavors since before I can remember, from getting me chemistry and microscopy kits when I was still in elementary school through encouraging me during tough times during my education at Kansas State University. My loving wife, whom I dedicated this document to, has been on the front lines as I have gone through the highs and lows of my early scientific career, and I truly would not be who I am today without her love and support.

My committee deserves a special place in this acknowledgment. Drs. Brase and Kirkpatrick spent hours reading through and editing this document in preparation for my defense, and provided challenging and enlightening comments and questions about the contents of this thesis. I thank them wholeheartedly for all of their hard work and dedication to me and my thesis.

Finally, the man who deserves the greatest acknowledgment for all of his hard work is my advisor Dr. Michael Young. Mike accepted me into graduate school at Southern Illinois University despite my sub-standard grade point average, and he has nearly single-handedly molded me into the scientist I am today. Early in my tenure in the Young lab, I assumed I knew more than I did, something that Mike did a good job of distilling out of me. Over continued experience in the lab, and through the direct tutelage of Mike I learned about decision making, impulsivity and the proper use of statistics. I learned how to program using the Torque video game platform, and I learned how to use multilevel modeling statistical techniques. Without this dedication from Mike, I can honestly say that I would be less of a scientist.

I thank everyone that has helped me get to this point.

Dedication

I dedicate this thesis to my loving wife Laura whose unerring dedication is evident in every page of this document.

Chapter 1 - Introduction

Errors are inevitable given the number of decisions individuals make across their lifetime. People are often faced with decisions with no obviously correct answer. For example, people must decide whether to purchase House A or House B- House A may be less expensive but require more repairs than House B. In this instance the correct answer is ambiguous at the time of the decision. Researchers often classify decisions where there is an inherent chance of adverse consequences as risky (e.g., Leigh, 1999). For example, assume the buyer decided to purchase House A, and it is later learned that it had shifted off its foundation, resulting in costlier repairs than previously determined. This outcome is deleterious as it results in a larger loss than previously thought. A complete avoidance of risk is, however, sub-optimal as it can result in worse outcomes than if a risk was taken. For example, consider a person who is saving for retirement and has $2,000.00 to put into either a certificate of deposit (CD) or a stock-based mutual fund. If the person decides to put the money in the CD the money will accrue interest slowly (e.g., 1.25%). This rate of return may be slower than inflation which would actually result in a loss. If however, the person invested in an index fund they are at a risk of losing money, but are likely to gain more money over time than if the money were placed in a CD.

These decisions are considered risky; they both have the potential for adverse consequences. When discussing risky choice, an error is often operationalized as having chosen an option which results in a sub-optimal outcome. Sometimes an individual knows immediately whether their choice was erroneous, but other times it may require additional information from the environment to determine whether a choice was indeed sub-optimal. In the laboratory, as in real life, decisions often include a safer option and a riskier option. The safe option typically has a smaller payoff (continuing to rent a home or buying a CD) whereas the risky option often has a larger possible payoff but can also lead to deleterious outcomes.

The purpose of this study was to better understand how errors are processed. These experiments employ electroencephalographic (EEG) techniques to assess the neural correlates of error processing. The event-related potential (ERP), a time-locked deflection in the continuous EEG signal (Luck, 2005), component of interest is known as the feedback-related negativity (FRN). This ERP component occurs roughly 250 ms after a participant receives feedback indicating an error was committed. The purpose of Experiment 1 was to assess the effect of error magnitude on FRN amplitude on error trials. The purpose of Experiment 2 was to assess the effect of error magnitude on FRN amplitude on so-called “correct” trials when the participant learned that a better outcome could have occurred.

Taken together, the experiments were designed to elucidate the effect of error magnitude on both correct and incorrect trials. Previous research has primarily focused on errors resulting in no or negative feedback, but have not examined the effects of deviation from the amount of reward possible as an error. For example, previous research has indicated that if a participant receives no reward when a reward of 100 points was possible an FRN is elicited. However, there is no known FRN research on the effect of deviations from possible- a participant receives 50 points when a reward of 100 points was possible. To better elucidate the purpose of these experiments, the literature on risky choice, reinforcement learning, error-related evoked potentials, the Reinforcement Learning Theory of the Error Related Negativity (RL-ERN), and the Balloon Analogue Risk Task (BART) will be discussed.

Risky Choice

Undue risk-taking can negatively affect individuals and society, and drug use is a prototypical example of risky behavior in which the rewards do not outweigh the costs. According to the National Institute on Drug Abuse website (National Institute on Drug Abuse, 2012), drug use (including alcohol, tobacco products, and illicit drugs) costs the United States over 600 billion dollars annually due to crime, lost work productivity, and healthcare. Drug use amongst teens and young adults is especially problematic. The use of marijuana by high school students has been on the rise with an increase of roughly 3% in 10^{th} graders and slightly over 4% increase in 12^{th} graders from 2007 to 2012. In 2012 roughly 15% of high school seniors had used a prescription drug non-medically in the last year with this statistic mainly being driven by the use of psychostimulants such as Adderall and prescription pain medications such as Vicodin. There are also areas with reduced drug use in the teenage population; use of alcohol and cigarettes has dropped in the population with historic lows for alcohol use, and fewer students smoking cigarettes than smoking marijuana. Also, aside from marijuana use, drug use amongst all Americans has remained unchanged or has decreased over the last decade.

Drug use is not the only risk-taking behavior with a detrimental effect; risky sexual behaviors amongst young adults is also a major concern (Centers for Disease Control and Prevention, Centers for Disease Control, 2013). According to the CDC 15% of young people (ages 10-24) have had sex with at least four separate partners during their lifetime, and 40% had not used a condom during their last sexual encounter. These statistics have culminated in roughly 1,000,000 young people in the United States living with a sexually transmitted infection (specifically gonorrhea, chlamydia, or syphilis). Also, the rates of HIV and AIDS increased in this population between the years 1997 and 2006.

These risky behaviors have cost the U.S billions of dollars over the course of the last few decades, and the question of what variables moderate proclivity towards risky behavior are now at the forefront of psychological research. One of the reasons why this research has not been as fruitful as possible is the difficulty of translating the laboratory findings on the neural mechanisms underlying risk-taking into successful real-world interventions, and thus this research has been slow to influence intervention practices. The purpose of using a variation of the BART is that it has been shown to differentiate clinical and non-clinical populations based on risk propensity. Hopefully, by using a task with high validity, the results of this and future research using this paradigm will translate to intervention practices more rapidly.

Recent research has indicated the amplitude of the error-related negativity (ERN), an ERP component related to error commission, is diminished in individuals with cocaine dependence (Franken, van Strien, Franzek, & van de Wetering, 2007). Franken and colleagues had participants (cocaine dependent individuals and normal controls) complete a version of the Eriksen Flanker task in which the target symbol was either an “S” or an “H” and was flanked by either Ss or Hs. Participants responded with either index finger based on the identity of the target letter. EEG was recorded while participants completed the task. The results indicated that cocaine dependent individuals had reduced ERN amplitude on error trials when compared to controls.

Reinforcement Learning

Drug use may impair an individual’s ability to learn the optimal strategy for completing a task. Using the history of reinforcement to learn an optimal behavior relies on reinforcement learning. Reinforcement learning is a type of machine learning (R. S. Sutton & Barto, 1988) in which an actor learns through interactions with the environment, and through trial and error learns the behaviors that lead to the highest rates of reinforcement. The best outcome is the outcome that best aligns with the learner's goal at the time of the choice. For example, if an individual has a choice between studying for an exam and watching a favorite TV show, the choice with the greater utility would be based on their goal at the time. If the goal is to do well on the test, then the maximizing option is to study. If the goal is to relax, the maximizing option is watching TV.

An important facet of reinforcement learning is the need to trade-off between exploration and exploitation. Exploration is generally the first step in solving the types of problems associated with reinforcement learning and involves selecting unknown-payoff options to assess the reward value of each option; thus exploration involves taking a risk. Exploitation generally occurs once an actor has found a rich option and selects that option repeatedly for a period of time. It benefits the actor to explore the environment to determine if there are richer choices available. There are two models of how actors are thought to explore their environment (R. S. Sutton & Barto, 1988). The first model suggests that the actor tends to choose the option with the richest reinforcement schedule most of the time, but when exploring, randomly chooses from the options that are less preferred (known as the ε-greedy strategy). This model does not describe human behavior (R. S. Sutton & Barto). The second model suggests that the actor tends to choose an option based on Luce's Decision Rule (Luce, 1959) (aka Softmax, Multinomial Logit). The equation for Luce's Decision Rule is shown below (Equation 1):

$$P(R_i) = \frac{e^{\theta v_i}}{\sum_{j=1}^{n} e^{\theta v_j}} \tag{1}$$

where *e* is a mathematical constant known as Euler's number, $P(R_i)$ is the probability of choosing option i and Θ is the exploitation parameter, which varies between 0 (equal probability of choosing each option) and ∞ (exclusive choice of the option with the highest known payoff).

The v_i is the value of the option under evaluation and v_j is the value of each available option. Luce's Decision Rule suggests that actors will choose options based on the known payoff rates. An actor will choose the option with the best payoff rate most often, the option with the second best payoff rate the second most often, etc. Consider a game of chance in which there are three options to choose from. Option 1 pays off 65% of the time, Option 2 pays off 45% of the time, and Option 3 pays off 25% of the time. In this example, the actor will choose Option 1 most often, followed by Option 2, and Option 3 the least often.

An example of a task in which reinforcement learning is well defined is the multi-armed bandit task (e.g., Anderson, 2012; Murray, 1971; Racey, Young, Garlick, Pham, & Blaisdell, 2011; Thomas, Kacelnik, & van der Meulen, 1985); in this task participants are presented with a number of alternatives with varying levels of reinforcement availability. Participants select any of the available options and receive the reward associated with the choice; on each subsequent choice participants can again select any of the available options. Using Luce's Decision Rule, the probability of a participant selecting the reward maximizing option is highest and the probability of a participant selecting the reward minimizing option is lowest. In this task, participants interact with the alternatives and learn which actions yield the richest outcomes. It is important to note that short-term rewards are greater during exploitation than during exploration, but in the long-term initial exploration will result in greater rewards due to greater understanding of the parameter space. The participant learns from previous outcomes and these outcomes allow the participant to estimate the size of the reward to be expected from a given choice on the current iteration. If the reward varies from what was expected (either better or worse) then the expectation of the outcome for the next iteration is updated; this method allows the participant to learn the appropriate action for the task at hand (R. S. Sutton & Barto, 1988).

It is evident that the behaviors present in the multi-armed bandit task are similar to those behaviors seen in drug use. In this task participants explore or exploit the choices present. Drug users tend not to adjust their behavior based on the negative outcomes from their behavior (e.g., Piazza & Deroche-Gamonet, 2013). In fact, one of the sets of criteria set forth by the Diagnostic and Statistical Manual of Mental Disorders 5th edition for substance use disorders is the continued use of the drug of abuse despite negative health, social/relational, vocational, and/or legal outcomes (American Psychiatric Association, 2013). This continued use despite negative outcomes can be conceptualized as a penchant towards exploiting known rewards. Addicott, Pearson, Wilson, Platt, and McClernon (2013) reported that smokers tend to over-exploit during the multi-armed bandit task. They found that non-smokers tended to explore the non-reward maximizing options more often than smokers who tended to primarily exploit the richest option.

The findings reported by Addicott et al. (2013) suggest that smokers may be more affected by the short-term consequences of their behavior compared to the long-term consequences of their behavior. Recall that initial exploration results in less reinforcement in the short-term but more reinforcement in the long-term, indicating that smokers behavior is affected more by short-term outcomes than long-term outcomes, which may explain why they ignore the long-term health consequences of their behavior. Research using the Balloon Analogue Risk Task (Fein & Chang, 2008; Hopko et al., 2006; Lejuez et al., 2007; Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, & Pedulla, 2003) has shown that performance on the BART is correlated with real-world risky behaviors such as drug use. In these studies, drug users and those individuals with predispositions towards drug use tend to behave in a more risky manner than controls. The extant literature on drug abuse seems to suggest that individual differences in predisposition to exploit (higher Θ values indicate a greater likelihood of exploiting known options) may be associated with a penchant to abuse drugs. Abusers fail to learn the reward maximizing option (abstinence) because they are focused on the short-term rewards of drug use and are insensitive to the long-term consequences. Research by Franken et al. (2007) indicating that cocaine dependent participants had a reduction in ERN amplitude may suggest that the amplitude of the ERN is related to reinforcement learning.

Error-related Evoked Potentials

Researchers using EEG to study various phenomena often have participants make decisions based on what was experienced, but until the early 1990s error trials were ignored (Falkenstein, Hohnsbein, & Hoorman, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). In go/no-go studies (e.g., Mishkin & Pribram, 1955) participants are required to respond when a “go” stimulus is presented and are required to omit a response when a “no-go” stimulus is presented. In typical studies using this paradigm, one of the trial types is rare whereas the other is common. Researchers focus on rare trials and analyze an event-related potential (ERP) component known as the P300 (S. Sutton, Braren, Zubin, & John, 1965). The P300 (see Figure 1.1) is a late posterior component occurring roughly 300 ms post-stimulus onset. The P300 occurs on trials where the stimulus presented was not expected or was surprising. Often error trials were omitted from the analysis to ensure that the participant processed the stimulus.

Ignoring error trials in this manner made it impossible to understand how people processed errors, which led to the relatively late discovery of the error-related negativity (ERN). This discovery was made independently by two teams of researchers, Falkenstein, Hohnsbein, and Hoorman in 1991 and Gehring, Goss, Coles, Meyer, and Donchin in 1993. The ERN has been shown to peak roughly 100 ms post-response and is elicited on error trials (see Figure 1.2). The ERN is often studied in reaction time tasks such as the Eriksen Flanker task (B. A. Eriksen

& Eriksen, 1974; C. W. Eriksen, Hamlin, & Daye, 1973). In the Flanker task participants see a centrally presented target object (e.g., "<") which is flanked by congruent (e.g., "<<_<<") or incongruent (e.g., ">>_>>") flankers. Participants have to respond by pressing the appropriate button. Participants are told to respond as quickly and as accurately as possible, and errors are readily noticed by participants. Research on the ERN eventually led to the discovery of the feedback-related negativity (FRN) (Miltner, Braun, & Coles, 1997).

The FRN is time-locked to a feedback stimulus rather than a response and occurs 250-300 ms post-feedback onset in studies where error commission is unknown until after feedback is received (Gehring, Liu, Orr, & Carp, 2012; Miltner et al., 1997). The FRN is commonly described as analogous to the ERN, and it has been suggested that the two components reflect the same error processing mechanism (Holroyd & Coles, 2002).

Previous research has examined the effects of error magnitude on the amplitude of the FRN; the findings have been mixed. Some research has indicated that the magnitude of the error has no effect on the magnitude of the FRN, indicating that the FRN is evoked from a binary outcome monitoring mechanism (e.g., Hajcak, Moser, Holroyd, & Simons, 2006) whereas other research seems to indicate that the amplitude of the FRN is associated with error magnitude (Bellebaum & Daum, 2008; Bellebaum, Polezzi, & Daum, 2010; Goyer, Woldorff, & Huettel, 2008).

More recently, the FRN has been studied using risky choice paradigms (e.g., Goyer et al., 2008; Hajcak, Moser, Holroyd, & Simons, 2007; Marco-Pallares, Cucurell, Munte, Strien, & Rodriguez-Fornells, 2011; Massar, Rossi, Schutter, & Kenemans, 2012; San Martin, Manes, Hurtado, Isla, & Ibanez, 2010). This research has focused on gambles in which the participant must select one of two choices and the participant is either rewarded or not. In one oft-used gambling task participants have to choose between two options, one with a high probability of obtaining a small reward (the “safe” option) and one with a low probability of obtaining a large reward (the “risky” option). In this task there are two possible outcomes for each choice-either the participant receives the amount gambled (win) or loses the amount gambled (loss). At the end of the trial participants are told whether they won or lost and how much—it is this feedback that elicits the FRN. The probability of receiving rewards or the magnitude of the reward is manipulated, and in most cases participants can learn to perform the task better.

These paradigms are ideal for testing the FRN as the outcome is unknown at the time of the decision, and the task can be learned. The FRN is associated with task learnability (Arbel, Goforth, & Donchin, 2013). It has been shown that the amplitude of the FRN elicited by an error is correlated with later performance on the task (van der Helden, Boksem, & Blom, 2010). In this line of research it has also been shown that the amplitude of the FRN is not associated with later performance when the underlying task structure is not learnable (Holroyd, Krigolson, Baker, Lee, & Gibson, 2009). For example, a task is learnable if there is some signal that indicates whether a given trial will be reinforced. If the probability of the trial being reinforced is randomly assigned then the task will not be learnable and thus will not elicit the FRN-performance association. Interestingly, it is possible for early trials to elicit an FRN because the response-reinforcement structure is unknown, but as the participant learns the response-reinforcement pairing, errors will begin eliciting an ERN (Holroyd & Coles, 2002). In this way, if the participant can learn which response to make on a given trial to obtain reinforcement, it is possible to see an FRN early, which will transition to an ERN once the correct response is known.

Reinforcement Learning Theory of the Error-Related Negativity

The Reinforcement Learning Theory of the ERN (RL-ERN), originally developed by Holroyd and Coles (2002), suggests that the ERN occurs due to an outcome monitoring system that relays information indicating that the outcome was worse than expected. For example, on the flanker task if the wrong response is selected, the error monitoring system indicates that an error has occurred and outcomes will be worse than expected. Errors are rare on the flanker task, so the expected value of any given trial should be close to the expected value of the correct response. However, on tasks where errors are more likely, the amplitude of the ERN may be smaller. Importantly, the error monitoring system responds to the earliest sign that an error has occurred. This theory parsimoniously explains the relationship between the ERN and the FRN. They are essentially the same error signal sent from the basal ganglia, and the signal occurs at the earliest sign that outcomes are going to be worse than expected. Although there is little published literature on the relationship of the ERN/FRN to substance abuse (Fein & Chang, 2008; Franken et al., 2007).

It is important to note that the RL-ERN theory suggests that the FRN elicited in EEG studies comes from the ventral surface of the anterior cingulate cortex. However, this signal comes rather late in the brain's processing of errors. The brain processes errors earliest in subcortical regions such as the ventral tegmental area (Wang & Tsien, 2011), but the electrical fields from these deep structures are too faint to be picked up by modern EEG, and thus EEG research must focus on later error processing. This reasoning, coupled with the theoretical relevance of the FRN is why we chose to focus on error processing in the anterior cingulate cortex.

BART as a Measure of Risk Propensity

The BART was initially developed by Lejuez et al. (2002) as a behavioral measure of risk propensity. Risk propensity is commonly assessed using surveys with questions about an individual's real-world risky behaviors and questions regarding known risk factors for behaving in a risky manner. It is, however, unknown to what extent someone can know the correct answer to these questions regarding risk factors. Often, it is thought that people look back on their last risky behavior and use that as an indicator of how to answer the questions. The drawback of asking directly about risky behaviors (e.g., in the past year how often have you used cocaine?) is it precludes the ability to ascertain the underlying causes of the risky behavior and does not allow for the measure of predisposition to new risky behaviors (Lejuez et al., 2002). Previous attempts at creating viable behavioral measures of risk-taking (e.g., the Iowa Gambling Task, Bechara, Damasio, Tranel, & Damasio, 1997) have had issues with convergent validity with survey-based measures. Of greater concern is the lack of association between the Iowa Gambling Task and other behavioral measures and self-reported real-world risk-taking behaviors (Lejuez et al., 2002). The BART, however, has consistently been shown to have good convergent validity with survey-based measures (e.g., Lejuez, Aklin, Zvolensky, et al., 2003; Lejuez et al., 2002).

In the BART, participants inflate a graphical representation of a balloon by pressing a key or by clicking the mouse and at any point during balloon inflation can cash-in by pressing a different key or clicking a button on the monitor. As the balloon increases in size the amount of points that can be earned by cashing-in increases linearly as a function of the number of presses or clicks, but at the same time the probability of the balloon popping and no points being awarded increases as well. The probability of the balloon popping has historically been governed by Equation 2:

$$Probability\ of\ Popping_i = \frac{1}{n-i+1} \qquad (2)$$

where *Probability of Popping* is the probability of the balloon popping on the i^{th} pump, *n* is the maximum number of pumps, and *i* is the pump number. Thus, the denominator of this equation indicates that as the participant pumps up the balloon, the probability of the balloon popping increases by a small amount contingent on how many times the participant has pumped up the balloon previously. The plus one is added to ensure that there is a 1/*n* probability of the balloon popping on the first pump. Figure 1.3(A) shows the function which defines the distribution of pop times. Figure 1.3(B) shows the probability of the balloon popping by a given pump number. In order to ensure the average pop time was similar across blocks, Lejuez and colleagues hand-picked pop times from the distribution of times created by the above equation. This procedure allowed the average pop times for all blocks to be artificially set at *n*/2.

BART as a Paradigm to Study Feedback-Related Negativity

The BART has high utility as a measure of risk propensity as is shown by the number of studies demonstrating its ability to differentiate between inherently risky populations and normal controls, such as, people with substance use disorders (Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Hopko et al., 2006; Lejuez et al., 2007; Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, et al., 2003), individuals diagnosed with psychopathy (Hunt, Hopko, Bare, Lejuez, & Robinson, 2005), and people with a family history of substance use disorders (Fein & Chang, 2008). The BART as it was originally conceptualized, however, is not feasible for EEG research. Pleskac, Wallsten, Wang, and Lejuez (2008) and McCoy and Young (In preparation) independently came to the conclusion that a version of the BART in which the participant passively observes the balloon expanding over time could have utility for use in neurophysiological paradigms. In Pleskac et al.'s (2008) version, the A-BART, participants indicate the number of pumps they would like the balloon to be inflated and then observe as the balloon is inflated that many times. Notably, participants are not allowed to defect (change their mind about the amount they want the balloon inflated) after their initial choice. In the version developed by McCoy and Young (In preparation), the participant passively watches the balloon expand over time and responds only to stop inflating the balloon. Both of these tasks involve fewer responses than the original BART because the balloon does not have to be pumped up, but show similar patterns of results. However, future research will be necessary to determine if the McCoy and Young version of the BART and the A-BART operate under similar neural mechanisms.

EEG research is sensitive to movement, and the pumping required by the BART represents a barrier to its use in EEG research. In pilot research, McCoy and Young (In preparation) tested whether the behavior of interest on the BART was tied to participants needing to press a button to pump up the graphical representation of a balloon. In Experiment 1 of their study, there was no behavioral difference in average latency to cash-in between whether the balloon inflated automatically and whether the participant had to manually inflate the balloon. In Experiment 2, there was no behavioral difference in average time to cash-in between whether the change in balloon thickness was cued or not. Participants in the no-cue condition took longer to reach asymptote than participants in the cued condition. EEG is sensitive to stimulus differences, thus the similarity between the cued and uncued versions in Experiment 2 was important.

The McCoy and Young version of the BART is used in the current study in lieu of the A-BART because it controls for stimulus confounds. EEG signals are sensitive to small differences in stimuli, so it is necessary for the stimuli on each trial to be identical. It is possible to make this change to the BART, but such a change to the A-BART would cause the balloon stimulus to become meaningless; fundamentally altering the task. Based on these results, the researchers determined that the BART would be a good candidate for EEG research on risky behaviors.

The variation on the BART used in the current study is different from the original BART in that the balloon inflates automatically and there are two balloon colors used. After each block of balloons the balloon color changes to indicate a change in the probability structure. This change was made to make it easier for participants to recognize and adapt to the change in balloon thickness. The two colors chosen (purple and green) were selected because they are unlikely to be confused and their discrimination is unlikely to be affected by color blindness (except in cases of monochromatic color blindness).

I made one additional change to the BART. The balloon always expanded to its maximum size on each trial regardless of when the participant cashed in or popped. This change ensured that, regardless of balloon thickness, participants would always see the same balloon animation, reducing stimulus related confounds.

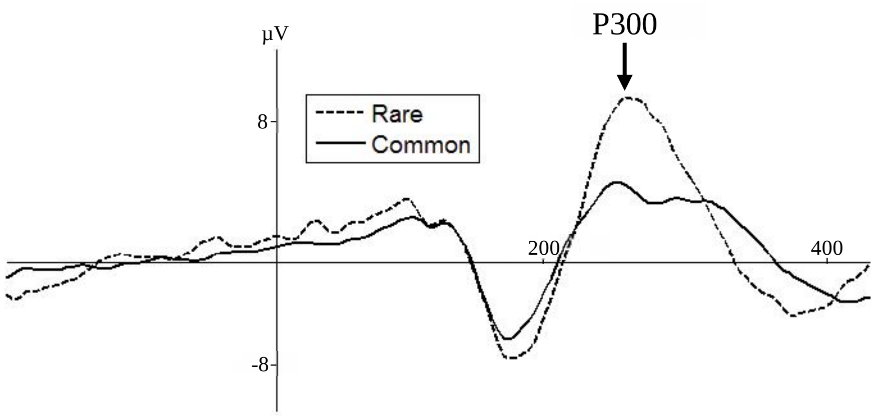


Figure 1.1. The P300 component evoked by a go/no-go study. The dashed line represents trials where the rare stimulus occurred. The solid line represents trials where the common stimulus occurred. The vertical bar represents stimulus onset.

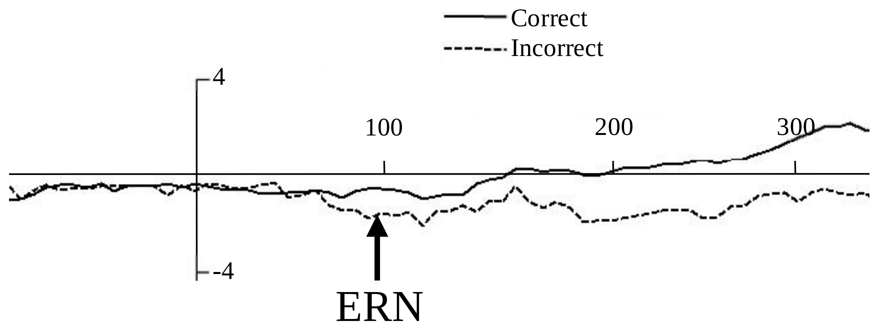


Figure 1.2. An error-related negativity evoked on error trials. The dashed line represents trials on which an error occurred, and the solid line represents trials on which an error did not occur. The vertical bar represents stimulus onset.

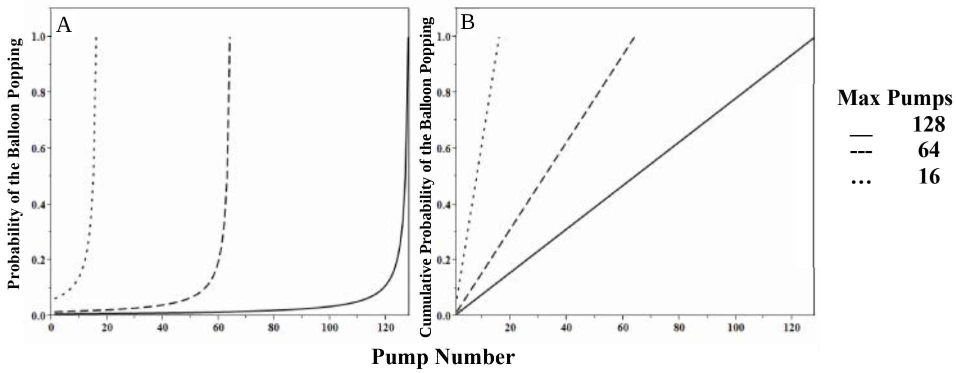


Figure 1.3. (A) Probability of the balloon popping as a function of pump number in the original BART. (B) Cumulative probability of the balloon popping as a function of pump number. Individual lines represent the three balloon "thicknesses" originally used.

Chapter 2 - Experiment 1

Previous research on the effect of error magnitude on the amplitude of the FRN has indicated that larger error magnitude is associated with greater FRN amplitude (e.g., Goyer et al., 2008). However, this finding has not been consistently found (e.g., Hajcak et al., 2006). The purpose of this experiment is to determine the effect of error magnitude on the amplitude of the FRN on error trials during a BART task. If high magnitude errors result in a larger FRN then this would be evidence in support of the RL-ERN theory. If, however, low magnitude errors result in a larger FRN or there is no difference in FRN amplitude across balloon thicknesses, then this would be evidence against the RL-ERN theory. While there is currently no alternative theory due in part to the relatively disparate results of recent research, further evidence against the RL-ERN theory may result in the formation of a new theory.

Methods

Participants

Participants were recruited from the General Psychology participant pool or through the campus community. A total of 17 undergraduates participated in the study. Two participants did not complete the study due to computer malfunctions.

Stimuli

In the current implementation of the BART, a trial comprised a single balloon animation with a randomly determined pop time with the maximum possible being 3400 ms. The balloon animation was a colored disk (presented on a black background) with a diameter that increased linearly with the passage of time. After 100 ms, the balloon was 0.5 cm in diameter (0.48˚ of visual angle) and grew at the rate of 0.5 cm (0.48° of visual angle) per 100 ms. The balloon always grew to its maximum size of 17 cm (15.82° of visual angle) in diameter regardless of when the participant cashed-in or when the balloon popped. This modification allowed us to ensure that the most prominent cue participants had to the outcome of the trial was the feedback we gave them at the end of the trial. Also, this change ensured that participants experienced the same visual stimulus on every trial, which reduces the potential for stimulus-related confounds.

Participants experienced a series of 450 balloons one at a time. These balloons expanded automatically and participants had to respond only to cash-in the balloon. The balloon color changed after each sub-block of 50 balloons alternating between purple and green to indicate there had been a change in the task. This procedure ensured that color was counterbalanced across the three balloon thicknesses (i.e., power values) to eliminate stimulus confounds. The participant then saw a fixation cross followed by a feedback slide. On the feedback slide the number of points the participant earned on the balloon was centrally displayed. This screen was followed by another fixation cross. Finally, participants saw a screen with the number of points they had accumulated presented at the bottom of the screen and the balloon number they had just completed at the top of the screen. The visual angle of the number of points earned was less than 2 degrees. The text displayed on all of the slides was white on a black background. Figure 2.1 shows the trial progression.

In the current study, pop times were determined using the superellipsoid function first adopted by Young, Webb, and Jacobs (2011) in their study of impulsive choice (see Figure 2.2A). This function has the advantage of maintaining the maximum possible pop time across variations in balloon thickness - likelihood that the balloon would pop sooner rather than later (cf. Lejuez et al., 2002). Thickness was varied by changing a single parameter, *power,* that altered the change in the probability across this duration:

$$Probability\ pop_t = \left(1 - \left(\frac{(3400-t)}{3400}\right)^{power}\right)^{\frac{1}{power}} \tag{3}$$

where *Probability pop* is the probability of the balloon popping at a given time, *t*, and *power* is the power parameter that defines the change in the probability of popping. As participants waited longer, the probability of the balloon popping increased. The power parameter adjusts the rate at which the probability changes across time. With higher powers (i.e., thinner balloons), the probability of successfully cashing in drops off quickly as time passes, whereas with lower powers (i.e., thicker balloons) the probability drops off slowly as time passes. Figure 2.2B shows the cumulative probability of the balloon having popped for various wait times for the three values of power used in the current study. For example, for the thinnest balloon, if the participant waited 2000 ms, the probability of the balloon having popped would be almost 100%. In contrast, for the thickest balloon, if the participant waited 2000 ms the probability of the balloon having popped would be roughly 5%.

The power parameter defined the changes in the likelihood of popping and was manipulated within-subject. A total of three different power values were experienced by participants spread across nine sub-blocks of trials. The expected value for each level of power value across latencies is presented in Figure 2.3. Depending on the power value, the expected value of cashing in at a given time was different. For example, if a participant waited 1000 ms to cash-in he or she should expect to obtain on average 50 points for the thinnest balloons, 500 points for the middle balloon thickness, and 1000 points for the thickest balloon.

Procedure

The study took place in an electrically shielded chamber (Ark Electronics Corporation). Upon arrival, participants were fitted for a Hydrocel Geodesic Sensor Net (EGI) (see electroencephalography procedures for description).

At the beginning of each trial a fixation cross was centrally displayed and the participant had to press any button on the response box to begin the balloon animation; the animation ran for 3400 ms. After the animation, a white fixation cross was displayed on a black background for 1000 ms. The termination of the fixation cross was jittered by 50 ms to attenuate the effects of alpha wave interference (Luck, 2005). The fixation cross was followed by a feedback slide presented for 1000 ms. If participants responded to cash-in before the balloon popped, this screen indicated the number of points earned on the trial (1 point per ms waited). If participants responded to cash-in after the balloon popped, then this screen indicated that the participant received zero points (shown simply as “0”). If participants did not respond while the balloon was on the screen, this screen said, “Please respond faster.” Trials in which participants did not respond during the animation were excluded from the analyses. After the feedback screen disappeared, a second fixation slide (identical to the first) was displayed for 1000 ms. Following the second fixation cross, a screen indicating the number of points the participant had accumulated and the balloon number that was just completed was displayed for 500 ms. After the accumulated points screen disappeared, the next trial began. After each block of 150 trials there was a rest break, during which the researcher checked the electrode impedance and corrected any that were above threshold (see the electroencephalography methods for details). At the end of the break, a 5 s countdown appeared indicating the experiment was about to recommence. During the countdown the numbers “5”, “4”, “3”, “2”, and “1” flashed sequentially for 1 s each in the center of the screen. Centrally displayed at the top of the screen were the words, “The next session will begin in:”

Participants experienced nine sub-blocks of balloons (see Figure 2.4). The order of power values experienced by participants was Latin Square counterbalanced within-subject. Each sub-block included 50 trials, and balloon color changed at the end of each sub-block to signal the transition from one sub-block to the next. Each power value was experienced three times and participants only experienced a given power value once (one sub-block of 50 balloon trials) within a block of 150 balloons (from one break to the next).

Electroencephalographic Methods

The Hydrocel Geodesic Sensor Nets (HCGSN) are designed as a minimally invasive measure of electrical activity on the scalp. The HCGSN uses AgCl electrodes. The nets have an array of 64 electrodes laid out based on the expanded 10/20 system (a standardized system used to ensure that across laboratories electrodes are placed in roughly the same area, see Figure 2.5). Once a net was selected, the researcher soaked the net in an electrolyte solution containing potassium chloride and baby shampoo; the baby shampoo breaks down the grease on the scalp and improves conduction. The net was soaked for 5 minutes before fitting it to the participant.

Participants were fitted for a Net (EGI) by measuring the circumference of his/her head using a measuring tape. Cz was determined using the 10/20 system; the researcher marked the location of Cz by measuring from nasion to inion and taking the center point and measuring from left preauricular point (left meatus) to right preauricular point (right meatus) and taking the center point. Where these two imaginary lines cross is the Cz.

After the cap was appropriately placed on the participant’s head, the researcher scrubbed (moved back and forth) each electrode to ensure good contact with the scalp. Once the net was fully placed, the participant took a seat in front of the stimulus presentation monitor that was located roughly 60 cm from the participant. Once seated, the researcher checked the impedance of the electrodes using Netstation's built-in capability; the researcher injected additional electrolyte solution and scrubbed any electrode with an impedance over 60 kΩ until all electrodes read under 60 kΩ. Once all impedances were within the acceptable range, the participant was shown their EEG record and told to perform various behaviors (e.g., eyeblink, eye movement, jaw clench). This step demonstrated the effect these behaviors can have on the EEG to discourage them from producing these behaviors during a trial. After being shown the EEG, participants read the onscreen instructions and began the study. For the purposes of this experiment, the FRN was operationalized as the voltage 250-300 ms post feedback onset. An FRN occurs when the voltage on trials where the balloon popped is lower (not necessarily negative) than on trials where the balloon was cashed-in (see Figure 2.6). Figure 2.6 shows an example epoch (segment of EEG time-locked to an event of interest). For the purposes of the analyses below, the FRN will be used to describe the difference in voltage between cashed-in and popped balloons.

Results

On average participants had 398 good EEG trials, roughly 88% of the 450 total trials experienced. Of the seventeen participants, five experienced computer malfunctions resulting in the experiment prematurely terminating. Two of these participants completed fewer than 75% of the 450 trials and were excluded. Additionally, two participants did not show any behavioral sensitivity to the manipulations and were also excluded, leaving 13 participants with usable data. Finally, the first block of the experiment was not analyzed as participants were not showing consistent sensitivity to the balloon thickness in this block.

Behavioral analyses

The behavioral analyses ensure that behavior on this task was similar to behavior in the original BART, and ensure that participants were sensitive to the balloon thickness manipulation. Three analyses were run on separate portions of the behavioral data. The first analysis assessed how long participants waited to cash-in the balloon. The second analysis assessed how much variability there was in participants' latency to cash-in the balloon. The third analysis assessed how often participants popped the balloon.

The first behavioral analysis assessed participants' latency to attempt to cash-in the balloon. Given that the dependent variable in this analysis is a type of reaction time, the data are not normally distributed, thus a log transformation was used to normalize the distribution. The best fitting multilevel model, as assessed by the Bayesian Information Criterion (BIC), included fixed effects of balloon thickness, trial within the current sub-block, and whether the previous balloon popped; the three-way interaction was dropped, but all two way interactions were kept. The random effects were intercept, balloon thickness, trial within the current sub-block, and whether the previous balloon popped.

The best fitting model predictions are presented in Figure 2.7. For the thickest balloon (solid line in figure), participants waited longer than for the neutral balloons (dashed line in figure), and waited the least for the thinnest balloons (dotted line in figure), $F(1, 12.3) = 551.02$, $p < .0001$. If the previous balloon popped (right panel of Figure 2.7), participants would cash-in the balloon earlier (700 ms on average) than if the previous balloon was cashed-in (left panel of Figure 2.7) (1168 ms on average), $F(1, 11.95) = 8.74$, $p = .0121$. There was an interaction between trial in sub-block and balloon thickness, $F(1, 3305) = 85.59$, $p < .0001$. Participants learned to respond slower before cashing-in the thickest balloons later in a sub-block, but learned to respond quicker before cashing-in the neutral and thinnest balloons ($p < .05$). As indicated in the figure, there was a significant interaction between whether the previous balloon popped and balloon thickness, $F(1, 3319) = 14.63$, $p < .0001$. If the previous balloon popped, participants were less sensitive to the balloon thickness manipulation than if the previous balloon was cashed-in (i.e., the gap between the different balloon thicknesses is smaller if the previous balloon popped than if the previous balloon was cashed-in). None of the other effects reached significance.

The second analysis of the behavioral data was aimed at understanding the variation in participant latencies. The dependent variable for this analysis was computed by taking the absolute value of the difference between the average latency in a sub-block for a given participant and the latency on the current trial. This resulted in the absolute deviation from average on a given trial. This dependent variable was used in the analysis of exploratory behavior here as it was in Frank, Doll, Oas-Terpstra, and Moreno (2009). Given that these deviation scores had a floor of zero and no ceiling, they were not normally distributed, and a log transformation was used to normalize the distribution. In the BART task, participants learn by exploring the range of possible latencies. This analysis assesses how participants explore (or exploit) the available latencies with larger deviations indicating more exploration. The best fitting model included fixed effects of balloon thickness, trial within the current sub-block, and whether the previous balloon popped. The random effects were intercept, balloon thickness, trial within the current sub-block, and whether the previous balloon popped.

As shown in Figure 2.8, participants tended to explore more for the thickest balloons (about 207 ms on average) than for the neutral balloons (155 ms on average), and tended to explore the least for the thinnest balloons (128 ms on average), $F(1, 12.43) = 7.75$, $p = .016$. This finding is not surprising given the broader range of latencies that lead to points for the thicker balloons compared to the thinner balloons (see Figure 2.3).

Participants tend to explore less, later in a sub-block, $F(1, 20.71) = 11.67$, $p = .003$. On average, participants tended to explore more if the previous balloon popped than if the previous balloon was cashed-in, $F(1, 13.03) = 9.16$, $p = .010$. As indicated in the figure, there was a three-way interaction, $F(1, 3312) = 9.01$, $p = .003$. Participants showed the greatest difference in exploration across balloon thicknesses after cashing-in a balloon late in a sub-block, and the second greatest difference after a balloon popped early in a sub-block. If the previous balloon popped late in a sub-block, participants showed the least difference in exploration across balloon thickness (all contrasts $p < .05$). None of the other effects reached significance.

The third analysis assessed the probability of the balloon popping on a given trial. Given the outcome variable was whether the balloon popped on any given trial, a binomial outcome, this analysis was a multilevel logistic regression. The best fitting model included fixed effects of balloon thickness, trial within the current sub-block, and whether the previous balloon popped; the three-way interaction was dropped from the analysis. The random effects were intercept, balloon thickness, trial within the current sub-block, and whether the previous balloon popped. On average participants tended to pop the balloon on 33% of trials. As shown in Figure 2.9, participants tended to pop the balloon most often for the thinnest balloons and least often for the thickest balloons, $Z = 9.34$, $p < .0001$. As participants gained experience in a sub-block, they tended to pop the balloon less often, $Z = -3.16$, $p = .0002$. There was a significant interaction between balloon thickness and trial in sub-block, $Z = -3.36$, $p = .0008$. Depending on the balloon thickness, participants either learned to pop the balloon less often (for the thinnest and the neutral balloons) or more often (for the thickest balloons) later in a sub-block. For the thinnest balloons, participants learned to pop the balloon less often, later in a sub-block, but for the thickest balloons participants learned to pop the balloon more often later in a sub-block ($p < .05$). Neither the main effect of whether the previous balloon popped nor the interactions including the variable were significant, but the model including the variable outperformed the model in which it was excluded.

EEG analysis

The EEG analysis involved assessing the magnitude of the FRN. The FRN is operationalized as the difference between popped and cashed-in trials in voltage (in microvolts) at fronto-central electrodes 250-300 ms after a feedback stimulus was displayed on screen. A main effect of whether the balloon popped would indicate a consistent FRN across the experiment. An interaction between whether the balloon popped and another variable would indicate that the size of the FRN changed as a function of the other variable.

The initial model included the effects of error magnitude and error probability on the amplitude of the FRN, but these were dropped from the best fitting model. The best fitting model included fixed effects of whether the balloon popped (cashed-in versus popped trials), and the balloon thickness. The random effects were intercept, whether the balloon popped, and the balloon thickness. Neither of the main effects were significant, but as indicated in Figure 2.10, the interaction between whether the balloon popped and balloon thickness was significant. The amplitude of the FRN was greatest for the thinnest balloons and smallest for the thickest balloons, $F(1, 3081) = 5.61$, $p = .018$. Figure 2.11 shows the waveform for cashed-in and popped balloons for each balloon thickness. According to planned contrasts, the difference in amplitude between popped and cashed-in trials was significant for the neutral and thinnest balloons ($p < .05$), but not the thickest balloons ($p > .05$).

Discussion

Before discussing the relevance of the findings of Experiment 1, it is important to discuss the findings of the behavioral results. The behavioral analyses were included to ensure that participants were behaving in a manner similar to the original BART. Participants were sensitive to the balloon thickness manipulation; participants responded slowest for the thickest balloons and fastest for the thinnest balloons, a conceptual replication of the behavioral findings from McCoy and Young (In preparation). In the original BART, participants tend to respond too quickly on a given trial, relative to optimal (Lejuez et al., 2002; Young, Webb, Rung, & McCoy, 2014). Participants were similarly sub-optimal in Experiment 1; responding earlier than optimal to cash-in the balloon. However, participants behaved more optimally later in a sub-block, indicating that participants were learning how to perform the task better. The pattern of behavioral results indicate that the changes made to the BART to make it compatible with EEG did not fundamentally alter behavior on the task.

In addition, participants appeared to explore more for thicker balloons than for thinner balloons, a result which replicated previous findings (McCoy & Young, In preparation). While it has been interpreted as participants exploring the alternatives, another potential explanation comes from the timing literature. Specifically, Scalar Timing theory (Gibbon, 1977; Gibbon, Church, & Meck, 1984; Kehoe, Olsen, Ludvig, & Sutton, 2009), this theory suggests that as participants or animals have to time an interval for a longer period of time, the variability in responding will increase in kind. Based on this interpretation, it makes sense that participants will “explore” more for thicker balloons, however, this does not explain the additional findings of the exploration data. Participants explored more following an error early in a sub-block. Also, participants explored more early on in a sub-block than they did later in a sub-block.

Contrary to the hypotheses of Experiment 1, the largest FRN occurred for the thinnest balloons. It was hypothesized that higher magnitude errors (i.e., errors associated with the thickest balloons) would result in a larger FRN. In fact, error magnitude was dropped from the model because of its poor performance as a predictor. Instead, the only predictors retained in the model were whether the balloon popped and balloon thickness. The inclusion of whether the balloon popped is not surprising because it is the basis of the FRN. The inclusion of balloon thickness, however, was surprising.

A likely explanation is that responding quickly for the thin balloons took more effort than responding slowly for the thick balloons. For extremely thin balloons (i.e., ones that popped in 100 ms), participants may have had difficulty responding quickly enough to cash-in the balloon. If this explanation is correct then there should be a physiological marker of effort that can be used to ascertain which balloons require the most effort. Beta band EEG, electrical activity at 15-30 Hz, has been associated with perceived mental effort (Howells, Stein, & Russell, 2010). In their study, Howells et al. (2010) found that when the demands of the task increased (i.e., became more difficult), participants said they expended more mental effort, and this was correlated with beta band power at electrodes over the left parietal lobe. Power in a given frequency band is directly associated with the amplitude of the EEG in that bandwidth across an epoch. EEG power can be thought of as amplitude squared per Hertz (V^2Hz).

To determine the perceived effort necessary to complete each balloon type, I measured the level of beta band power over the left parietal lobe for each balloon type. I then correlated the power in the beta band to the magnitude of the difference (in μV) between cashed-in and popped trials for each of the three balloon thicknesses (see Figure 2.12). Beta power was highest for the thinnest balloons, and lowest for the neutral balloons. This finding supports the supposition that the thin balloons required more effort than the thickest and neutral balloons. Furthermore, the correlation between beta power and FRN amplitude was high, but thin balloons having a much higher beta power than thick and neutral balloons primarily drove this finding. There was little difference between thick and neutral balloons.

The evidence presented seems to suggest a role of effort in the amplitude of the FRN. This result may seem common sense, but to date there is limited literature on the importance of engagement in the task and the magnitude of the FRN. Shephard, Jackson, and Groom (2014) looked at the performance differences between adults and children on a reinforcement learning task with reversal learning. Adults tended to perform better on the task than did children (as indicated by accuracy scores). Interestingly, an investigation of their results indicate a correlation between task difficulty and FRN amplitude ($R^2 = .40$) for children but not adults. This suggests when a task is difficult, the amplitude of the FRN will increase correspondingly, but when the task is considered easy, there may be no correlation between difficulty and FRN amplitude.

It does not take a leap of intuition to assume that more effort will be expended in a difficult task than an easy task. Thus, the correlation found by Shephard et al. (2014) for children may be explained by an increase in beta activity for the difficult task that is not occurring when the task is considered easy. When an error is made on a more effortful task, it may be more salient than when an error is made on an easier task. Previous research on the FRN has suggested that the FRN may encode the salience of the prediction error (Talmi, Atkinson, & El-Deredy, 2013). The findings of the current study seem to agree with this conclusion. Also, fMRI studies seem to suggest that the FRN is associated with mesencephalic dopamine projections to the dorsal anterior cingulate cortex (dACC), but since the publication of the RL-ERN theory, further evidence has distinguished between areas in the dACC associated with reinforcement learning and motivational salience (Hauser et al., 2014). Thus it seems plausible that mesencephalic dopamine projections to the dACC, an initial basis for the RL-ERN theory of the FRN, that are associated with the FRN may in fact be processing the motivational salience of errors.

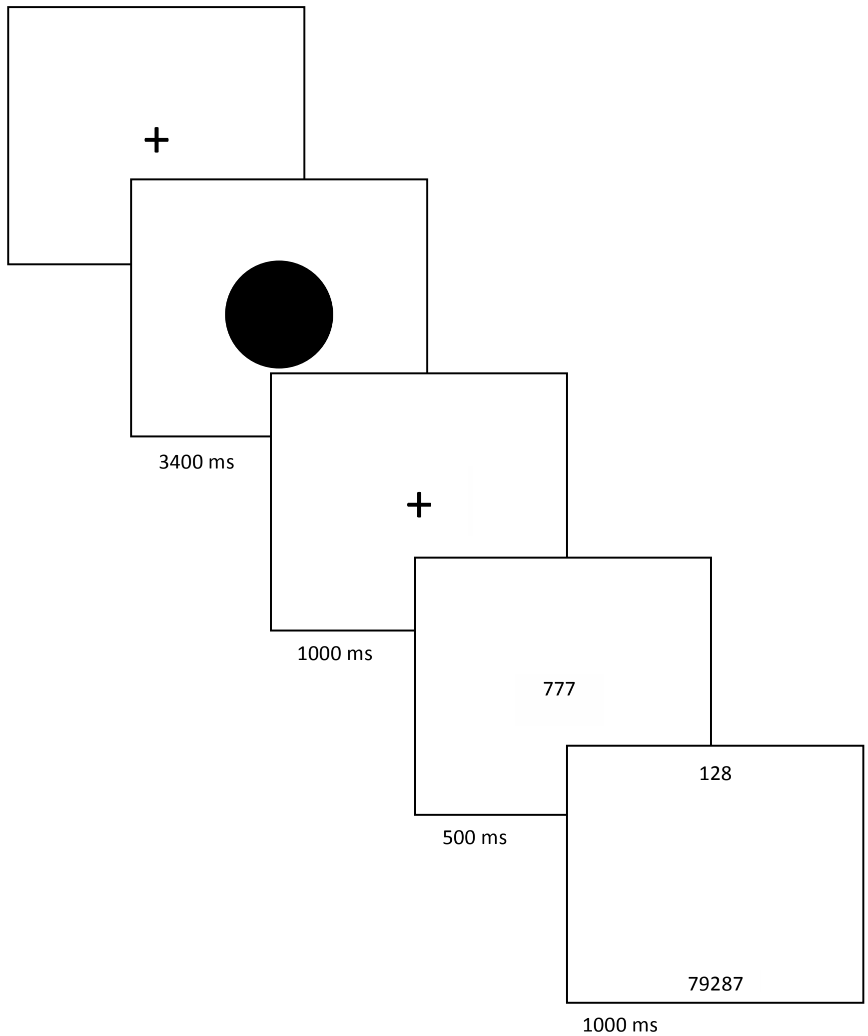


Figure 2.1. Trial progression. Each trial began with a fixation cross and participants had to press any key to start the balloon animation.

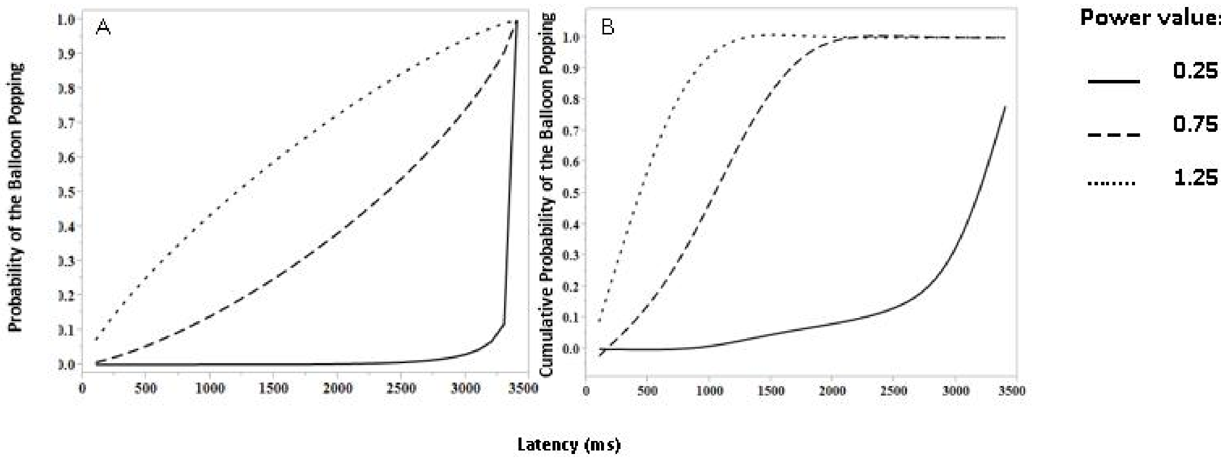


Figure 2.2. (A) Probability of the balloon popping as a function of latency as determined by a super-ellipsoid function (see Equation 3). (B) The cumulative probability of a balloon popping as a function of latency. Power values are represented by separate lines.

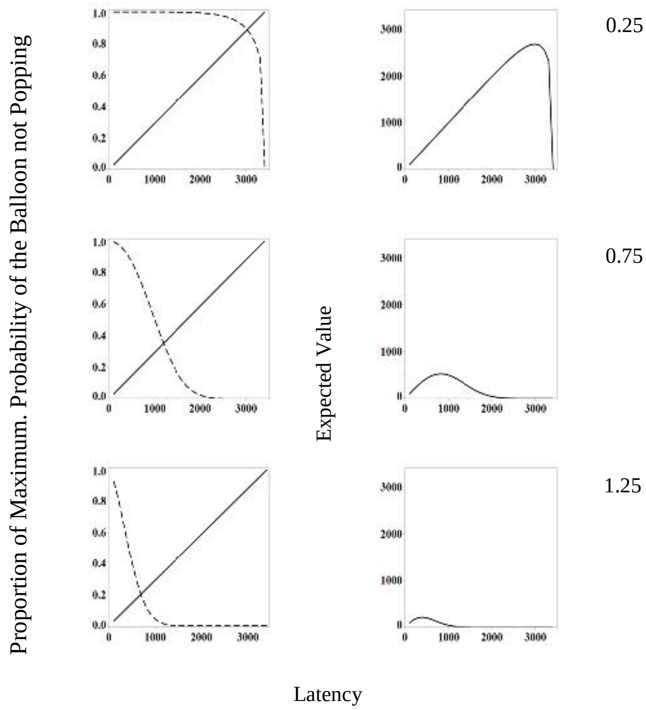


Figure 2.3. The left column displays the probability of the balloon not popping before a given latency (dashed line) and the proportion of points earned for cashing-in at a given latency (solid line). The Right column displays the expected value associated with cashing-in at a given latency. Power values are listed to the right of the figure.

Block 1			Block 2			Block 3		
0.75	1.25	0.25	1.25	0.25	0.75	0.25	0.75	1.25
Sub-Block 1		Sub-Block 3		Sub-Block 5		Sub-Block 7		Sub-Block 9

Figure 2.4. Experiment set-up showing block and sub-block structure for Experiment 1. Between each block there was a rest break. Each block contained 150 trials and each sub-block consisted of 50 trials.

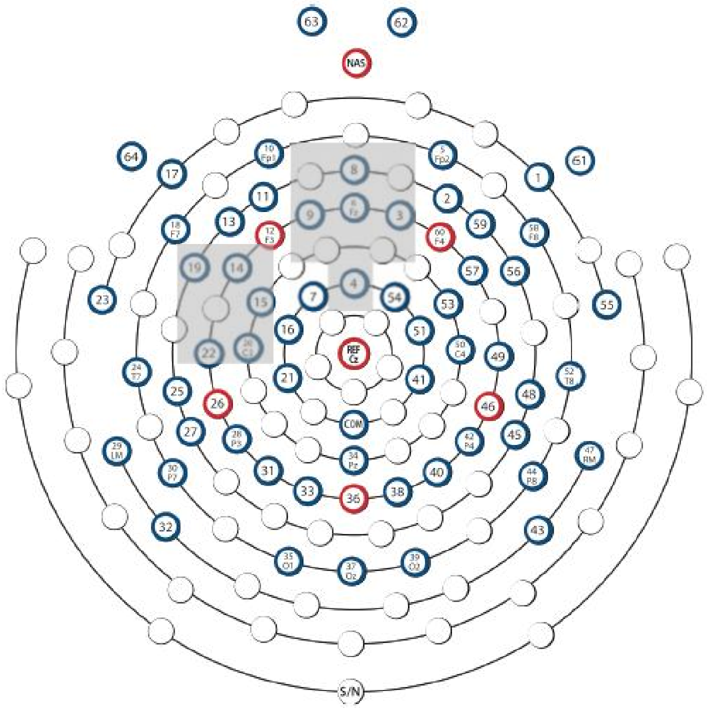


Figure 2.5. Diagram of 64 channel net electrode layout. Labeled electrodes based on traditional 10/20 electrode layout. The front of the head is at the top of the figure. The shaded region over fronto-central electrodes indicates the electrodes used in determining the FRN. The shaded region over left-parietal electrodes indicates the electrodes used in determining beta power.

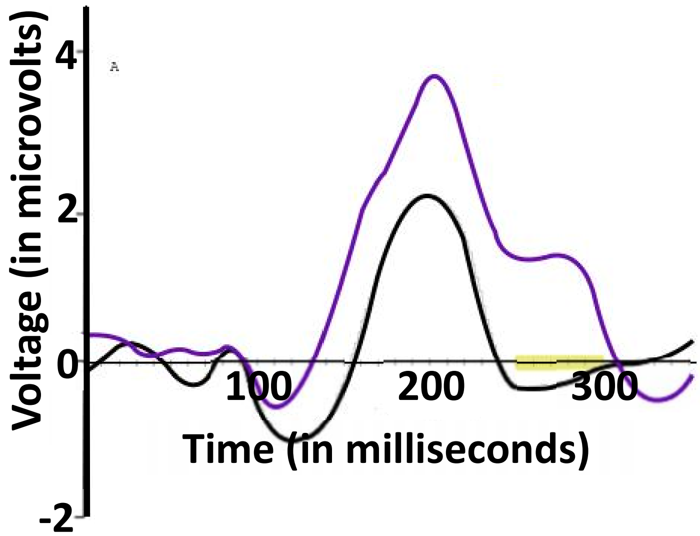


Figure 2.6. Voltage reading (in µV) on trials in which the balloon popped (Black line) and trials in which the balloon was cashed-in (Purple line). Trials segmented to feedback onset with a 200 ms baseline period prior to feedback onset (not shown). Segment ends 800 ms post feedback onset (truncated in figure). Shaded region indicates approximate time region of interest 250-300 ms post feedback onset.

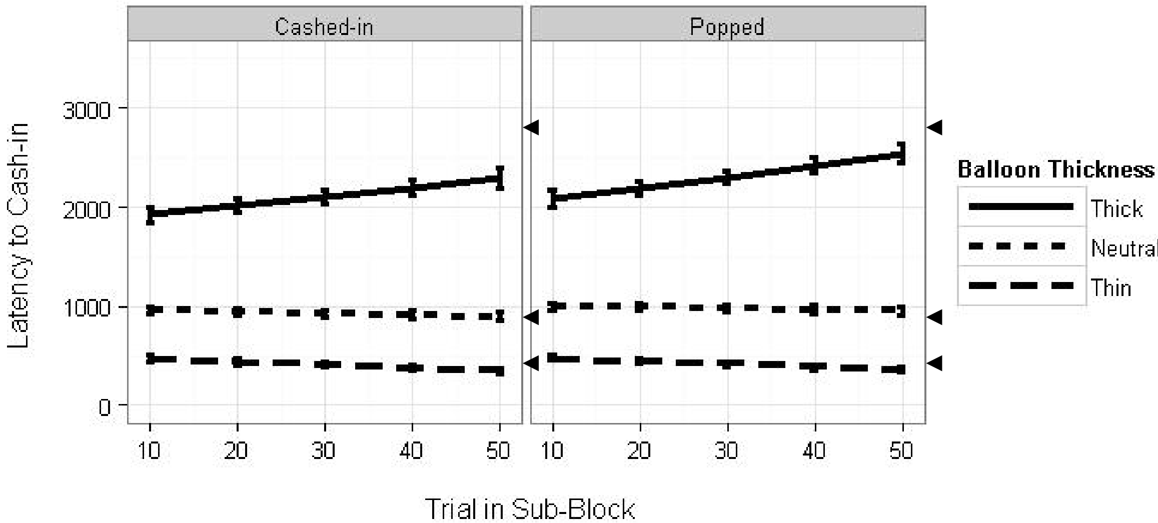


Figure 2.7. Effect of trial in a sub-block, whether the previous balloon popped, and balloon thickness on latency to cash-in the balloon on the current trial. The left panel of the figure shows latencies when participants cashed-in the previous balloon, while the right panel of the figure shows latencies when participants popped the previous balloon. Different balloon thicknesses are represented by separate lines. The arrows at the right side of each panel indicate the optimal latency to cash-in the balloon for each balloon thickness. The highest arrow (at 2800 ms) represents optimal behavior for the thickest balloons, and the lowest arrow (at 400 ms) represents optimal behavior for the thinnest balloons.

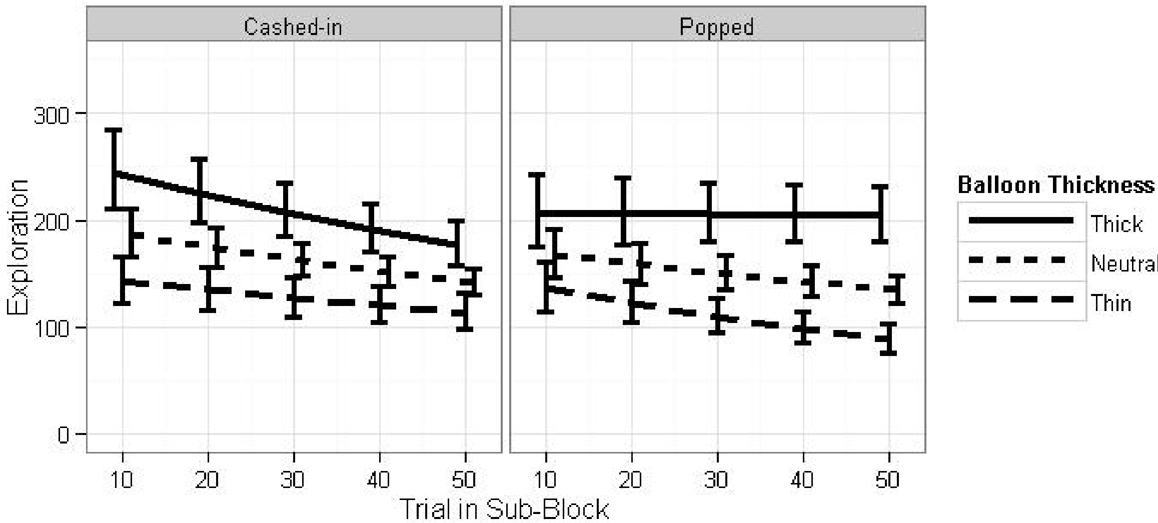


Figure 2.8. The effect of trial in sub-block, whether the previous balloon popped and balloon thickness on exploratory behavior. The left panel of the figure shows exploratory behavior when participants cashed-in the previous balloon, while the right panel of the figure shows exploratory behavior when participants popped the previous balloon. Different balloon thicknesses are represented by separate lines.

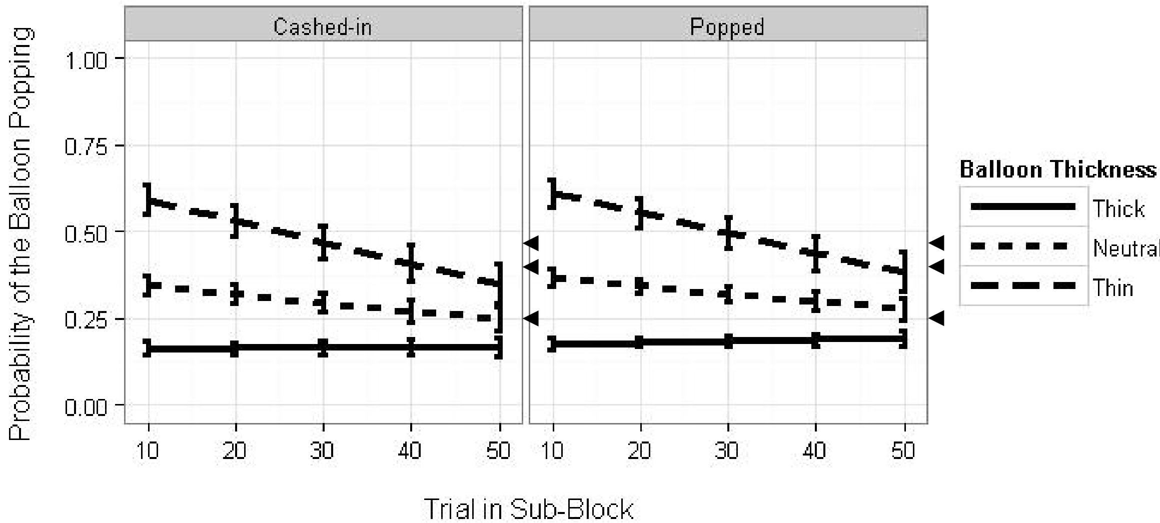


Figure 2.9. The effect of trial in sub-block, whether the previous balloon popped, and balloon thickness on the probability of the participant popping the balloon on a given trial. The left panel of the figure shows the probability of participants popping the balloon when they cashed-in the previous balloon, while the right panel of the figure shows the probability of participants popping the balloon when they popped the previous balloon. Different balloon thicknesses are represented by separate lines. The arrows at the right side of each panel indicate the optimal probability of popping a balloon for each balloon thickness based on the optimal wait time. The highest arrow (at .46) represents optimal behavior for the thinnest balloons, and the lowest arrow (at .25) represents optimal behavior for the thickest balloons.

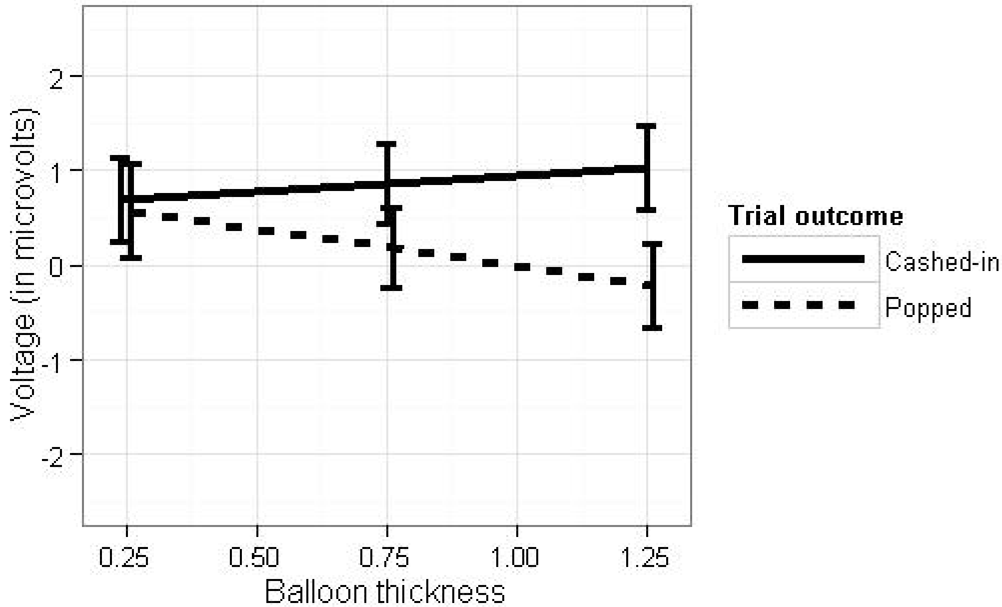
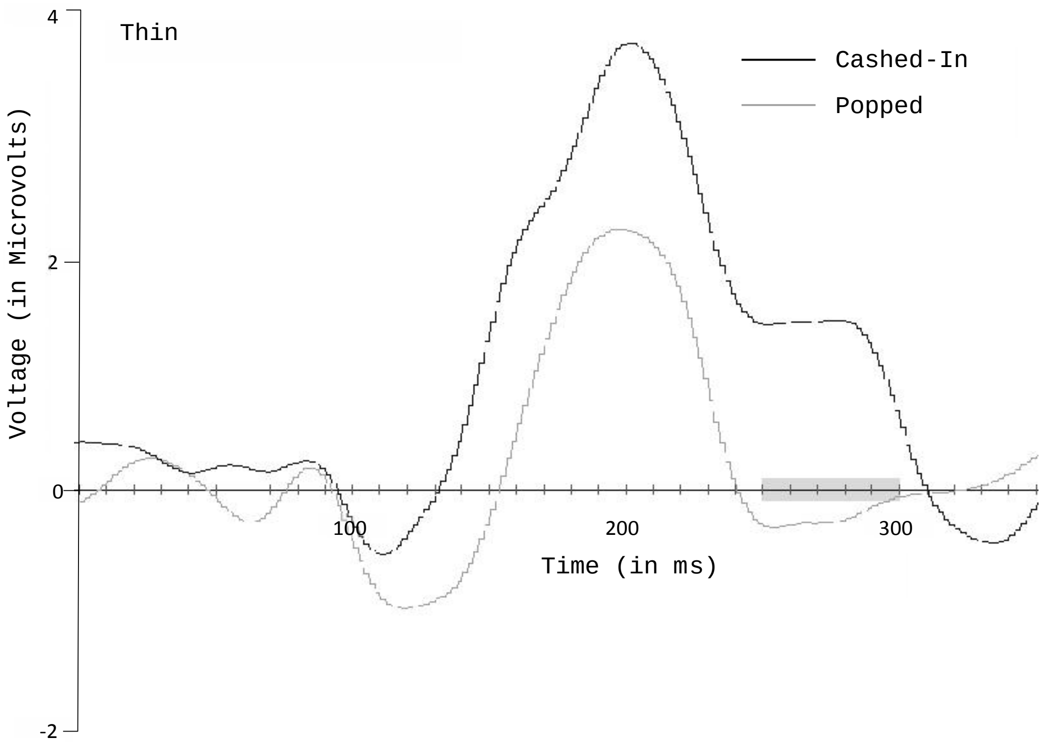
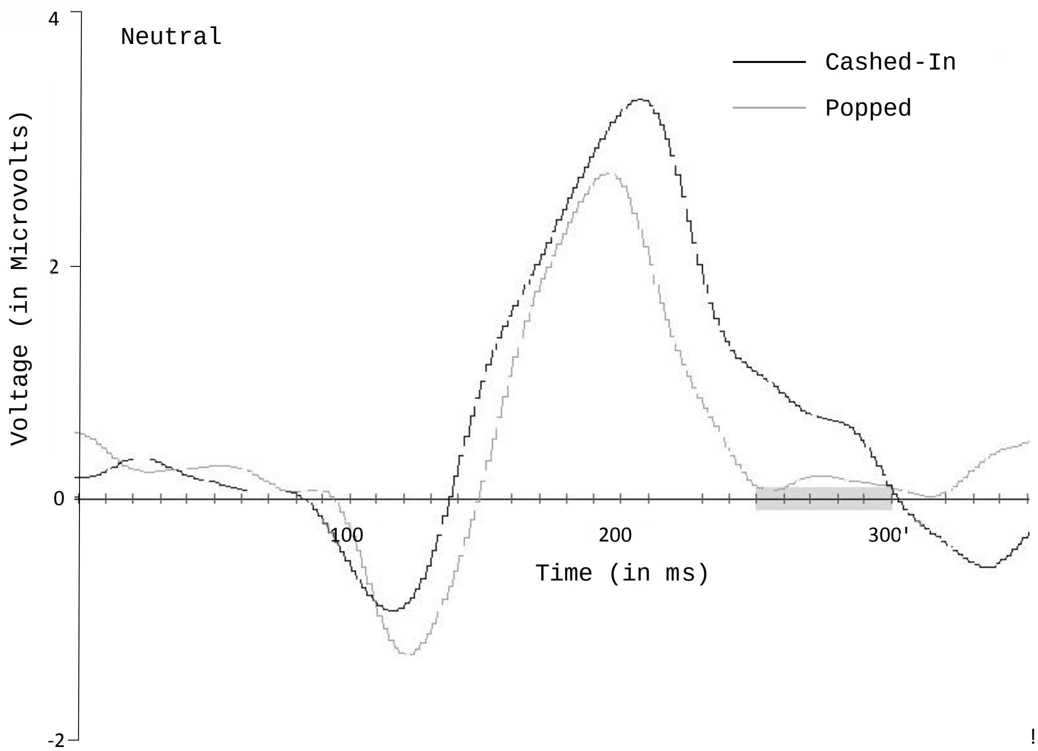


Figure 2.10. Voltage (in μV) at each balloon thickness. The separate lines represent different trial outcomes. The amplitude of the FRN is the difference between the two trial outcomes when cashed-in balloons result in higher voltage than popped balloons.



Thin
Cashed-In
Popped
Voltage (in Microvolts)
4
2
0
-2
100
200
300
Time (in ms)



Neutral
Cashed-In
Popped
Voltage (in Microvolts)
4
2
0
-2
100
200
300'
Time (in ms)

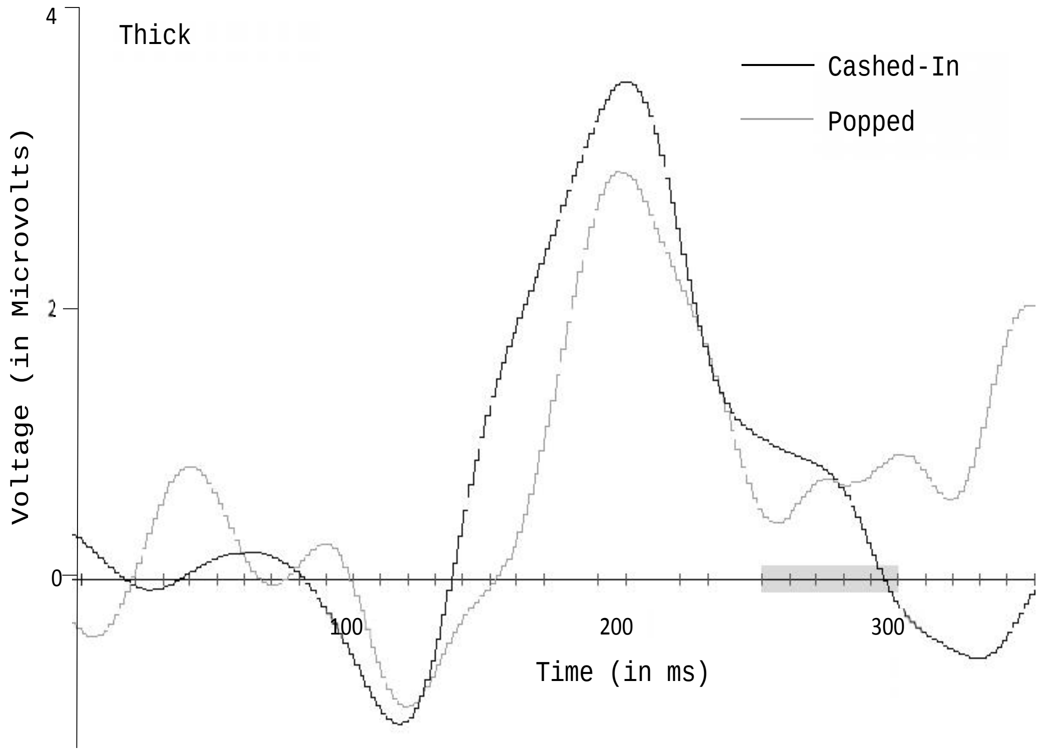


Figure 2.11. Averaged EEG waveforms for each balloon thickness. The darker lines represent cashed-in balloons and the lighter lines represent popped balloons.

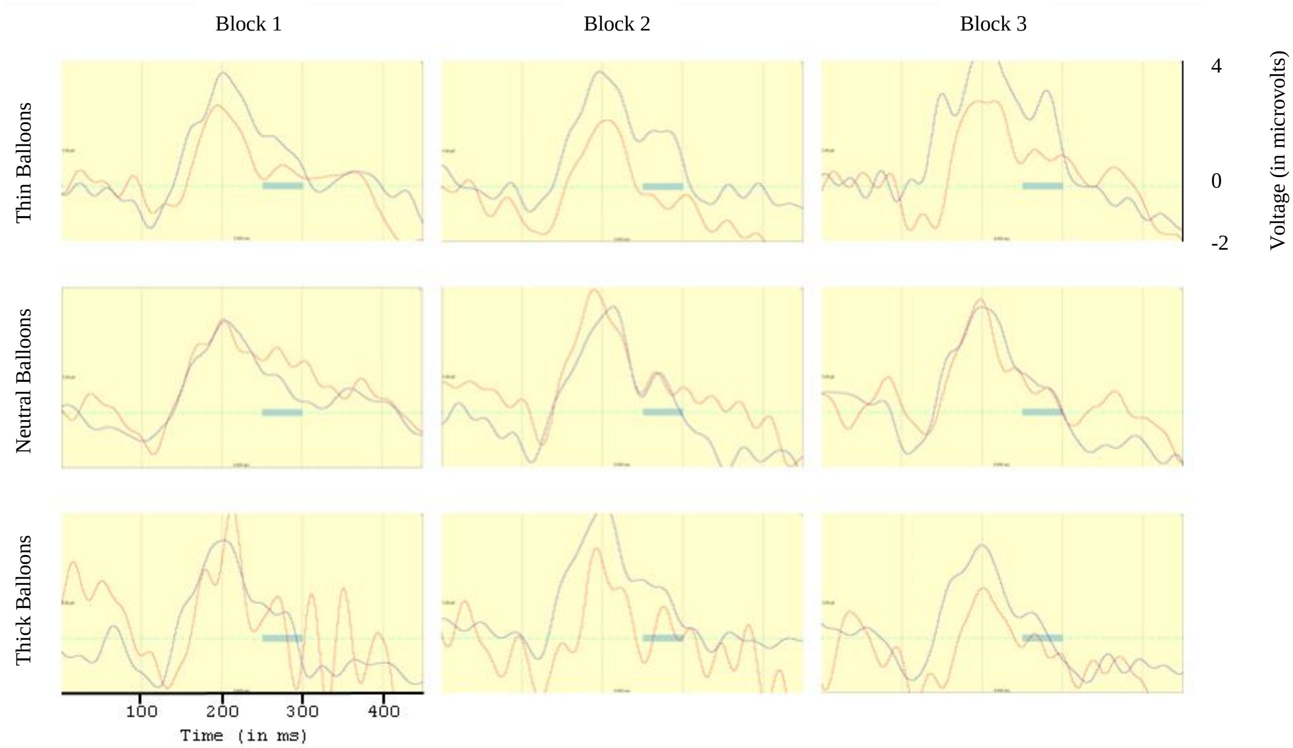


Figure 2.12. Averaged EEG waveforms for each balloon thickness during each block. The blue lines represent trials where the balloon was cashed-in and the red lines represent trials where the balloon popped. The shaded region is the time region associated with the FRN.

Chapter 3 - Experiment 2

In Experiment 2, I again tested the effect of error magnitude and error probability on the amplitude of the FRN, but focused on trials where the participant cashed-in the balloon. Experiment 2 had two within-subjects conditions. The masked points possible condition was a replication of Experiment 1, in that participants experience the same feedback as in Experiment 1. In the unmasked points possible condition, participants were presented with information on the number of points they could have earned (i.e., The number of points they could have earned for waiting until just before the balloon popped to cash-in) on the balloon before they were shown how many points they received. Thus in the unmasked condition, participants could experience an error signal even on trials where the balloon was cashed-in and points were earned. It was hypothesized that participants would show an FRN on cashed-in trials in the unmasked condition, but not in the masked condition. It is also hypothesized that, if the RL-ERN theory is accurate, the amplitude of the FRN in the unmasked condition should vary based on the magnitude of the error. However, if the amplitude of the FRN does not vary based on the magnitude of the error, this would be further evidence against the RL-ERN theory.

Methods

Participants

Participants were 20 undergraduates recruited from the General Psychology participant pool. Participants obtained course credit for their voluntary participation.

Stimuli and procedure

The stimuli used in Experiment 2 were similar to those used in Experiment 1 with the following exceptions. Participants experienced three feedback screens instead of two. The additional feedback screen occurred before the primary feedback screen, and lasted for 500 ms, and was followed by a 1000 ms buffer screen (see Figure 3.1). This additional buffer screen played a similar role to the initial buffer screen in that it attenuated the evoked potentials associated with the feedback screen. In the unmasked condition, the first feedback screen centrally displayed the number of points the participant could have earned. In the masked condition, the first feedback screen had the number of points the participant could have earned centrally displayed, but the number was masked (“8888” overlaid with “888”). The use of the mask allowed both conditions to have the same amount of visual stimulation (important for avoiding confounds) while manipulating the amount of information available to the participant. Participants were randomly assigned to one of two orders. Half of the participants experienced the unmasked condition followed by the masked condition while the other half experienced the opposite order.

To accommodate the additional condition, participants experienced 6 blocks of 75 balloons each (25 balloons in each sub-block). Participants experienced 3 blocks in each condition. This resulted in the same number of balloons experienced as in Experiment 1. The revised block/sub-block structure is presented in Figure 3.2.

Results

On average participants had 408 good EEG trials, roughly 91% of the 450 total trials experienced. Of the twenty participants, one had excessive artifacts in their data, resulting in fewer than 75% usable trials; this participant was excluded. There were 19 participants with useable data.

Behavioral analyses

The behavioral analyses ensured that behavior on this task was similar to behavior in Experiment 1 and ensured that participants were sensitive to the balloon thickness manipulation. Three analyses were run on separate portions of the behavioral data. The first analysis assessed how long participants waited to cash-in the balloon. The second analysis assessed how much exploration participants engaged in during the task. The third analysis assessed how often participants popped the balloon.

The first behavioral analysis assessed participants' latency to attempt to cash-in the balloon. The best fitting multilevel model included fixed effects of balloon thickness, trial within the current sub-block, whether the previous balloon popped, and whether the participant was completing the masked or the unmasked portion of the experiment; the three-way and four way interactions were dropped, but all two way interactions were kept. The random effects were intercept, balloon thickness, trial within the current sub-block, whether the previous balloon popped, and condition.

The best fitting model predictions for this analysis are presented in Figure 3.3. For the thickest balloon (solid line in figure), participants waited longer than for the neutral balloons (dashed line in figure), and waited the least for the thinnest balloons (dotted line in figure), $F(1, 17.14) = 438.24$, $p < .0001$. If the previous balloon popped (right panel of Figure 3.3), participants would cash-in the balloon earlier (634 ms on average) than if the previous balloon was cashed-in (left panel of Figure 3.3) (1071 ms on average), $F(1, 17.23) = 8.03$, $p = .011$. There was a significant interaction between balloon thickness and trial in sub-block, $F(1, 6951) =$

111.04, $p < .0001$. Participants learned to cash-in later for the thickest balloons later in a sub-block, but learned to cash-in sooner for the neutral and thinnest balloons (all contrasts were significant, $p < .05$). As indicated in the figure, there was a significant interaction between whether the previous balloon popped and balloon thickness, $F(1, 3319) = 14.63$, $p < .0001$. If the previous balloon popped, participants were less sensitive to the balloon thickness manipulation than if the previous balloon was cashed-in, $p < .05$. There was a significant interaction between condition and balloon thickness, $F(1, 6944) = 5.35$, $p = .021$. Participants were more sensitive to the balloon thickness manipulation in the unmasked condition than the masked condition, $p < .05$. Also, an examination of Figure 3.3 shows that in the unmasked condition participants behaved more optimally (lines more closely approach the optimality arrows) than in the masked condition. This result is not surprising given that participants have more information to inform their behavior in the unmasked condition than in the masked condition. None of the other effects reached significance. Figure 3.4 shows the effect across the two conditions based on the order of exposure to the conditions.

The second analysis of the behavioral data was aimed at understanding the variation in participant latencies. The best fitting model was identical to the best fitting model in Experiment 1 for the same analysis. The model initially included the condition variable, but it was dropped from the best fitting model. As shown in Figure 3.5, participants tended to explore more for the thickest balloons (about 279 ms on average) than for the neutral balloons (216 ms on average), and tended to explore the least for the thinnest balloons (167 ms on average), $F(1, 20) = 48.03$, $p < .0001$.

Participants tended to explore less later in a sub-block, $F(1, 29.72) = 15.06$, $p = .0005$. On average, participants tended to explore more if the previous balloon popped than if the previous balloon was cashed-in, $F(1, 7008) = 75.13$, $p < .0001$. As indicated in the figure, there was a three-way interaction, $F(1, 3312) = 9.01$, $p = .003$. If the previous balloon was cashed-in later in a sub-block, the difference in exploration across balloon thicknesses was largest, the next largest difference occurred if the previous balloon popped, early in a sub-block, the smallest difference was found for trials where the previous balloon popped late, and where the previous balloon was cashed-in early (all contrasts significant, $p < .05$).

The third analysis assessed the probability of the balloon popping on a given trial. The best fitting model included fixed effects of balloon thickness, trial within the current sub-block, and whether the previous balloon popped, the three-way analysis was dropped from the analysis. The random effects were intercept, balloon thickness, trial within the current sub-block, and whether the previous balloon popped. The condition variable was included in the initial model, but it was dropped from the best fitting model.

On average participants tended to pop the balloon on 33% of trials. As shown in Figure 3.6, as expected participants tended to pop the balloon most often for the thinnest balloons and least often for the thickest balloons, $Z = 13.86$, $p < .0001$. There was a significant interaction between balloon thickness and trial in sub-block, $Z =-5.88$, $p < .0001$. Depending on the balloon thickness, participants either learned to pop the balloon less often (for the thinnest and the neutral balloons) or more often (for the thickest balloons) later in a sub-block. The trial slope was most positive for the thickest balloons and most negative for the thinnest balloons (contrasts between all three balloon thicknesses were significant, $p < .05$). There was a significant interaction between whether the previous balloon popped and balloon thickness, $Z = -2.97$, p - .003. If the previous balloon popped, there was a larger difference in probability of popping the balloon between the thickest and thinnest balloons than when the previous balloon was cashed-in, p <

.05. There was a significant interaction between trial in sub-block and whether the previous balloon popped, $Z = 2.55$, $p = .011$. There was also a significant three-way interaction, $Z = -3.09$, $p = .002$. All trial slopes were significantly different from the other trial slopes ($p < .05$), The most negative slope occurred when the previous balloon popped for the thinnest balloons. All of the slopes were negative with the exception of the 0.25 slope when the previous balloon was cashed-in. Participants were more likely to pop the balloon early in a sub-block than later in a sub-block for all of the balloon thicknesses if the previous balloon popped, but if the previous balloon was cashed-in, participants were more likely to pop the balloon early in a sub-block only for the thinnest and the neutral balloons, but not for the thickest balloons. None of the other effects approached significance.

EEG analyses

Analysis of Correct versus Incorrect trials

The first EEG analysis involved assessing the magnitude of the FRN across cashed-in and popped balloons. The FRN was operationalized as the difference between popped and cashed-in trials in voltage (in microvolts) at fronto-central electrodes 250-300 ms after the feedback stimulus indicating the number of points earned on the trial was displayed on screen. A main effect of whether the balloon popped would indicate a consistent FRN across the experiment. An interaction between whether the balloon popped and another variable would indicate that the size of the FRN changed as a function of the other variable.

The initial model included effects of balloon thickness, condition, order, trial in sub-block, whether the previous balloon popped, and whether the current balloon popped, The best fitting model included fixed effects of whether the current balloon popped and whether the previous balloon popped, and random effects of intercept, whether the previous balloon popped and whether the current balloon popped. As shown in Figure 3.7, the task elicited an FRN (i.e., the voltage over fronto-central sites was lower when feedback indicated the balloon popped), $F(1, 18.47) = 9.38$, $p = .007$ (main effect). There was an interaction between whether the previous balloon popped and whether the current balloon popped such that the amplitude of the FRN was maximal when the previous balloon popped, $F(1, 5405) = 7.78$, $p = .005$. In fact, contrasts indicate that a significant FRN only occurred on the current trial if the previous balloon popped. The main effect of whether the previous balloon popped was not significant ($F < 1.0$). Figure 3.8 shows the wave forms similar to Experiment 1 for each balloon thickness and for whether the balloon was cashed-in or popped.

Analysis of Correct Trials only

In order to determine whether an FRN would occur on trials where participants received points, a multilevel model was run on trials where the participant cashed-in the balloon. For this analysis, an overt error did not occur because participants received points based on how long they waited to respond, but in the unmasked condition, participants were shown how many points they could have earned on each trial. Because participants saw the number of points possible, any deviation from the total possible could be viewed as an error and could elicit an FRN.

For the purpose of this analysis, the FRN is still operationalized as the magnitude of the difference in voltage between correct and incorrect trials, but the definition of what constitutes an incorrect trial is redefined; the definition of error is not categorical, but continuous. Any deviation from the number of points possible constitutes an error, thus there are degrees of “incorrectness” such that a small deviation should be viewed as less of an error than a large deviation.

The best fitting model for this analysis included fixed effects of condition and magnitude of the deviation from the number of points possible. As indicated by Figure 3.9, there was a significant interaction between condition and deviation from the number of points possible, $F(1, 4733) = 4.35$, $p = .037$. In the masked condition, in which the participant did not receive information about whether an error occurred on a correct trial (i.e., cashing-in too early) nor the magnitude of said error, there was no change in the voltage across the range of deviations from points possible (Slope = .0001, $p > .05$). However, in the unmasked condition, in which participants did have information about the magnitude of the error involved, there was a change in voltage across the range of deviations such that voltage decreased from small deviations from points possible to large deviations from points possible (indicating an increase in the amplitude of the FRN across the range), slope = -.0004, $p < .05$.

Discussion

The behavioral results of Experiment 2 replicate the findings of Experiment 1. Participants were sensitive to the balloon thickness manipulation, and tended to explore more early on in a sub-block than later in a sub-block. Also, participants tended to pop the balloon at a similar rate in both experiments.

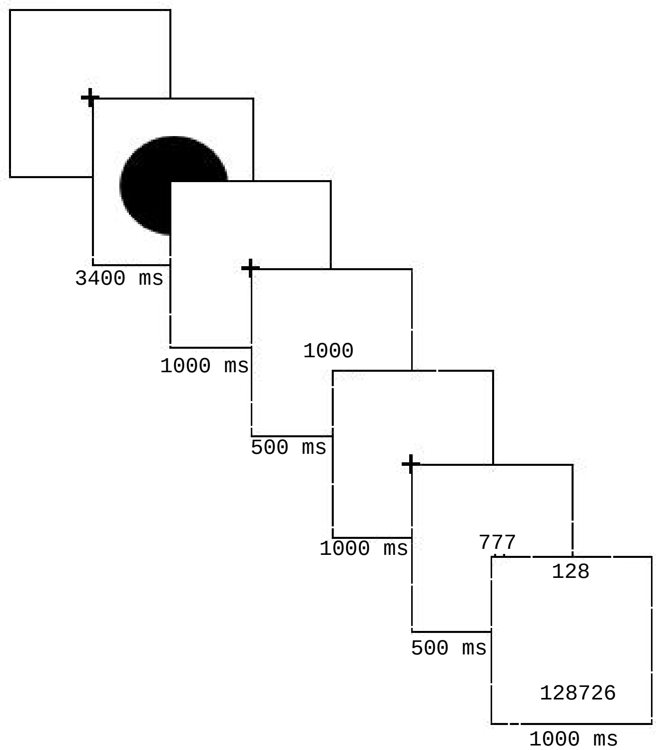
The FRN analysis on the popped versus cashed-in trials did not replicate in Experiment 2. In Experiment 2, the best fitting model for the traditional FRN included whether the previous balloon popped and the outcome on the current trial. In this case, errors that were preceded by an error resulted in an FRN whereas errors that were not preceded by an error did not result in an FRN. This difference is interesting in that it may indicate that the occurrence of multiple errors in a row may be more important for changing behavior than a single error. Another potential explanation is that duplicate errors are a rare event. Previous research on the amplitude of the

FRN has indicated that the amplitude of the FRN is greater following rare errors than following common errors (San Martin et al., 2010). There were a total of 2,431 popped balloons across the entire experiment for all of the participants. Of those popped balloons, 1053 (or 43%) occurred after the previous balloon was popped. This result suggests that it is unlikely that the difference in FRN amplitude based on whether the previous balloon popped was due to error probability. The failure to find an effect of error magnitude in Experiment 2 for the traditional FRN analysis is further evidence against the RL-ERN theory, but the results seem disparate with the results of Experiment 1.

Another finding in Experiment 1 was that Beta power was positively associated with the amplitude of the FRN, but this finding was not replicated in Experiment 2 (data not shown). In Experiment 1, beta power was used as a possible explanation for why the results were disparate with the RL-ERN theory, but failure to replicate this finding in Experiment 2 calls this interpretation into question. However, beta power was still highest for the thinnest balloons in Experiment 2, which does replicate a portion of the finding from Experiment 1. This indicates that the amount of effort necessary to perform well on the thinnest balloons is higher than the effort necessary to perform well on the neutral and thickest balloons.

One of the most interesting findings of Experiment 2 was the evidence that an FRN occurs on cashed-in trials. One of the primary goals of Experiment 2 was to assess whether an FRN would occur when participants received points, but failed to maximize the number of points they could have earned. Participants were sensitive to the number of points they could have earned in the unmasked condition, but not in masked condition. Furthermore, participants exhibited a larger FRN for larger errors than for smaller errors in the unmasked condition, a finding in direct opposition to the findings of Hajcak et al. (2006). In their study, participants were tasked with choosing between four doors. After choosing a door, participants were told whether they earned 25 cents, earned 5 cents, lost 25 cents, or lost 5 cents. There was a significant FRN (voltage was more negative following loss trials than gain trials), but there was not a significant main effect of magnitude nor was the interaction significant. There is a methodological difference between these two studies. Whereas Hajcak et al. (2006) used four (five in Experiment 2) separate categorical outcomes, I used a continuous outcome ranging from a difference of zero points to a difference of 3400 points. This method allowed for a more complete understanding of the effect of error magnitude on the amplitude of the FRN. Another potential cause for different findings across the two studies is that in their study the large error was a 25 cent loss which may not be sufficient for participants to react differently. In my study, participants can experience very large errors (i.e., 3000 points less than they could have earned) or small errors (i.e., 100 points less than they could have earned), this means that large errors are truly much larger in magnitude than small errors.

A


3400 ms
1000 ms
1000
500 ms
1000 ms
777
128
500 ms
128726
1000 ms

B

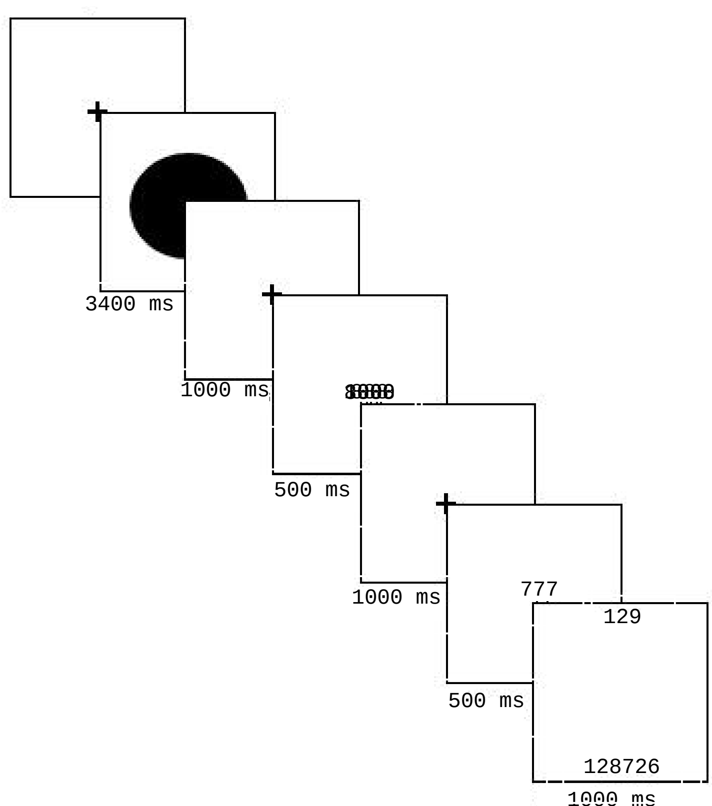

3400 ms
1000 ms
500 ms
1000 ms
777
129
500 ms
128726
1000 ms

Figure 3.1. Trial progression. Each trial began with a fixation cross and participants had to press any key to start the balloon animation. (A) The trial progression for unmasked trials. (B) The trial progression for masked trials.

Block 1			Block 2			Block 3			Block 4			Block 5			Block 6		
0.75	1.25	0.25	1.25	0.25	0.75	0.25	0.75	1.25	0.75	1.25	0.25	1.25	0.25	0.75	0.25	0.75	1.25
Sub-Block 1		Sub-Block 3		Sub-Block 5		Sub-Block 7		Sub-Block 9		Sub-Block 11		Sub-Block 13		Sub-Block 15		Sub-Block 17	

Figure 3.2. Experiment set-up showing block and sub-block structure. Between each block there was a rest break. Each block contained 75 trials and each sub-block consisted of 25 trials. After block 3, participants switched to the next condition (e.g., if the first condition experienced was the unmasked condition then after block 3 participants switched to the masked condition).

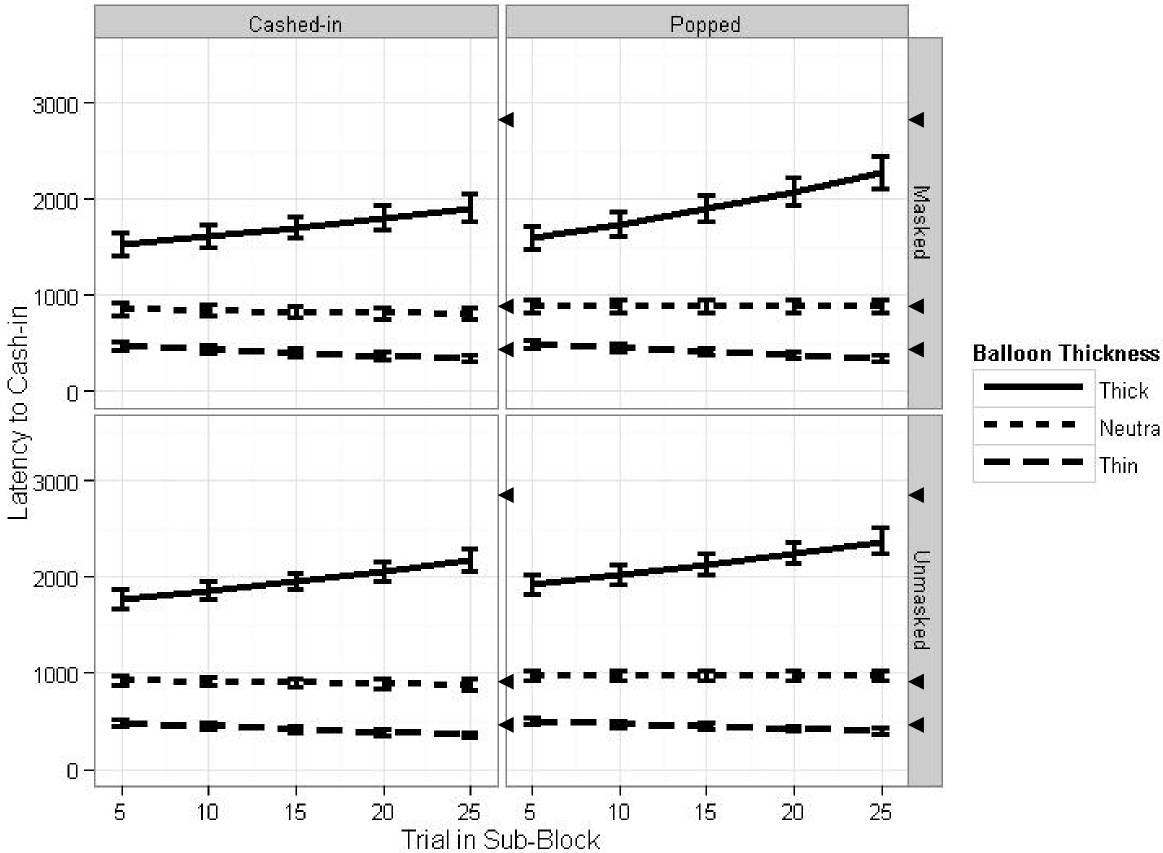


Figure 3.3. Effect of trial in a sub-block, whether the previous balloon popped, and balloon thickness on latency to cash-in the balloon on the current trial. The left panel of the figure shows latencies when participants cashed-in the previous balloon, whereas the right panel of the figure shows latencies when participants popped the previous balloon. The top row of the figure shows data for the masked condition, and the bottom row of the figure shows data for the unmasked condition. Different balloon thicknesses are represented by separate lines. The arrows at the right side of each panel indicate the optimal latency to cash-in the balloon for each balloon thickness. The highest arrow (at 2800 ms) represents optimal behavior for the thickest balloons, and the lowest arrow (at 400 ms) represents optimal behavior for the thinnest balloons.

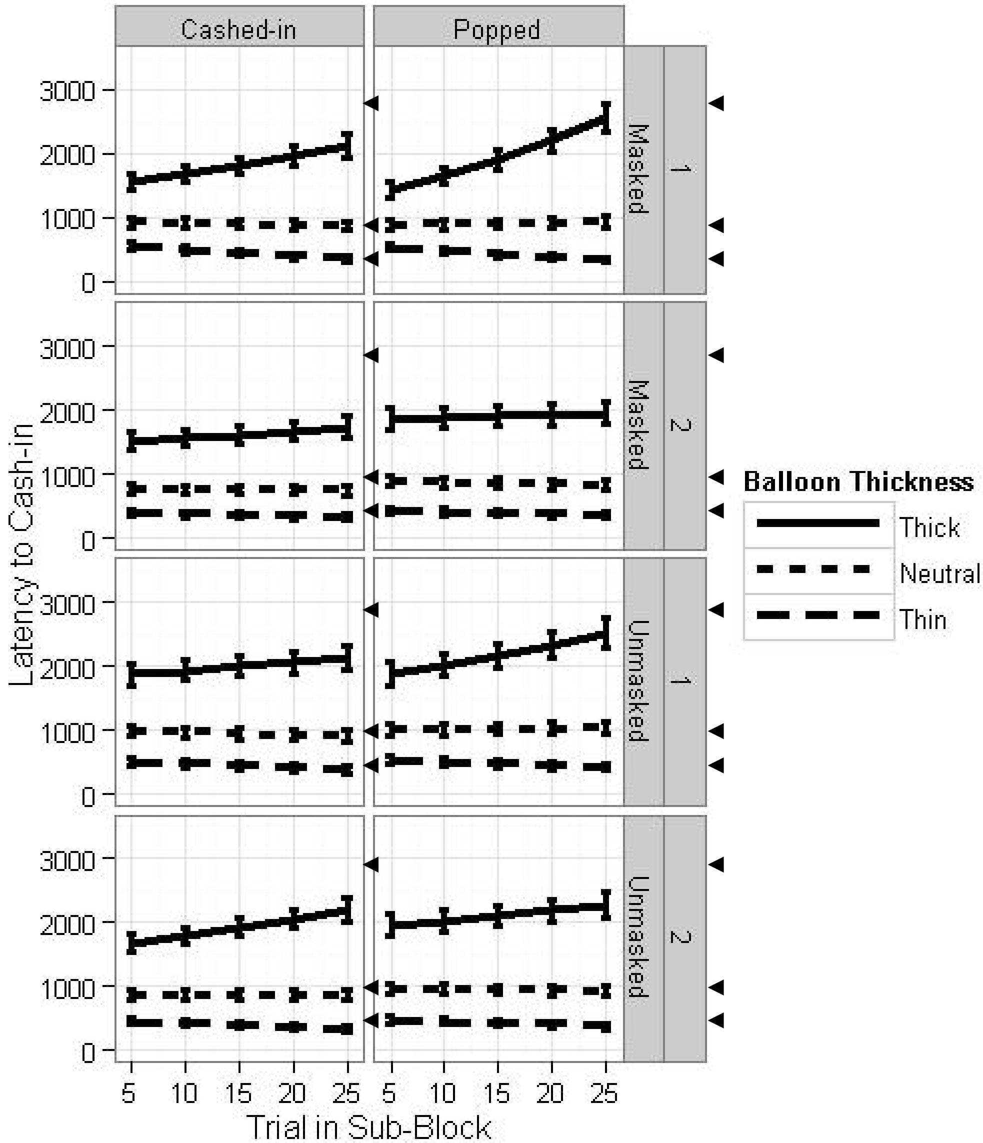


Figure 3.4. Effect of trial in sub-block, whether the previous balloon popped, balloon thickness, condition, and session on latency to cash-in the balloon. The panels on the left side of the figure show latencies when participants cashed-in the previous balloon, whereas the right panels of the figure show latencies when participants popped the previous balloon. The different rows of the figure represent the separate conditions and sessions of the experiment. Different balloon thicknesses are represented by separate lines. The arrows at the right side of each panel indicate the optimal latency to cash-in the balloon for each

balloon thickness. The highest arrow (at 2800 ms) represents optimal behavior for the thickest balloons, and the lowest arrow (at 400 ms) represents optimal behavior for the thinnest balloons.

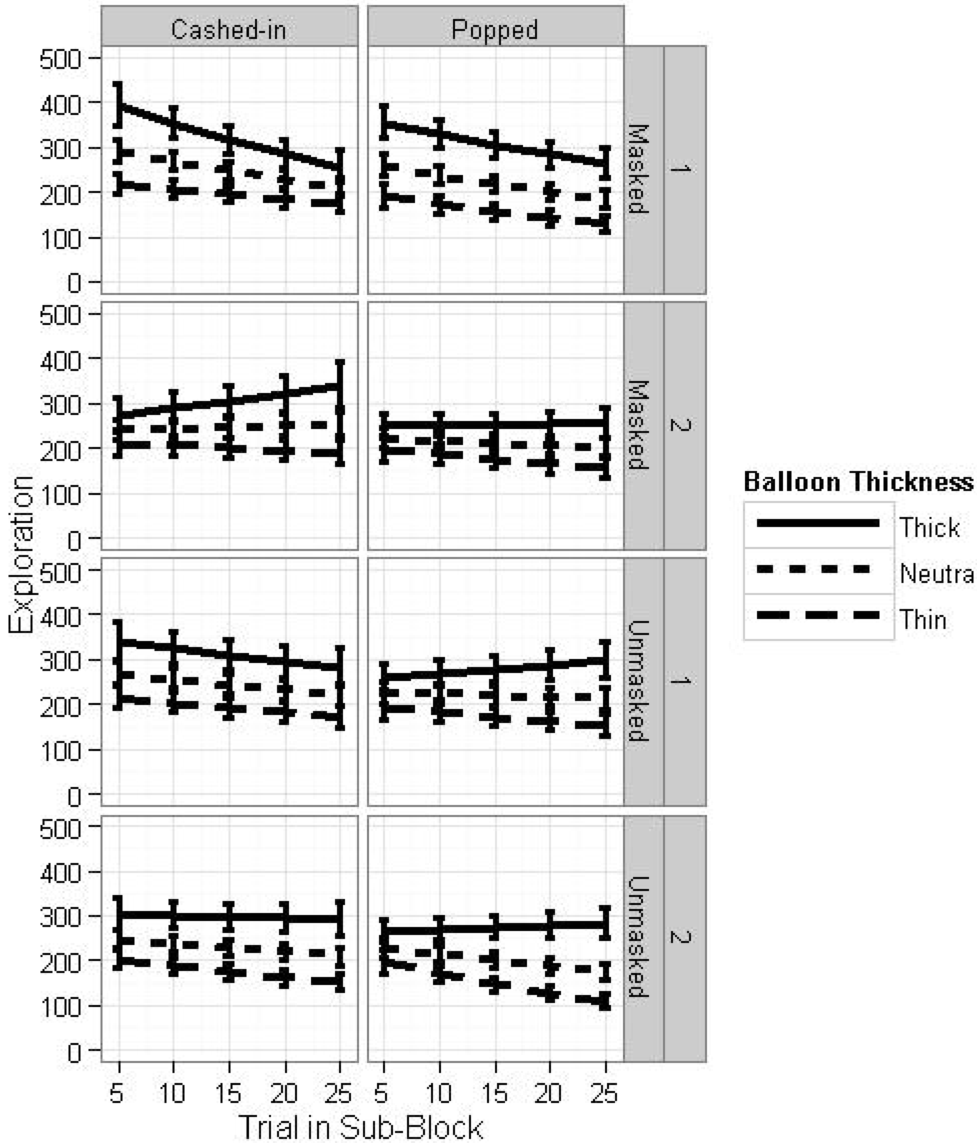


Figure 3.5. The effect of trial in sub-block, whether the previous balloon popped and balloon thickness on exploratory behavior. The left panels of the figure shows exploratory behavior when participants cashed-in the previous balloon, whereas the right panels of the figure shows exploratory behavior when participants popped the previous balloon. The different rows of the figure represent the separate conditions and sessions of the experiment. Different balloon thicknesses are represented by separate lines.

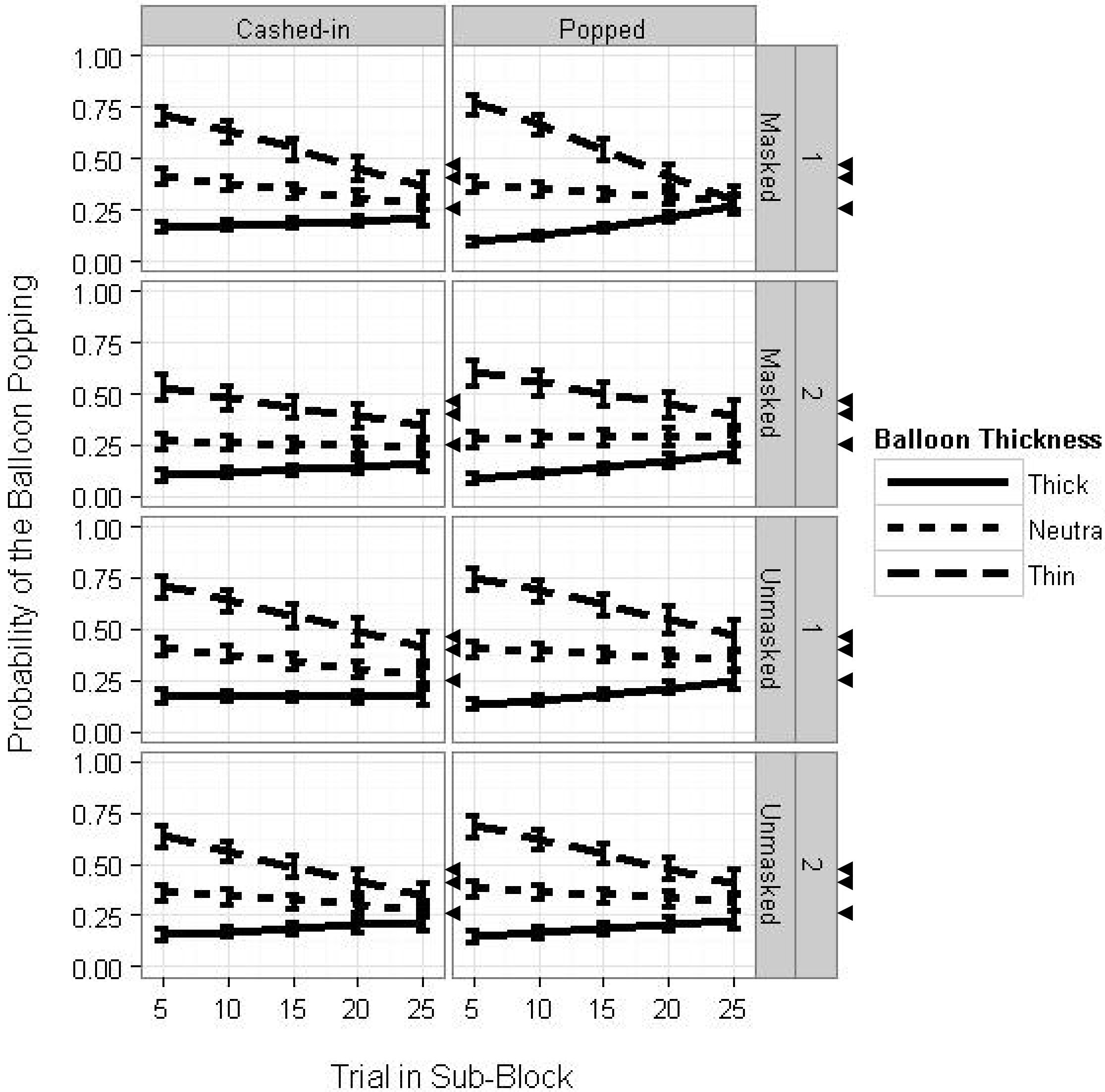


Figure 3.6. The effect of trial in sub-block, whether the previous balloon popped, condition, session, and balloon thickness on the probability of the participant popping the balloon on a given trial. The panels on the left side of the figure show probabilities when participants cashed-in the previous balloon, whereas the right panels of the figure show probabilities when participants popped the previous balloon. The different rows of the figure represent the separate conditions and sessions of the experiment. Different balloon thicknesses are represented by separate lines. The arrows at the right side of each panel indicate the

optimal probability of popping the balloon for each balloon thickness. The highest arrow (at .46) represents optimal behavior for the thinnest balloons, and the lowest arrow (at .25) represents optimal behavior for the thickest balloons.

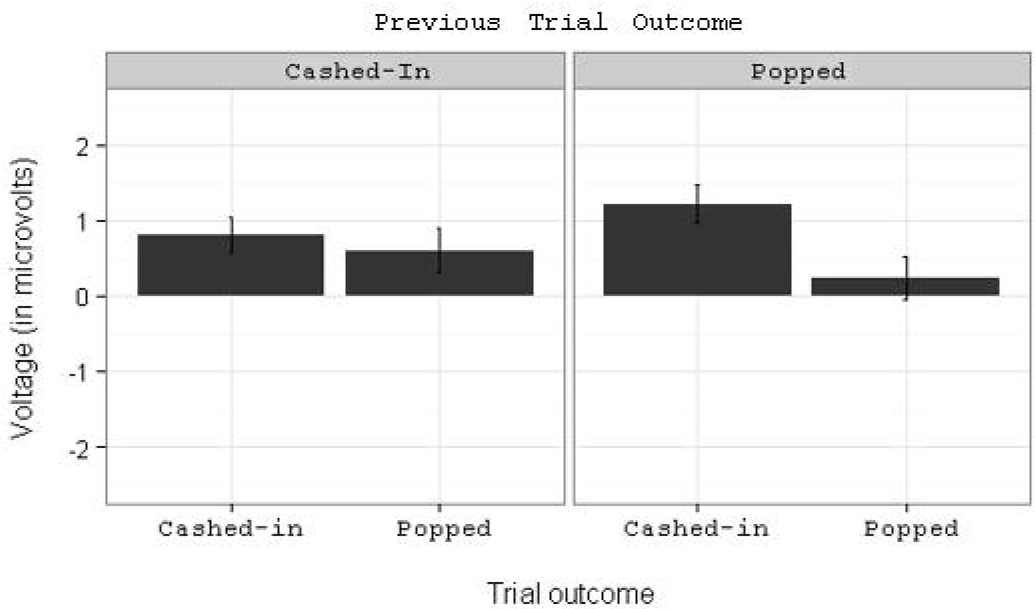
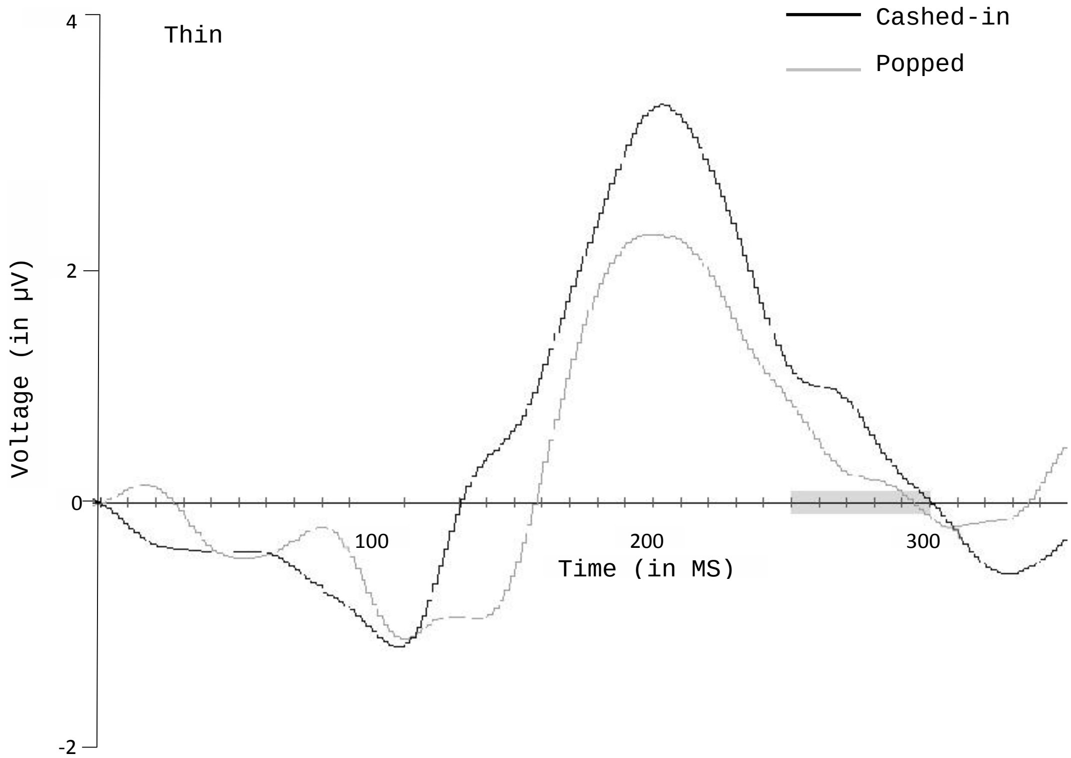
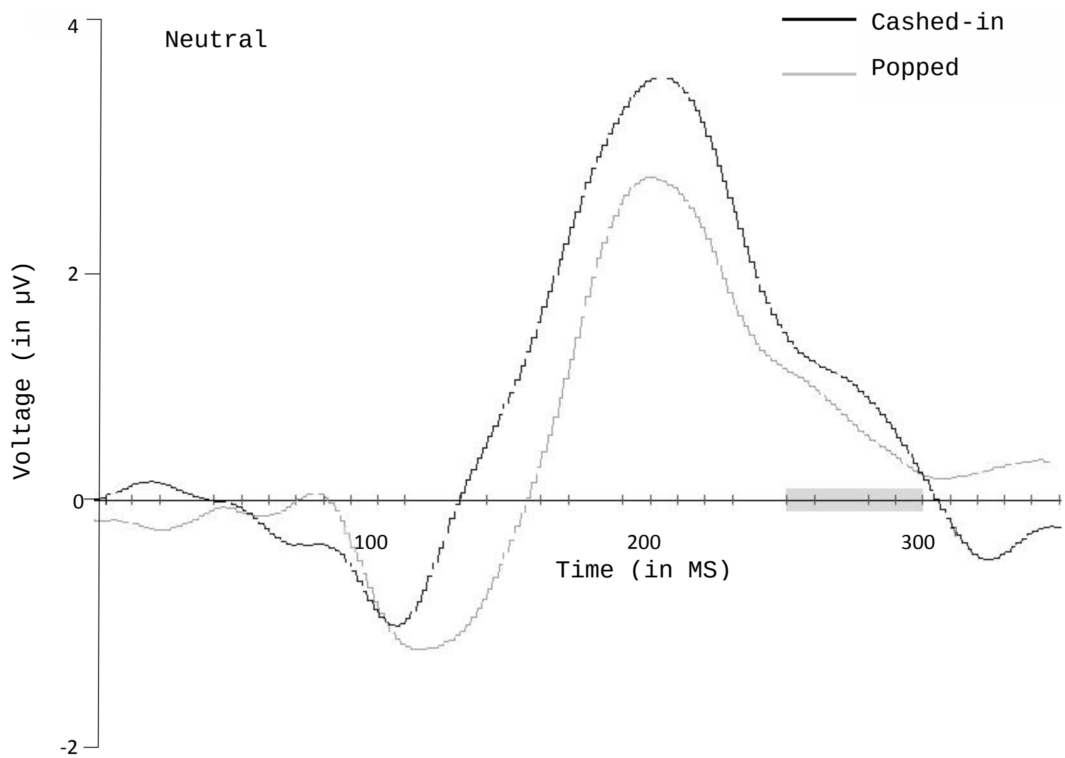


Figure 3.7. The effect of whether the previous balloon popped on the amplitude of the FRN (difference between cashed-in and popped trials). The left panel of the figure shows the amplitude of the FRN when the previous balloon was cashed-in. The right panel of the figure shows the amplitude of the FRN when the previous balloon was popped.


Thin
Cashed-in
Popped
Voltage (in µV)
4
2
0
-2
100
200
300
Time (in MS)


Neutral
Cashed-in
Popped
Voltage (in µV)
4
2
0
-2
100
200
300
Time (in MS)

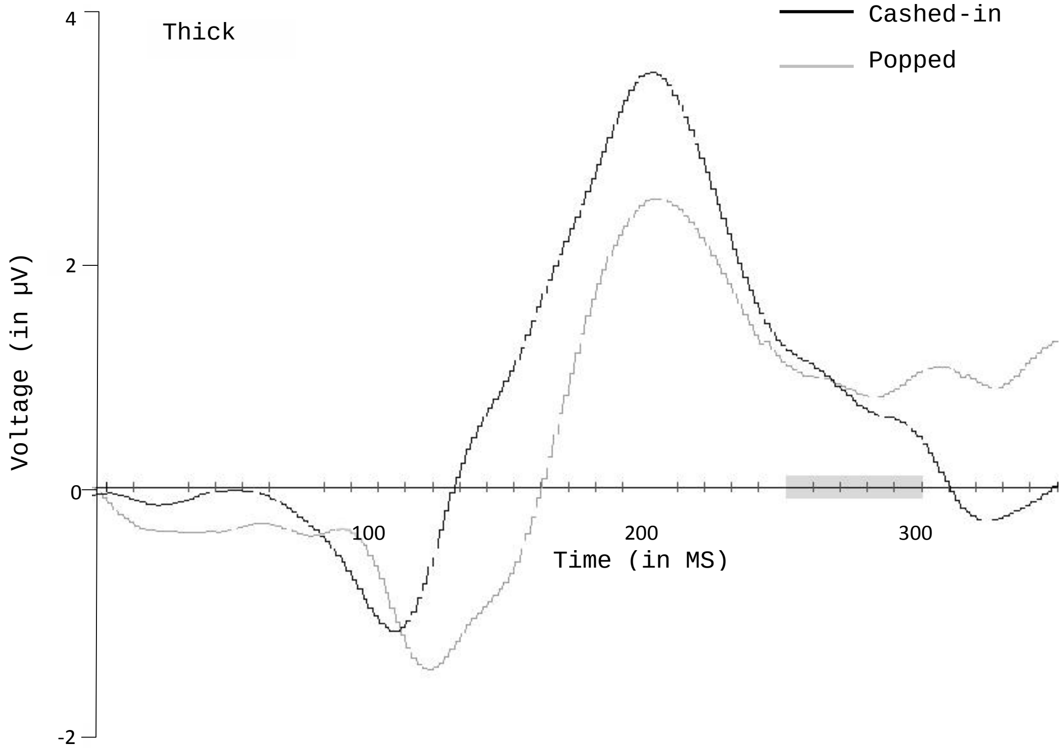


Figure 3.8. The average voltage for cashed-in balloons (dark lines) and popped balloons (light lines) across an average epoch. The shaded time region is the time region commonly associated with the FRN. (A) The average voltage across for the thinnest balloons. (B) The average voltage across for the neutral balloons. (C) The average voltage across for the thickest balloons.

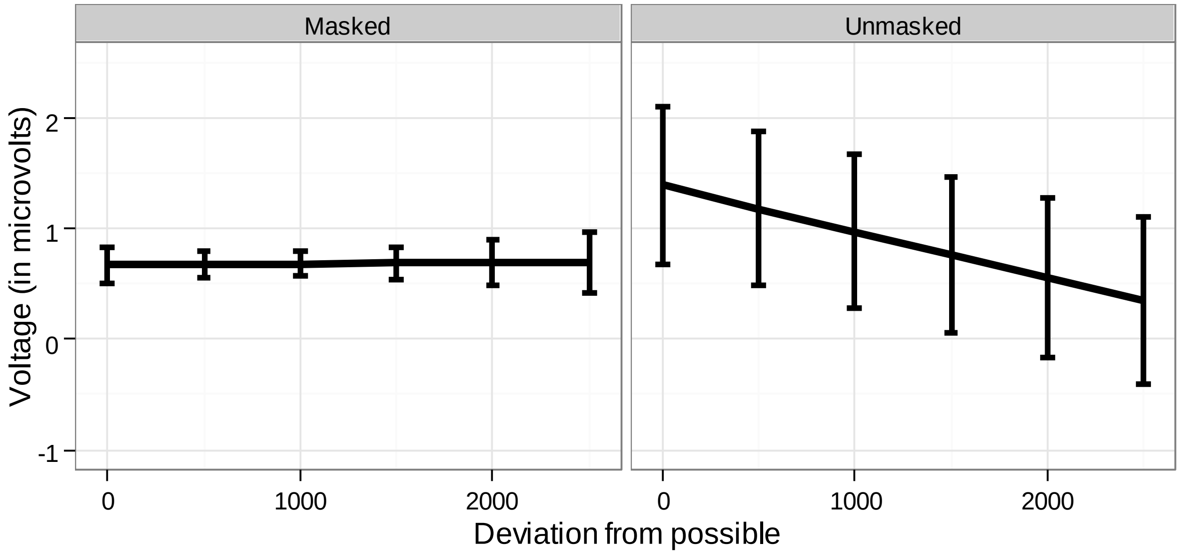


Figure 3.9. The effect of condition on the amplitude of the FRN. The FRN is operationalized as the difference in amplitude across the error continuum, for example, if the voltage is lower for larger errors than smaller errors then an FRN occurred. The left panel shows the effect in the masked condition. The right panel shows the effect in the unmasked condition.

Chapter 4 - General Discussion

The reigning theory on the FRN is the RL-ERN theory which suggests that the magnitude of errors affects the amplitude of the FRN. There is literature that suggests that the magnitude of errors does affect the amplitude of the FRN (Goyer et al., 2008) and also research that suggests that it does not affect the amplitude of the FRN (Hajcak et al., 2006; San Martin et al., 2010). The current set of experiments was designed to assess the effect of error magnitude on the amplitude of the FRN on a task that had not previously been used to study the phenomenon. The purpose of these experiments was to better understand the neural mechanisms underlying error processing in sequential risky decision making tasks involving dynamic environments. Both experiments utilized a variation of the BART methodology in which the balloon grew to full size on every trial. Although there were some initial concerns about the viability of this methodology, participants behaved in a similar manner on the current version of the BART as they had on previous versions. Specifically, there were worries about whether participants would be sensitive to the manipulation without experiencing the balloon popping. Figures 2.7 and 3.3 suggest that participants were sensitive to the balloon thickness in both experiments despite the loss of the overt cue.

BART Behavioral

The purpose of the behavioral analyses was to ensure that behavior on the task was similar to behavior on previous versions of the task. Specifically, I wanted to ascertain whether participants were sensitive to the balloon thickness manipulation, how they responded to errors (i.e., popped balloons), and variability in behavior on the task.

In general, the behavioral results indicate that behavior was similar across the two experiments. In both experiments, participants waited longest to cash-in the thickest balloons and shortest for the thinnest balloons, consistent with task contingencies. Participants also learned to respond more quickly later in a sub-block for the neutral and thinnest balloons, while learning to respond more slowly later in a sub-block for the thickest balloons. This result indicates that participants learned to behave more optimally through experience with the task. Participants also responded more quickly following a popped balloon than following a cashed-in balloon in both experiments. Looking at the latencies to respond, it is evident that participants' behavior changed after a balloon popped in the absence of immediate feedback indicating that an error occurred. In this study there was no cue that the balloon popped until participants received feedback, and there was some concern that this would result in participants not learning to perform the task. However, participants did recognize when an error occurred and corrected behavior in response to the error. These latency results replicate previous results which suggested that responding to inflate the balloon is not necessary for decision making on the task (McCoy & Young, In preparation). These results beg the question regarding the method of behavioral change. One means of assessing how participants are learning the task is assessing the variability in their wait times to determine if any patterns emerge.

When looking at variability in wait times (i.e., exploration), an interesting pattern emerged for the thick balloons. Early in a sub-block participants explored much more after the previous balloon popped than after the previous balloon was cashed-in. However, later in a sub-block participants shifted their strategy and explored less after a balloon popped than after the previous balloon was cashed-in. This seems to suggest a shift from a win-stay lose-shift strategy to some other strategy. This may be an important behavioral finding for future research using the BART, and may shed further light on how people change strategies while completing the task. Early in a sub-block, participants employed a win-stay, lose-shift strategy that likely helped them ascertain the most appropriate wait time for the sub-block of balloons. Later in a sub-block, when participants already knew how long to wait on average, shifting behavior following an error likely resulted in poorer performance causing participants to abandon the strategy. This shift in strategy in the BART is an intriguing finding. Is there a change in sensitivity to errors (e.g., habituation to the balloon popping) or are participants becoming more sensitive to a different cue?

For the thickest balloons, later in a sub-block, participants popped the balloon more often than they did earlier in a sub-block. This result is interesting given that one would expect participants to learn to avoid popping the balloon, an overt error. This behavioral result foreshadows the EEG results for so-called "correct" trials from Experiment 2. Specifically, this result seems to indicate that participants tend to view failing to obtain as many points as possible as an error. In fact, participants may view this as a larger error than the balloon popping for the thickest balloons, which would explain participants popping the balloon more often in order to obtain more points.

Taken together, the behavioral results suggest behavior in this version of the BART is relatively stable and highly replicable. This finding bodes well for the future of this task as a paradigm for studying decision making in the BART. These results also pose additional questions that merit future research.

Brain-Behavior Relationships

The EEG results were, in general, far less replicable across experiments than their behavioral counterparts. The only common parameter from the models for each experiment was the outcome of the current trial. Given that this is the variable that indicates whether an error occurred, it is not surprising that the two models share this parameter. However, the parameter estimates for the trial outcome variables are similar across the two experiments suggesting that the variable is having a similar effect in both experiments.

For the traditional FRN analysis (i.e., comparing voltage on popped and cashed-in trials), neither of the best fitting models included error magnitude. In Experiment 1 the best fitting model for the FRN data included fixed effects of trial outcome and balloon thickness. However, the best fitting model for the FRN data in Experiment 2 included fixed effects of trial outcome and whether the previous balloon popped. Neither of these results are supportive of the RL-ERN theory. The RL-ERN theory predicts that the magnitude of an error is most predictive of the subsequent FRN amplitude (Holroyd & Coles, 2002), a result which was not found in the current set of experiments. In the discussion of Experiment 1, it was suggested that perhaps the reason for the result may have been due to the additional effort required to cash-in the thinnest balloons. However, this finding was not replicated in Experiment 2 where beta power was not associated with FRN amplitude. Additionally, in Experiment 2 balloon thickness was not retained in the model. Digging deeper, it was evident that while the beta finding did not replicate in Experiment 2, the pattern of effects looked most similar for the masked condition when participants experienced it first. In this group, the amplitude of the FRN was correlated with power in the beta range (not statistically significant). Also, when looking at the effect of balloon thickness on the amplitude of the FRN in each condition and across sessions, the interaction replicated (not statistically significant) during the first session, but not the second session (see Figure 4.1). The difference appears to be caused by the loss of power in Experiment 2 due to the reduced experience with each condition. It also appears as though participants are processing the task differently in the second half of the experiment.

Interestingly, in Experiment 2 it was possible to assess the effect of error magnitude on so-called correct trials. In the unmasked condition participants were given information about how many points they could have earned on the current balloon before seeing the number of points they earned on the current balloon. The absolute difference between the number of points the participant could have earned and the number of points they did earn was used as a deviation score for assessing the magnitude of the error on cashed-in trials. On correct trials where the deviation was larger, the voltage was more negative than when the deviation was smaller. This finding is indicative of a magnitude effect on the amplitude of the FRN. This is supportive of the RL-ERN theory. This finding also indicates that participants process failures to maximize the number of points possible as errors.

The results from the traditional FRN analyses and the FRN analysis on correct trials is seemingly at odds. In the first case, there is no evidence of an effect of error magnitude of the amplitude of the FRN contrary to the predictions of the RL-ERN theory, but in the latter case there is a clear magnitude effect. Taken together these results suggest that the RL-ERN theory does not explain the breadth of phenomena that may affect the amplitude of the FRN.

Another difference between the experimental EEG results emerged. The traditional FRN analysis for Experiment 2 indicated an effect of whether the previous balloon popped on the amplitude of the FRN. It was suggested that successive errors may be indicative of a greater need to change behavior. However, this effect was not found in Experiment 1.

The disparity in EEG results across these two experiments suggests something fundamentally different may be going on. The most likely explanatory difference is the inclusion of two separate within-subjects conditions in Experiment 2. When assessing the parameter estimates for participants who experienced the masked condition first, the estimates for the variables are somewhat more similar to the same parameter estimates for Experiment 1. This limited experience with each condition may have resulted in a loss of power that produced the non-significant finding. Also, participants had less experience with a sub-block before moving to the next sub-block, which may have resulted in different strategies being employed for changing behavior. It may be that extended, uninterrupted experience with a given balloon thickness evokes a different strategy. Specifically, I hypothesize that when participants have limited exposure to a balloon thickness before changing to a new balloon thickness, having two errors occur in succession becomes a more valid cue for adapting future behavior than balloon thickness. For example, consider when a new sub-block of trials has just started and the participant is tasked with identifying the optimal strategy. Given that the participant has experienced at least one sub-block before, he or she knows there is a limited amount of time to identify a viable strategy. After a balloon pops, the participant adjusts his or her behavior, but the next balloon still pops. Because the participant knows there is a limited number of balloons of each type, he or she adjusts more strongly following the second error. This pattern of behavior is not found in Experiment 1, because participants have more chances to learn the optimal strategy.

The results of the current study suggest that the RL-ERN theory is not sufficient for describing the breadth of results in this and a multitude of other studies. Previous research has disclosed shortcomings in the RL-ERN theory. Hajcak et al. (2006) reported that the amplitude of the FRN is not moderated by the magnitude of the error. San Martin et al. (2010) found that the amplitude of the FRN was also affected by the probability of an error occurring. Specifically, they found that when the probability of an error was low, the amplitude of the FRN was greater than when the probability was high. The current results add the need to assess the amount of effort needed to complete the task. Many of the previous studies used simple gambling tasks in which participant effort was relatively low, but the current study used a task where participants had to respond quickly, requiring greater effort. It seems that the amount of effort necessary to complete the task may affect the amplitude of the FRN. This result suggests a need for caution in interpreting results in other studies where a task may be more difficult for one group of participants than for the other. Additionally, this study found that participants treat failures to maximize as an error. This result deserves future attention to better elucidate when a failure to maximize is considered an error and when it is not.

Error Importance

These results suggest something interesting about what is being coded by the FRN. Previous research has suggested a multitude of disparate phenomena being coded by the FRN, from whether an error occurred (Hajcak et al., 2006), the probability of an error occurring (San Martin et al., 2010), or the magnitude of an error (Goyer et al., 2008). The current results suggest that the amount of effort necessary to complete the task and consecutive errors may both influence the amplitude of the FRN. These findings suggest the FRN may be encoding error importance. When performing a task in which a variation in error magnitude is possible, it is likely that larger errors will be more important in determining optimal behavior than smaller errors. Likewise, if errors are unlikely to occur by chance in a task, when one does occur, it may signal a necessary change in behavior. In contrast, if errors are more likely to occur, then an error occurring may not signal a need to change behavior or it may simply signal a small change in behavior. The current results suggest that if the task is difficult and an error occurs, a change in behavior may be necessary, but this change in behavior may be harder to implement. When the behavioral change is difficult to implement, perhaps the FRN will be larger as is evidenced in the current study. It is postulated that if an error signals a need for behavioral change, it will be interpreted as more important than an error that does not signal a need for behavioral change. Additionally, an error that signals the need for a larger or more difficult change in behavior will be interpreted as more important than one that signals the need for a small or easy change in behavior.

If error importance is the driving factor affecting the amplitude of the FRN then how do the results of Experiment 2 fit the narrative? How are consecutive errors “more important” than single errors? If participants are using a strategy based on limited experience with a sub-block, then this type of error may be the most valid cue that behavior needs to change, and thus the most important error. The FRN on correct trials seems to support this hypothesis as well. These results indicated that participants interpreted deviations from the number of points possible as errors. Thus, it makes sense that the larger deviations indicate a need for a larger change in behavior, and are thus more important. Taken together, the results from the research on the FRN suggest a need to reconceptualize what the FRN is encoding, specifically how does the importance of the error for aligning behavior with current task goals affect the amplitude of the FRN.

Future research will be necessary to determine whether the error-importance hypothesis described above is better suited to describing error processing than the RL-ERN. One testable prediction that could differentiate between the RL-ERN theory and the error-importance hypothesis is whether the error affects the participant’s goals for the task. For example, a large error that is important for the participant’s goals should elicit a larger FRN than a large error that is irrelevant to the participant’s goals. Thus a follow up to the current experiment could utilize the unmasked condition for Experiment 2 in which participants know how many points they could have earned on the current balloon. Using this methodology, participants could be placed into one of two between subjects conditions. In the important error condition, participants would be informed that they need to get as many points as possible. In the unimportant error condition, participants would be told that they need to avoid popping the balloon. Participants should elicit an FRN in the important error condition regardless of which theory is correct. However, if the RL-ERN theory is correct then an FRN should be evoked on error trials in the unimportant error condition, but not if the error importance hypothesis is correct.

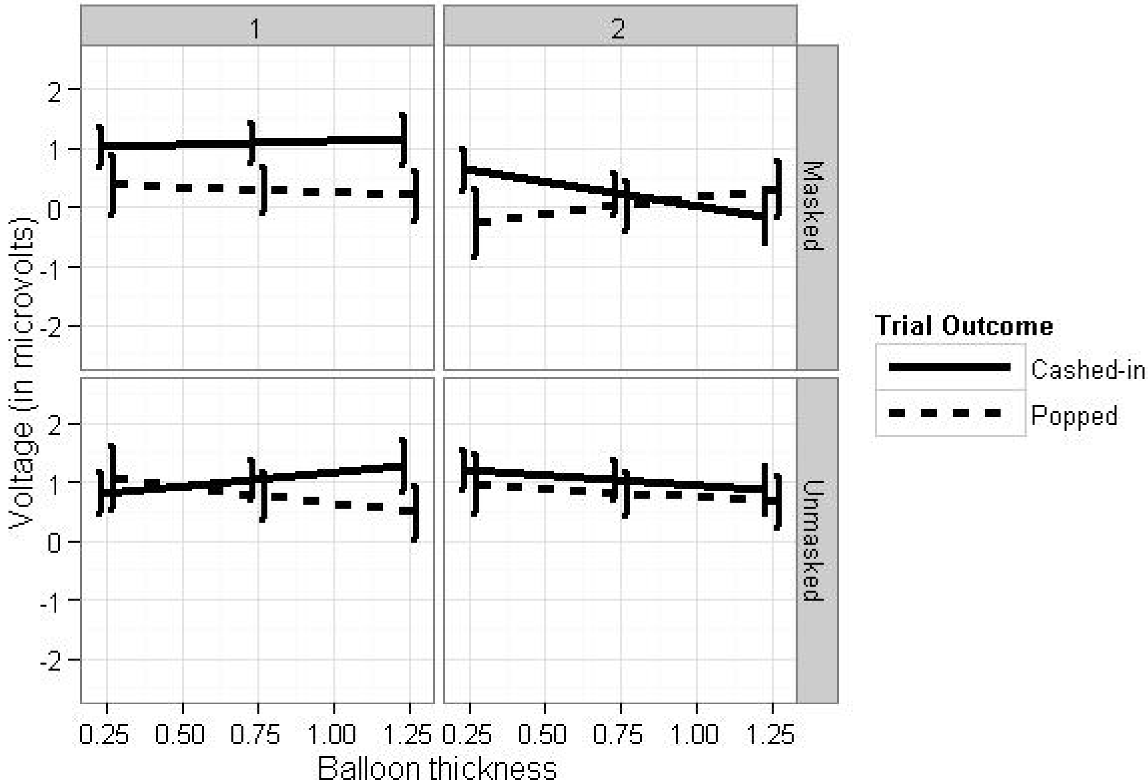


Figure 4.1. The effect of balloon thickness on the amplitude of the FRN. The left panels display the effect in session one and the right panels display the effect in session two. The top panels display the effect in the masked condition and the bottom panels display the effect in the unmasked condition.

References

Addicott, M. A., Pearson, J. M., Wilson, J., Platt, M. L., & McClernon, F. J. (2013). Smoking and the bandit: A preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Experimental and Clinical Psychopharmacology, 21*, 66-73. doi: 10.1037/a0030843.supp

Aklin, W. M., Lejuez, C. W., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour Research and Therapy*, *43*(2), 215-228. doi: 10.1016/j.brat.2003.12.007

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of mental disorders* (5 ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, C. M. (2012). Ambiguity aversion in multi-armed bandit problems. *Theory and Decision, 72*, 15-33.

Arbel, Y., Goforth, K., & Donchin, E. (2013). The good, the bad, or the useful? The examination of the relationship between the feedback-related negativity (FRN) and long-term learning outcomes. *Journal of Cognitive Neuroscience, 25*, 1249-1260.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293-1295.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience, 27*, 1823-1835.

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*, 3343-3350.

Centers for Disease Control. (2013). Sexual Risk Behavior Data & Statistics.   Retrieved 10/28/2013, 2013

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143-149.

Eriksen, C. W., Hamlin, R. M., & Daye, C. (1973). The effect of flanking letters and digits on speed of identifying a letter. *Bulletin of the Psychonomic Society, 2*, 400-402.

Falkenstein, M., Hohnsbein, J., & Hoorman, J. (1991). Effects of crossmodal divided attention on late ERP components. 2. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78*, 447-455.

Fein, G., & Chang, M. (2008). Smaller feedback ERN amplitudes during the BART are associated with a greater family history density of alcohol problems in treatment-naive alcoholics. *Drug and Alcohol Dependence, 92*(1-3), 141-148. doi: 10.1016/j.drugalcdep.2007.07.017

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

Franken, I. H. A., van Strien, J. W., Franzek, E. J., & van de Wetering, B. J. (2007). Error-processing deficits in patients with cocaine dependence. *Biological Psychology, 75*, 45-51.

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*, 385-390.

Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components*. New York, NY: Oxford University Press.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279-325.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences, 423*, 52-77.

Goyer, J. P., Woldorff, M. G., & Huettel, S. A. (2008). Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *Journal of Cognitive Neuroscience, 20*, 2058-2069.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, G. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology, 71*, 148-154.

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*, 905-912.

Hauser, T. U., Iannaccone, R., Stampfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage, 84*, 159-168.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709. doi: 10.1037//0033-295x.109.4.679

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, and Behavioral Neuroscience, 9*(59-70).

Hopko, Derek R., Lejuez, C. W., Daughters, Stacey B., Aklin, Will M., Osborne, Amanda, Simmons, Burnetta L., & Strong, David R. (2006). Construct Validity of the Balloon Analogue Risk Task (BART): Relationship with MDMA Use by Inner-City Drug Users in Residential Treatment. *Journal of Psychopathology and Behavioral Assessment, 28*(2), 95-101. doi: 10.1007/s10862-006-7487-5

Howells, F. M., Stein, D. J., & Russell, V. A. (2010). Perceived mental effort correlates with changes in tonic arousal during attentional tasks. *Behavioral and Brain Function, 6*, 1-15.

Hunt, M. K., Hopko, D. R., Bare, R., Lejuez, C. W., & Robinson, E. V. (2005). Construct validity of the Balloon Analog Risk Task (BART): associations with psychopathy and impulsivity. *Assessment, 12*(4), 416-428. doi: 10.1177/1073191105278740

Kehoe, E. J., Olsen, K. N., Ludvig, E. A., & Sutton, R. S. (2009). Scalar timing varies with response magnitude in classical conditioning of the nictitating membrane response of the rabbit (Oryctolagus cuniculus). *Behavioral Neuroscience, 123*, 212-217.

Leigh, B. C. (1999). Peril, Chance, Adventure: Concepts of risk, alcohol use and risky behavior in young adults. *Addiction, 94*, 371-383.

Lejuez, C. W., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., & Gwadz, M. (2007). Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART-Y) in the assessment of risk-taking behavior among inner-city adolescents. *Journal of Clinical Child and Adolescent Psychology, 36*, 106-111.

Lejuez, C. W., Aklin, Will M., Jones, Heather A., Richards, Jerry B., Strong, David R., Kahler, Christopher W., & Read, Jennifer P. (2003). The balloon analogue risk task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology, 11*(1), 26-33. doi: 10.1037/1064-1297.11.1.26

Lejuez, C. W., Aklin, Will M., Zvolensky, Michael J., & Pedulla, Christina M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of Adolescence, 26*(4), 475-479. doi: 10.1016/s0140-1971(03)00036-8

Lejuez, C. W., Read, Jennifer P., Kahler, Christopher W., Richards, Jerry B., Ramsey, Susan E., Stuart, Gregory L., . . . Brown, Richard A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75-84. doi: 10.1037//1076-898x.8.2.75

Luce, R. D. (1959). *individual choice behavior*. Oxford, England: John Wiley.

Luck, S. J. (2005). *An introduction to the event related potential technique*. Cambridge, MA: MIT Press.

Marco-Pallares, J., Cucurell, D., Munte, T. F., Strien, N., & Rodriguez-Fornells, A. (2011). On the number of trials needed for a stable feedback-related negativity. *Psychophysiology, 48*, 852-860.

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

McCoy, A. W., & Young, M. E. (In preparation). Modifying the balloon analogue risk task for use in electroencephalographic paradigms.

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

Mishkin, M, & Pribram, K. H. (1955). Analysis of the effects of frontal lesions in monkeys: I. Variations of delayed alterations. *Journal of Comparative and Physiological Psychology, 48,* 492-495.

Murray, F. S. (1971). Multiple probable situation: A study of a five one-armed bandit problem. *Psychonomic Science*, *22*, 247-249.

National Institute on Drug Abuse. (2012). National institute on drug abuse: Trends and statistics. 2013

Piazza, P. V., & Deroche-Gamonet, V. (2013). A multistep general theory of transition to addiction. *Psychopharmacology, 229*, 387-413.

Pleskac, T. J., Wallsten, T. S., Wang, P., & Lejuez, C. W. (2008). Development of an automatic response mode to improve the clinical utility of sequential risk-taking tasks. *Experimental and Clinical Psychopharmacology, 16,* 555-564.

Racey, D., Young, M. E., Garlick, D., Pham, J. N., & Blaisdell, A. P. (2011). Pigeon and human performance in a multi-armed bandit task in response to changes in variable interval schedules. *Learning and Behavior, 39*, 245-258.

San Martin, R., Manes, F., Hurtado, E., Isla, P., & Ibanez, A. (2010). Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage, 51*(3), 1194-1204. doi: 10.1016/j.neuroimage.2010.03.031

Shephard, E., Jackson, G. M., & Groom, M. J. (2014). Learning and altering behaviours by reinforcement: Neurocognitive differences between children and adults. *Developmental Cognitive Neuroscience, 7,* 94-105.

Sutton, R. S., & Barto, A. G. (1988). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150,* 1187-1188.

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *The Journal of Neuroscience, 33*, 8264-8269.

Thomas, G., Kacelnik, A., & van der Meulen, J. (1985). The three-spined stickleback and the two-armed bandit. *Behaviour, 93,* 227-240.

van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex, 20*(1596-1603).

Wang, D. V., & Tsien, J. Z. (2011). Convergent processing of both positive and negative motivational signals by the VTA dopamine neuronal populations. *PLOS One, 6,* 1-14.

Young, M. E., Webb, T. L., & Jacobs, E. A. (2011). Deciding when to "cash in" when outcomes are continuously improving: an escalating interest task. *Behavioural Processes, 88*(2), 101-110. doi: 10.1016/j.beproc.2011.08.003

Young, M. E., Webb, T. L., Rung, J. M., & McCoy, A. W. (2014). Outcome probability versus magnitude: When waiting benefits one at the cost of the other. *PLOS One, 9,* 1-7.

Appendix A - Processing EEG Data

One of the most important concepts to remember about EEG is that it cannot inherently differentiate between sources of electricity. It is important to understand that the information present in the raw data is an amalgamation of the output of multiple electrical sources, and to analyze the electrical source of interest it is necessary to isolate the sources of the signal and to remove the portions not attributed to the source of interest.

EEG data are inherently messy. The data is often distorted by ocular artifacts, muscular artifacts and internal/external environmental differences between participants. Ocular artifacts include eye blinks and eye movements. The eyeball can be conceptualized as a battery with the front of the eyeball being positively charged and the back of the eyeball being negatively charged. When a participant moves his or her eyes, the ipsilateral electrodes will show a large, fast, positively valenced change in the EEG wave with a slow return to baseline that is not temporally linked to any specific component or event. The contralateral electrodes will show an oppositely valenced shift of similar magnitude. The size of the shift is negatively correlated with distance from the eyes. The other ocular artifact is blinks; eye blinks can be evidenced by their rapid, large peaks in the waveform that are not temporally linked to any specific component or event. Both eye blinks and movements are considered step-like artifacts based on their appearance in the waveform. They are most easily dealt with by reducing their number during data acquisition. Prior to data acquisition participants are asked to limit their eye movements and blinks during important phases of the trial. Another issue when dealing with ocular artifacts is the variable nature of their amplitude.

The amplitude of a participant's ocular artifacts is unique to him or her with the size of the artifact being on average between 75 μV and 200 μV. Because of this variability, a generic threshold for artifact rejection is inappropriate. When rejecting ocular artifacts, the first step must be to set up a signal detection type approach in which a criterion is set – the amplitude of the abrupt change in voltage or the slope of the change in voltage. With this criterion set the researcher must visually inspect a percentage of the raw data and determine the rates of hits, misses, correct rejections, and false alarms. For the detection of ocular artifacts, a hit is an instance in which an artifact is present and the voltage criterion indicates that section of the data should be rejected. A miss is an instance in which an artifact is present, but does not reach the threshold for rejection. A correct rejection is an instance in which no artifact is present and the voltage criterion is not met, this category is likely to be the most common category for the data to fall into if too much data is outside of this category there may be major problems with the data and it may need to be rejected completely. The final category is false alarms in which no artifact is present, but the voltage threshold is met and the data section is rejected. Too many misses or false alarms indicates that the criterion used is too conservative or too liberal, respectively, and needs to be adjusted. To determine the validity of the criterion, bias will be assessed by dividing the number of hits by the number of false alarms.

Once an effective voltage criterion is determined, an artifact rejection tool will be run. This tool, available in EGI's NetStation, allows a custom voltage threshold to be set and the type of artifact rejection to be selected. The next step involves the NetStation's Ocular Artifact Rejection (OAR) tool; this tool sets a slope threshold for the removal of eye artifacts. Their distinctive shape and scalp distribution allows for the modeling of eye artifacts using a technique known as independent components analysis (see Appendix). This technique allows for the removal of eye artifacts without the rejection of a trial, conserving trials which would otherwise be rejected. The artifact rejection tool is then rerun to re-label trials for rejection.

Another artifact of importance is myogenic activity. The volume of electrical activity created by muscles is much greater than that given by the ERP components of interest. When a participant clenches her or his jaw or similarly contracts any muscle group near or on the skull, a large series of peaks occurs in rapid succession washing out the data for the duration of the muscle movement. When these artifacts occur, the only option is to reject the trial. The best practice for the reduction of muscle artifacts is to ask participants to try not to clench their muscles during the important phases of the trials.

Finally, there are some artifacts that are best ignored or removed during the filtering process. Slow voltage shifts are caused by changes in impedance across time at a single electrode. These artifacts are often caused by the participant moving or sweating. The best practice for reducing these artifacts is to tell the participant not to move during the study and to keep the chamber at a cool temperature throughout the session to reduce the likelihood of the participant sweating. If large enough, these artifacts will need to be dealt with by removing the offending electrode from the analyses. The next type of artifact is line noise and other environmental electrical noises. Line noise is caused by the AC electric current in wires and is often 50 or 60 Hz. Another common source of electrical noise is computer monitors. Computer monitors emit electrical noise due to their refresh rate between 50- and 120- Hz. These sources of electrical noise are easily dealt with by employing a low-pass filter. A low-pass filter allows activity below a certain threshold to pass and be included in the final data. For the current research a low-pass filter with a half attenuation set at 30 Hz will be used. This filter type reduces the occurrence of window edge-related artifacts in which data appears to be bunched at the top of the distribution of electrical frequencies by attenuating frequencies at the filter threshold instead of having a hard stop at the threshold. The half-attenuation point (in this case

30 Hz) is the point at which half of the power, amplitude squared, of the frequency is attenuated. The attenuation curve is a smooth curve from complete attenuation to no attenuation. There is also high-pass filtering which removes extremely slow frequencies; for this study, a high-pass filter will be set at 0.1 Hz to attenuate the effects of very slow waves. The final type of artifact is not generally detrimental to the ability to interpret the results. Alpha waves are frequencies around 10 Hz which indicate that the participant is fatigued. These artifacts are not amenable to filtering out nor can they be picked up as obvious artifacts. They cannot be filtered out because of their frequency – a they are too close to the frequency range of interest, and filtering alpha waves out would result in the data being skewed. They often negatively correlated with mental activity such as perceiving, decision making, and learning because these activities tend to be engaging. The best practice for dealing with alpha waves is to keep them from occurring in the first place, alpha waves can be avoided by using well-rested participants and by jittering the start time for the time-locking event by 50 ms.

Tools provided with EGI's NetStation software will be used for cleaning the data. The first step is to visually inspect the data and determine a sufficient voltage threshold and slope for the artifact rejection tools. The data are then high (0.1 Hz) and low (30 Hz) pass filtered followed by segmentation to the onset of the feedback screen, or the response for frequency analyses. After segmentation the data undergo an initial artifact rejection that marks the segments with artifacts as bad, OAR that checks the marked segments for ocular artifacts, and another artifact rejection that overwrites the initial artifact rejection tool's marks based on the removal of ocular artifacts by the OAR tool. The data are then re-examined to determine if any other trials need to be rejected based on artifacts that were not caught. The data then undergo bad channel replacement (BCR). In bad channel replacement, a technique known as spherical spline interpolation is used to determine the voltage at the bad channel. Spherical spline interpolation involves taking the voltage readings at the nearby electrodes and interpolating the voltage at the bad channel from the voltage of the other channels. This technique is more accurate with a denser array of channels (e.g., the value at a bad channel would be most accurate when interpolated from a 256-channel cap, moderately accurate when interpolated using a 64-channel cap and poorly interpolated using a 19-channel cap).

The data will then undergo montage operations to re-reference the data. During data acquisition the voltage is on-line referenced to the Cz electrode, but leaving this electrode as the reference electrode results in biased data. Voltage cannot be calculated based on a single point because it is the movement of current between two points. The activity at each electrode (other than Cz) is compared to the activity at Cz to determine the voltage at a given electrode. This reference results in a bias; amplitude of a given deflection increases as a function of distance from Cz. If the voltage reading at a given electrode is the difference between activity at that electrode and Cz, then the size of a given deflection will necessarily be larger as distance from Cz increases. To correct for this bias, the data will be offline average referenced. An average reference compares the activity at a given electrode to the activity at each electrode resulting in a less biased reference. According to Luck (2005) the average reference is a poor choice due to its bias; however, the author fails to consider that every known reference is biased in some manner and the average reference has been suggested as the least biased estimate available.