Event-Related Potential Correlates of Catecholinergic Neuromodulators Norepinephrine and Dopamine

by

Christopher Michael Warren
BAH, University of Guelph, 2006
MSc., University of Victoria, 2008

A Dissertation Submitted in Partial Fulfillment of the Requirements of the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

Event-Related Potential Correlates of Catecholinergic Neuromodulators Norepinephrine and Dopamine

by

Christopher Michael Warren
BAH, University of Guelph, 2006
MSc., University of Victoria, 2008

Supervisory Committee

Dr. Clay B. Holroyd, Supervisor
(Department of Psychology)

Dr. Michael E. J. Masson, Departmental Member
(Department of Psychology)

Dr. D. Stephen Lindsay, Departmental Member
(Department of Psychology)

Dr. Raad Nashmi, Outside Member
(Department of Biology)

Supervisory Committee

Dr. Clay B. Holroyd, Supervisor
(Department of Psychology)

Dr. Michael E. J. Masson, Committee Member
(Department of Psychology)

Dr. D. Stephen Lindsay, Committee Member
(Department of Psychology)

Dr. Raad Nashmi, Committee Member
(Department of Biology)

ABSTRACT

Adaptive decision making depends on multiple processes, including fast and efficient processing of stimulus events for effective responding and slow trial-to-trial learning of action values for optimization of the selection process. I applied the event-related brain potential (ERP) technique to investigate the involvement of two neuromodulatory systems in learning and decision making: The locus coeruleus-norepinephrine (LC-NE) system and the mesencephalic dopamine system (DA system). I present evidence that the "oddball" N2, a negative deflection in the ERP elicited by task-relevant events that begins approximately 200 ms after onset of the eliciting stimulus and that is sensitive to low-probability events, is a manifestation of *cortex-wide* noradrenergic modulation recruited to facilitate the processing of unexpected stimuli. Further, Holroyd and Coles (2002) proposed that the impact of DA reinforcement learning signals on the anterior cingulate cortex (ACC) produces a component of the ERP called the feedback-related negativity (FRN). I present electroencephalographic evidence that both the DA system and the LC-NE system act in concert when learning from rewards that vary in expectedness, but that the DA system is relatively more exercised when subjects are highly engaged/challenged by learning tasks, whereas the impact of the LC-NE system is attenuated by this manipulation.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I want to thank my PhD adviser, Clay Holroyd, for inspiring me, for teaching me, and for patience and diligence in working with me. I want to thank my Master's adviser, Mike Masson, also for teaching me and for encouraging me to work on the questions that interested me most. I also thank Steve Lindsay for showing unwavering confidence in me through my 5 years of graduate school at the University of Victoria, and both Steve and Raad Nashmi for serving on my committee. And finally, I thank Sander Nieuwenhuis for the two brilliant papers cited herein that provoked this work, for agreeing to serve as my external examiner, and for the stimulating and insightful discussion that resulted.

DEDICATION

For Nikki.

Because you did everything while I did nothing but work on this.

General Introduction

The central question of this thesis concerns how specific neuromodulators act to influence brain activity in an adaptive manner. Neuromodulators are neurotransmitters that target large populations of neurons and regulate the way those neurons respond to direct synaptic transmission of other neurotransmitters such as glutamate and γ-aminobutyric acid (GABA) (e.g. Aston-Jones & Cohen, 2005; Briand, Gritton, Howe, Young, & Sarter, 2007). Thus, neuromodulators change communication between neurons, and can have a profound effect on cognition and behaviour. For example, many clinical disorders have been associated with and are treated with pharmacological interventions that target specific neuromodulators, such as selective serotonin reuptake inhibitors for increasing serotonin levels to alleviate depression (e.g. Aston-Jones, & Cohen, 2005; Briand et al., 2007; Fournier et al., 2010). In this thesis I work toward establishing and further developing links between specific electrophysiological manifestations of brain activity and the two catecholinergic neuromodulators, norepinephrine (NE) and dopamine (DA), respectively. In particular, I propose that an event-related brain potential (ERP) component called the N2 is a manifestation of noradrenergic modulation. In addition, I provide support for the theory that phasic DA activity, which is intimately involved in reinforcement learning, is indexed by an ERP component known as the feedback-related negativity (FRN). The experiments reported herein specifically address the relationship between NE and the N2, and the interaction of NE and DA in producing the N2 and FRN to better understand both components. At a more general level, this work is directed toward validating these ERP components as indexes of specific neuromodulatory activity that can be used to examine non-invasively the function or malfunction of these neuromodulatory systems in diverse populations and situations.

The locus coeruleus-norepinephrine system

The primary source of NE to the forebrain is the locus coeruleus (LC), a nucleus in the pontine region of the brainstem composed of approximately 10,000 to 15,000 NE containing neurons per hemisphere (Berridge & Waterhouse, 2003). The LC-NE system projects to virtually the entire central nervous system except the basal ganglia and is the exclusive source of noradrenergic innervation to the hippocampus and neocortex. The LC receives its major afferent connections from the prefrontal cortex, especially the anterior cingulate cortex (ACC) and the orbital frontal cortex (OFC), which is held up in support of the view that LC activity can be mediated by high-level cognitive and affective processes (Aston-Jones & Cohen, 2005; Aston-Jones, Foote, & Bloom, 1984; Berridge & Waterhouse, 2003).

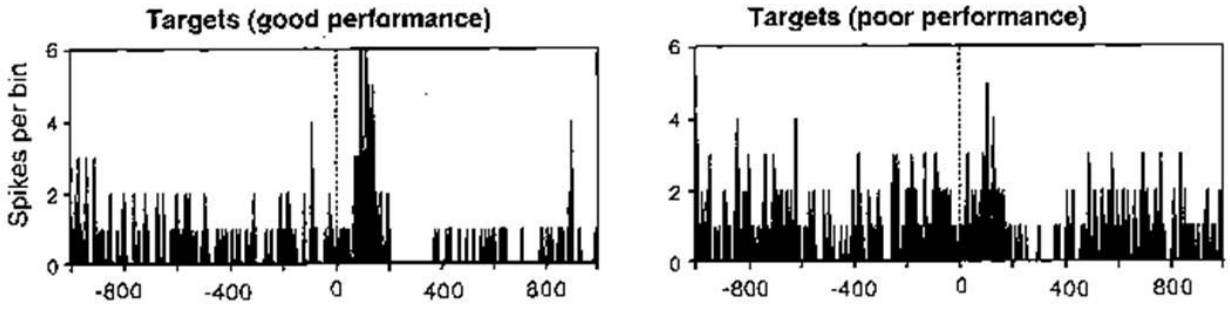

Figure 1. LC activity recorded from monkeys during a visual target-discrimination task, during epochs of good target-detection performance (left) and bad target-detection performance (right). The dotted vertical line indicates target onset. Activity prior to target onset is representative of tonic LC activity, whereas the pronounced peak in firing following target onset is representative of phasic LC activity. Immediately following the phasic burst of activity, the LC firing is inhibited, leading to a short, refractory-like period of reduced activity. Periods of poor performance are marked by high tonic activity, attenuated phasic activity, and reduced auto-inhibition. This figure was adapted from Usher, Cohen, Servan-Schreiber, Rajkowski, and Aston-Jones (1999).

The LC is associated with two characteristic components of activity: Tonic activity and phasic activity (see Fig. 1). Tonic activity consists of firing rates between 0 and 5 Hz, and is associated with the level of arousal and the sleep-wake cycle of an organism (Berridge & Waterhouse, 2003). Within the tonic component of LC activity, the relationship between LC firing

rate and the behaviour of an organism parallels the Yerkes-Dodson curve relating arousal to performance (Yerkes & Dodson, 1908). Greater LC firing frequencies are associated with greater arousal, alertness, wakefulness, and vigilance up to about 3 Hz, when further increases in firing rate are associated with distractibility and erratic behaviour (Aston-Jones, Rajkowski, Kubiac, & Alexinsky, 1994).

LC phasic activity is characterized by bursts of rapid firing (up to 20 Hz) in response to task-relevant or especially salient environmental stimuli. These bursts tend to involve 2 to 3 action potentials approximately 100 ms to 150 ms following onset of the eliciting stimulus (in monkeys). The bursts are followed by a sudden period of suppressed firing lasting approximately from 200 ms to 500 ms post-stimulus onset (Berridge & Waterhouse, 2003). The period of suppressed firing is due to the self-inhibitory nature of the LC-NE system. Although NE increases the responsivity of target neurons outside the LC itself, LC neurons are inhibited by the NE they release, reducing the chance that they will fire again for a brief refractory-like[1] period (Aghajanian, Cedarbaum, & Wang, 1977).

NE impacts brain function by potentiating neuronal responses to both excitatory and inhibitory inputs while having little to no impact on the background firing rate (e.g. Foote, 1987; Waterhouse, Hoffer, & Freedman, 1979a; Woodward, Moises, Hoffer, & Freedman, 1979b). For example, Woodward et al. (1979b) iontopheretically applied NE to Purkinje neurons in the cerebellum. NE did not affect the spontaneous firing rate of the neurons, but when glutamate was also applied to excite the neuron, the excitation was enhanced relative to when glutamate was applied in the absence of NE. Similarly, when GABA was applied, inhibition was increased in the presence of NE. This impact of NE is similar across all cortical regions (Foote & Morrison, 1987; Waterhouse et al., 1980, 1981).

[1] As opposed to the potassium-mediated refractory period common to individual neurons throughout the brain.

Servan-Schreiber, Prinz, and Cohen (1990) demonstrated that at the level of the neural network, increases in the responsivity of each individual neuron improves the signal-to-noise ratio of the network. They modelled the effect of NE as an increase in the *gain* of the activation function of the neural units. The activation function that describes neural responsivity is a sigmoidal curve relating input power to output power, giving the appearance of asymptote at extreme positive or negative inputs, while approximating a linear relationship at inputs close to zero. Gain is a parameter that impacts the shape of the activation function, such that when gain is higher, smaller differences in input will cause the system to respond maximally (steepening the slope of the line at values of input close to zero). When applied to neurons, an increase in gain means that small inhibitory or excitatory inputs yield larger inhibitory or excitatory outputs. When such an increase in gain is applied to a multilayered neural network, the network demonstrates an improvement in signal detection. Thus, the presence of NE in target neural networks will improve performance of the network. However, Servan-Schreiber and colleagues noted three costs associated with this kind of improvement. First, the enhancement of a signal relative to background noise encourages rigid responding to the signal only, at the expense of variability cultivated by noise. Noise can be important for eliciting random responses in an exploratory manner (e.g. Aston-Jones & Cohen, 2005; Yu & Dayan, 2005). A second cost of noradrenergic boosts in signal detection involves the case when a weak signal carries important information that is overshadowed by a stronger signal. In the presence of NE, the weaker signal would suffer and ultimately lose a competition with the stronger signal in early layers of the network, whereas in the absence of noradrenergic modulation it would have a chance to gain traction through the layers of the network to ultimately provide an important influence on the outcome of network processing. Finally, there is evidence that maintaining enhanced signal detection is metabolically demanding (e.g. McCulloch, 1982, as cited by

Servan-Schreiber et al.). These costs illustrate why the recruitment of NE to improve signal detection should be used in a sparing and strategic manner.

Usher, Cohen, Servan-Schreiber, Rajkowski, and Aston-Jones (1999) characterized the phasic burst of NE as a temporal attention filter that selects for the occurrence of motivationally salient events, responding with a NE release that briefly facilitates responding to these events by speeding/improving signal detection throughout the cortex. When monkeys are trained to respond to infrequent visual targets but not frequent distracters, single-cell recordings from the LC demonstrate increased phasic firing to targets (Aston-Jones et al., 1994). Using the same paradigm, Usher et al. (1999) categorized epochs of behavioural performance (within subjects) into periods of good performance and periods of bad performance on the basis of false alarms and misses. Periods of better performance were associated with low tonic activity overall and high phasic activity in response to presentation of a target stimulus. In contrast, poor performance was associated with high tonic activity and decreased phasic responses to target presentation. Periods of good behaviour were also associated with a narrower distribution of reaction times, interpreted as the effect of low tonic activity in eliminating anticipatory responses, and the effect of strong phasic activity in speeding up slow responses. Usher et al. noted that the timing of NE-mediated improved signal detection corresponded closely with the period of effective target processing that precedes a cognitive phenomenon known as the attentional blink (e.g., Raymond, Shapiro, & Arnell, 1992). When two targets are embedded within a rapid serial visual presentation, the first target is processed effectively, and so is the second if it is presented within approximately 200 ms of onset of the first target (called "lag-1 sparing"). If the second target is presented between approximately 200 ms and 600 ms after onset of the first, it suffers from a deficit in processing that Usher et al. implied is due to the LC refractory-like period. This relationship was developed extensively by Nieuwenhuis, Gilzenrat, Holmes, and Cohen (2005b) and will be discussed further below.

The mesencephalic dopamine system

The DA system is composed of several nuclei of DA-releasing neurons including the substantia nigra, the pars compacta, and the ventral tegmental area. These nuclei project mainly to the basal ganglia and to frontal structures associated with cognitive control such as the ACC and prefrontal cortex (e.g., S. M. Williams & Goldman-Rakic, 1993). Dopaminergic modulation can weaken or strengthen neural connections when changes in dopamine levels at the relevant synapse coincide with associated neural activity (Calabresi, Pisani, Mercuri, & Bernardi, 1996; Wickens, Begg, & Arbuthnott, 1996; Wickens & Kotter, 1995). In some cortical regions, particularly the prefrontal cortex, the impact of DA on cortical processing is similar to NE, increasing the signal-to-noise ratio of target neural networks (Durstewitz & Seamans, 2008; Servan-Schreiber et al., 1990). The dual-state theory of DA system function (Durstewitz & Seamans, 2008) distinguishes between the effect of DA on two classes of dopamine receptors. D1 receptors are preferentially activated by moderate concentrations of DA (<500 nmol/L), and produce an enhancement of the signal-to-noise ratio in prefrontal cortex that stabilizes working memory representations in the face of noise or interference. D2 receptors are preferentially activated by low or high concentrations of DA(>1μmol/L, < 10 nmol/L), and reduce the signal-to-noise ratio in target networks to allow for flexible integration of new information when the contents of working memory need to be updated.

Single cell recordings from primates show that DA neurons exhibit phasic firing in response to rewards or reward-predicting stimuli (see Fig. 2), and exhibit decreases in firing in response to aversive events, such as pain (e.g., Schultz, 1998, 2002). Phasic bursts of DA release in response to reward tend to be elicited within 50 ms to 110 ms after onset of the eliciting stimulus, and last approximately 200 ms or less (Schultz, 2002). Depressions in the firing of DA neurons in response to pain tend to occur with similar latency as to rewards, but with longer duration (Schultz, 2002; Schultz & Romo, 1987). In conditioning paradigms, before a task is learned DA neurons fire at

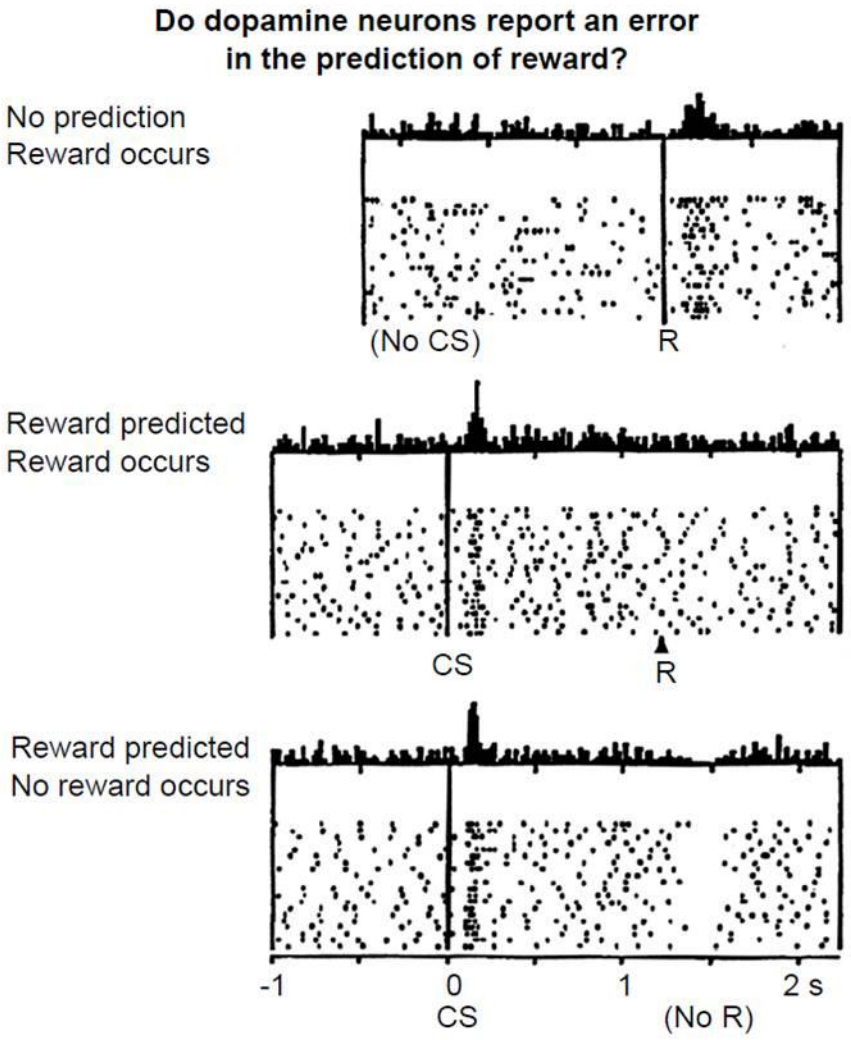

Figure 2. Dopamine neuron activity recorded from monkeys in a visual discrimination task. The y-axis indicates the sequence of trials plotted from top to bottom, and the x-axis represents time. R stands for onset of reward, and CS stands for presentation of a conditioned stimulus. Each dot represents an action potential. Before conditioning, a R produces an increase in DA-neuron firing (top). Once the animal has learned to predict the R from the CS, the CS but not the R produce increased DA-neuron firing (middle). If the CS is presented without the predicted R, a dip in DA-neuron firing is observed (bottom). This figure was taken from Schultz, Dayan, and Montague (1997).

reward presentation, whereas after a task has been learned the DA neurons fire at presentation of the conditioned stimulus. At intermediate points in learning, DA neurons fire at both the conditioned stimulus and the reward. Furthermore DA neurons show decreased firing from baseline activity at the time of the expected reward if the reward is not provided (Schultz et al., 1993; Schultz

2002). Similarly, the probability of a reward (based on previous experience with a task or conditioned stimulus) impacts the degree to which DA neurons respond to the presentation or lack of presentation of a reward (Fiorillo, Tobler, & Schultz, 2003; Romo & Schultz, 1990; Schultz, 2002). Fiorillo et al. presented monkeys with distinct visual stimuli that indicated the probability of a subsequent reward. The probabilities varied from 0, to .25, .5,.75, and 1. DA activity in response to the stimuli increased as probability of reward increased, whereas DA activity in response to reward decreased as probability of reward increased. Furthermore, suppression of DA firing in response to the absence of reward was greatest when the probability of reward was high. These observations indicate that DA neurons respond to rewards contingent on the disagreement between the actual and predicted reward outcome (Schultz, Dayan, & Montague, 1997).

Schultz and colleagues (1997) argued that the firing patterns of DA neurons in response to rewards operated like temporal difference (TD) error signals, a concept that originated as a means by which engineered systems could learn to predict action outcomes (Sutton, & Barto, 1981; 1990; see also Rescorla & Wagner, 1972). TD errors are computed as the difference between the experienced "value" of ongoing events and the predicted value of those events. A positive TD error indicates that an event has greater value than originally predicted, whereas a negative TD error indicates that an events has less value than predicted. The TD hypothesis of the DA system was presented with a computational model that successfully demonstrated a diverse pattern of animal behaviour in conditioning paradigms, including "blocking," the phenomenon in which an animal will not learn the relationship between a new predictor and a reward if the new predictor is redundant to a previously learned predictor. Furthermore, the structure and implementation of the model maps well onto the structure and function of the DA system (Holroyd, 2001; Holroyd & Coles, 2002). Under the TD hypothesis of the DA system, the value of an event is computed in the basal ganglia, and the TD

error is computed either in the basal ganglia or in the DA-releasing neurons themselves, which refer the TD signal via dopaminergic projections to target structures, including the ACC.

The ACC

Neuroimaging investigations of ACC function have implicated the ACC in decision making and the exertion of cognitive control (e.g., Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Holroyd & Coles, 2002; Rushworth, Behrens, Rudebeck, & Walton, 2007; Holroyd & Yeung, in press). The ACC is activated by error commission or by feedback indicating a failure to obtain reward (e.g., Debener, Ullsperger, Siegel, Fiehler, von Cramon, & Engel, 2005; Holroyd, Nieuwenhuis, Yeung, Nystrom, Mars, & Coles, 2004). Lesion studies with monkeys have also demonstrated a critical role for the ACC in learning specifically from rewards (e.g., Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006).

However, the ACC is also activated by infrequent, unexpected or conflict-inducing events, where conflict is defined as the simultaneous activation of competing neural processes (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Braver, Barch, Gray, Molfese, & Snyder, 2001; Carter, Braver, Barch, Botvinick, Noll, & Cohen, 1998; Clark, Fannon, Lai, Benson, & Bauer, 2000; Crottaz-Herbette & Menon, 2006; Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003; Kiehl, Laurens, Duty, Forster, & Liddle, 2001; Ullsperger & von Cramon, 2001; van Veen, Cohen, Botvinick, Stenger, & Carter, 2001; van Veen & Carter, 2002a, 2002b; Yeung et al., 2004; Yeung & Cohen, 2004). This second set of findings has led to the conflict-monitoring hypothesis of ACC function (Botvinick et al., 1999), that the ACC monitors for response conflict and recruits additional control mechanisms to resolve that conflict.

The topic of ACC function is controversial but there is general consensus that the ACC plays a central role in behavioural control. A new theory of ACC function has recently been proposed that integrates key characteristics of four major classes of ACC theories (Holroyd

&Yeung, in press). The hierarchical reinforcement learning theory of ACC function holds that the ACC is involved in motivating extended sequential behaviours and that DA facilitates this function (Holroyd & Yeung, in press).

Catecholinergic modulation and the event-related potential

The event-related brain potential (ERP) technique is particularly well suited to investigating catecholinergic modulation because of its high temporal resolution. In particular, phasic bursts of NE are thought to impact processing in the cortex within 170 ms of onset of the eliciting event (Aston-Jones & Cohen, 2005), and the impact is thought to only last approximately 180 ms (Usher et al., 1999). Such a brief change in activity (that involves both a brief enhancement but also a subsequent dampening of activity due to the LC-NE system's refractory-like period) would likely not be detected with imaging techniques that operate on longer timescales such as functional magnetic resonance imaging (fMRI). Hence the ERP technique provides a means for assessing changes in neural activity in real time. Further, ERP components are mainly believed to be generated by cortical activity, and the broad efferent projection schemes of the LC-NE and DA systems to cortical areas suggest that phasic NE and DA releases should modulate cortical areas with a neural organization most suitable for producing changes in the ERP. For example, Holroyd and Coles (2002) note that layer 5, "gigantopyramidal" neurons in the ACC are structured particularly well to contribute to the EEG. Furthermore, the putative impact of NE and DA on neuronal responsivity suggest their action should produce detectable changes in the ERP.

Relevant ERP Components

In discussing the electrophysiological impact of catecholinergic modulation, I will refer to several ERP components. It is useful to give a general overview of these components. In response to any attended stimulus, the ERP will display several positive and negative deflections. Among these are the P2, N2, and P3, which get their names from being the second positive deflection,

second negative deflection, and third positive deflection after onset of the eliciting stimulus, respectively. Each of these deflections can be categorized into specific variants, or components, defined by the scalp distribution and the experimental manipulation that exercises their amplitude. For example, the N2 is the second negative deflection elicited by a task-relevant stimulus, whereas the conflict N2 is a modulation of the N2 by conflict-inducing events that is maximal over frontal-central electrode channels. In addition, some ERP components are defined as a difference between two conditions: The component is revealed in a difference wave, where the "raw" ERP from one condition is subtracted from another to isolate the difference in the ERP specifically due to the manipulation. Examples are the error-related negativity (ERN) and the feedback-related negativity (FRN), constructed by subtracting the ERP to positive events from the ERP to negative events, revealing a prominent negative difference over frontal-central electrodes approximately 200 ms to 300 ms after the eliciting stimulus. It is important to distinguish between raw ERP components, and difference-wave ERP components, as I will discuss further below.

The LC-P3 Theory

The LC-NE system has been associated with an ERP component known as the P3 (Nieuwenhuis, Aston-Jones, & Cohen, 2005a; Nieuwenhuis et al., 2005b; Pineda, Foote, & Neville, 1989). The P3 is a positive deflection typically peaking around 300 ms. It exhibits a broad scalp distribution that is maximal over parietal cortex and is thought to represent the summation of activity from widely distributed areas (e.g., Johnson, 1993; Kok, 2001). The P3 has been strongly associated with the successful updating of working memory, or "context updating" (Donchin & Coles, 1988).

The LC-P3 theory (Nieuwenhuis et al., 2005a) proposes that the arrival of NE in the cortex and the resulting change in processing dynamics is manifested electrophysiologically at the scalp as the P3. In support of this hypothesis, Nieuwenhuis et al. (2005a) presented a comprehensive review

of the literature to support the LC-P3 theory, marshalling abundant evidence that phasic LC activity shares antecedent conditions with the P3. They characterize four important mediating antecedent events that the P3 has in common with LC phasic activity: Subjective probability (infrequent events elicit a larger P3 than frequent events), motivational salience (targets elicit a larger P3 than distracters), applied attention (attended stimuli elicit a larger P3 than ignored stimuli, and targets elicit a larger P3 under conditions that demand full attention compared to dual-task conditions), and attention-capturing stimuli (task-irrelevant stimuli that are highly deviant from the stimulus context elicit a larger P3 than less deviant stimuli). Additionally, Nieuwenhuis and colleagues. (2005a) refer to psychopharmacological and animal lesion studies in support of the link between the P3 and noradrenergic modulation. Generally speaking, noradrenergic agonists such as clonidine and direct lesions of the LC have been reported to reduce the amplitude of a P3-like potential observed in monkeys (e.g. Pineda, Foote, & Neville, 1989; Pineda & Westerfield, 1993; Swick, Pineda, & Foote, 1994; Swick, Pineda, Schacher, & Foote, 1994). Further, in one psychopharmacological study involving human participants, Halliday et al. (1994) found that clonidine reduced the amplitude of the P3.

Issues with the LC-P3 Theory

The LC-P3 theory possesses considerable explanatory power and accounts for a wide range of existing data. However, at least three issues need further explication. First, the LC-NE system is auto-inhibitory, such that phasic bursts of NE are followed by a refractory-like period lasting approximately 500 ms characterized by reduced or arrested NE supply to the cortex. If enhanced signal detection produces electrophysiological changes detectable at the scalp (the P3), then degraded signal detection associated with reduced NE supply should presumably likewise be visible in the ERP. However, no such deflection is visible in the ERP following the P3.

Second, the attentional blink, a deficit in target processing attributed to the LC refractory period (e.g. Nieuwenhuis et al., 2005a, 2005b; Usher et al., 1999; Warren et al., 2009), typically lasts from 200 to 600 ms, corresponding closely to the temporal profile of the P3 (Fig. 3).

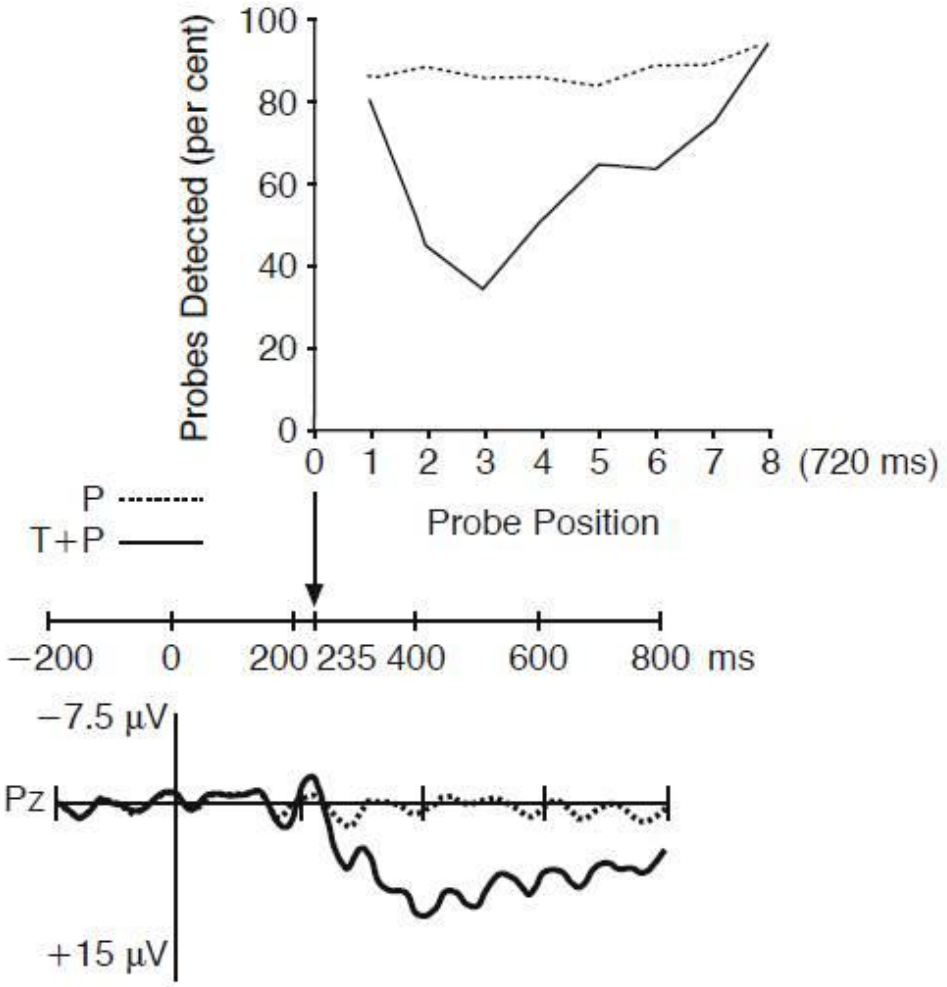

Figure 3. Accuracy and electroencephalographic data from an attentional blink study. The two plots compare second-target accuracy over time (top plot) with the P3 elicited by the first target (bottom plot). P stands for probe target (second target), and T stands for the first target, which was present in the "T+P" condition, but not in the "P" condition. When the onset of the first target in the accuracy plot is aligned with the peak of the N2 in the ERP plot (see arrow), accuracy correlates with P3 amplitude across time. This figure was taken from MacArthur, Budd, and Mitchie (1999).

Nieuwenhuis et al. (2005a, 2005b) proposed that the onset of the first target in the AB task elicits a LC-NE system phasic response, with the subsequent flood of NE to the cortex benefitting

processing of the first target, and also the second target if the second target is presented within approximately 100 ms of the first (lag-1 sparing). After this initial period of effective target processing (~200 ms from onset of the first target to offset of a second target presented 100 ms later), the LC is inhibited and cortical levels of NE are not sufficient to effectively process any targets presented 200 ms to 600 ms after onset of the first target. This description of the timing of the attentional blink relative to the LC-NE system response raises the question: Why does the timing of the P3 correspond so closely with the timing of the attentional blink, rather than the timing of effective processing preceding the attention blink, as it should if it were an electrophysiological manifestation of effective target processing? Adding to this seeming contradiction, MacArthur, Budd, and Mitchie (1999) showed a significant negative correlation between the amplitude of the first-target P3 and second-target accuracy, such that the size and temporal profile of any subject's attentional blink mirrored the size and temporal profile of that subject's first-target P3, highlighting the close relationship between the time course of the P3 and the time course of the attentional blink.

The fact that the P3 occurs during the attentional blink, rather than before it, may not be as large a discrepancy as it first appears because the relative timing of the P3 and the attentional blink should be corrected for the time it takes target-related information to reach the cortex (MacArthur et al., 1999; Nieuwenhuis et al., 2005a, 2005b). MacArthur and colleagues (1999) estimated that it takes about 235 ms for information related to a stimulus to reach the decision stage in cortex that determines whether that stimulus is a target. Thus, electrophysiological activity 235 ms after onset of the target reflects neural processing at that critical decision stage. Therefore, the neural processing relevant to the identification of a target presented at time 0 ms is exhibited at time 235 ms in the ERP, and 335 ms for a target presented at time 100 ms, and so on. Following from this, MacArthur and colleagues calculated the correlation between P3 amplitude and second-target accuracy across time after adjusting the timing of the ERP relative to the attentional blink to align time 0 ms in the

attentional blink with time 235 ms in the ERP "to account for the propagation delay between probe stimulus onset and the arrival of the signal at the cortex, and [their] best estimate of the latency of neural processing underlying probe discrimination" (p. 3692). This correction was done assuming the P3 is a manifestation of a processing *deficit* that produces the attentional blink, but I suggest the correction should have been smaller. Aston-Jones and Cohen (2005) maintain that the timing of NE arrival in cortex (~170 ms after target onset) is effective for facilitating a behavioral response to the eliciting stimulus, including "internal" responses such as coding the stimulus into memory. Further, examination of Figure 3, taken from MacArthur and colleagues (1999) suggests the correlation between the temporal profile of the attentional blink and of the P3 would be stronger if they had used an adjustment of 170 ms (onset of the N2) instead of 235 ms (peak latency of the N2). In contrast, in assuming the P3 is a manifestation of facilitated processing that *precedes* the attentional blink deficit Nieuwenhuis et al. (2005b) implicitly suggest such a correction should be larger. If the peak of the P3 (~ 400 ms in Fig. 3) reflects the peak of noradrenergic modulation, the peak should correspond to processing of the target exhibiting the highest accuracy, typically the lag-1 target at 100 ms. This suggests a correction of 300 ms is required to map the P3 onto the processing enhancement that precedes the attentional blink, conspicuously longer than the estimates I have cited of how long it should take target-related information to reach the cortex.

Not only the latency, but also the duration of the P3 is inconsistent with the timing of the attentional blink. A conservative estimate of the duration of the P3 in the attentional blink paradigm is about 300 ms to 400 ms (e.g., MacArthur et al., 1999), much longer than the period of effective target processing preceding the attentional blink deficit (~200 ms). It may be possible to reconcile the theory with this discrepancy – for example, Nieuwenhuis and colleagues. (2005b, p. 516) speculated that the duration of the P3 may be longer than NE facilitation due to "the slow speed of electronic conduction of current from the active zone at the soma through the apical dendrites of

cortical pyramidal neurons," or that the "depolarization (associated with the P3 potential) may outlast the increased responsivity of pyramidal neurons to extrinsic inputs due to recruitment of inhibitory interneurons or other unknown mechanisms." However, MacArthur et al.'s position, that the P3 reflects a processing deficit rather than the period of effective processing that precedes the deficit does not require additional explanation.

Finally, the NE phasic burst appears to arrive in cortex too early to be manifested by the P3. Aston-Jones and Cohen (2005) suggest NE should reach the cortex within approximately 170 ms of target onset, but the P3 typically doesn't begin for another 60 ms until about 230 ms following target onset (if onset of the P3 is taken as the peak latency of the preceding negativity, the N2). Although this estimate of the timing of NE arrival is based on single-cell recordings in monkeys, Aston-Jones, Foote, and Segal (1985) demonstrated that conduction speeds in NE-releasing neurons vary across species such that the actual timing of NE arrival in cortex is relatively preserved despite varying axonal distances. Further, P3 onset sometimes occurs after motor response initiation suggesting that the underlying mechanism does not directly implement the stimulus-response mapping but rather is involved in a subsequent, related process (Duncan-Johnson, & Donchin, 1982; Krigolson, Holroyd, Van Gyn, & Heath, 2008; Ritter, Simson, Vaughan, & Friedman, 1979). In contrast, phasic LC activity consistently precedes behavioural responding and has been strongly associated with processes that lead to the response (e.g., Clayton, Rajkowski, Cohen, & Aston-Jones, 2004; Rajkowski, Majczynski, Clayton, & Aston-Jones, 2004).

Nieuwenhuis and Jepma (2010) suggested two possible explanations of this temporal discrepancy. First, they argued that overt behaviours can occur despite minimal NE prior to the arrival of an LC phasic release, such that the P3 need not be generated before the response. Although this is a reasonable allowance to consider, support for this possibility is not evident in single cell recording studies of the LC in animals (e.g., Clayton et al., 2004; Rajkowski et al., 2004).

Second, they have pointed out that even when the peak of the P3 occurs after the behavioural response, the *onset* of the P3 often occurs before the response, suggesting that some part of the response generation process may have benefitted from NE modulation. However, this argument does not appear to be true for the results from Krigolson and colleagues (2008), in which the corrective movement was initiated even before onset of the P3 (Fig. 4). Nevertheless,

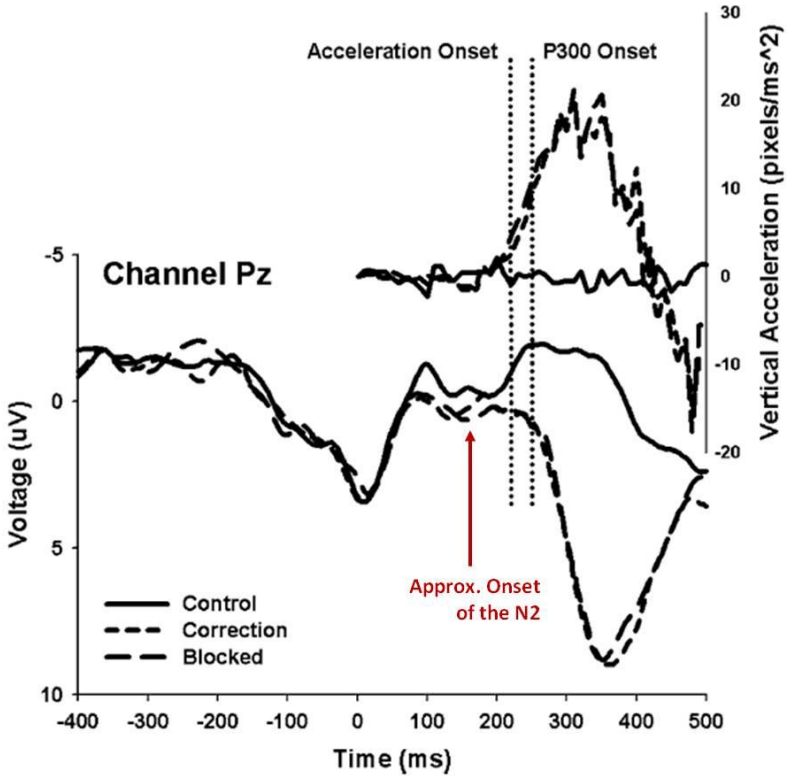

Figure 4. Comparison of the onset of a movement correction (top) with the timing of onset of the P3 and the N2 (bottom). Subjects were required to move an on-screen cursor from a starting point to a target, and the target could either remain stationary (Control), change location immediately after subjects began to move the cursor, forcing them to correct their aim (Correction), or the target could move but the cursor would not respond to a correction (Blocked). The comparison of the timing of movement associated with a correction/attempted correction (Vertical Acceleration) and the timing of associated ERPs reveals the onset of the corrective movement preceded onset of the P3, but followed onset of the N2. Adapted from Krigolson, Holroyd, Van Gyn, and Heath (2008).

Nieuwenhuis holds that the P3 is comprised of both the P2 (the positive deflection preceding the N2) and the P3, with the N2 appearing as a "dent" on the rising flank (Nieuwenhuis, personal

communication). This position ameliorates the timing issues I have raised in that it moves the putative onset of the NE arrival approximately 100 ms forward in time, which is more consistent with the relative timing of behavioural responses and also with observations from single-cell recordings in monkeys. On the other hand, this position widens the temporal window of NE facilitation by approximately 100 ms (to about 450 ms) and thus is even more inconsistent with the duration of effective target processing preceding the attentional blink, as discussed above. Additionally, the view that the P2 is part of the P3, though compelling, does not appear to be widely held as researchers typically analyze the two ERP phenomena as distinct ERP components (e.g., Luck & Hillyard, 1994; Rozenkrantz, & Polich, 2008; Simson et al., 1976).

The Modified LC-P3 Theory

The P3 is typically preceded by a negative deflection in the ERP known as the N2. The N2 typically peaks between 200 ms and 300 ms after onset of the eliciting stimulus. There are several variants of the N2 that are defined according to their specific antecedent events and variations in scalp distribution (for reviews see Pritchard, Shappell, & Brandt, 1991; Folstein & Van Petten, 2007). For example, the "mismatch N2" is observed when comparing the presentation of a stimulus that was preceded by exactly the same stimulus versus a different stimulus (mismatches produce a larger N2), is distributed frontal centrally, and in the auditory modality can be elicited when subjects are not attending to the stimuli (for a review see Naatanen, 2007). The "oddball" N2 is a larger N2 in response to the presentation of a stimulus that is representative of a relevant, seldom-presented category, such as a male face when 80% of stimuli are female faces. The oddball N2 has a frontal central scalp distribution in the auditory modality, but in the visual modality has been observed sometimes with a frontal-central distribution (e.g., Holroyd et al., 2008; Nieuwenhuis et al., 2003) and sometimes with a more posterior distribution (Folstein & Van Petten, 2007; Luck & Hillyard, 1994; Ritter, Simson, & Vaughan, 1983). The "novelty" N2 is elicited by stimuli that are entirely

unexpected in the experimental context, such as a picture of a rainbow appearing within a set of numerical stimuli when the task is for the subject to count prime numbers. These novel stimuli produce a larger N2 that is frontally distributed. Finally, the frontal-central "conflict" N2 has been identified as having an amplitude exercised by infrequent, unexpected, or conflict inducing events (e.g., Nieuwenhuis et al., 2003; Yeung, Botvinick, & Cohen; 2004).

In early studies of the impact of stimulus probability on the ERP, the N2 and P3 were collectively termed the N2/P3 complex because of their tendency to co-vary in amplitude and latency (e.g. Duncan-Johnson & Donchin, 1977; see also Ritter et al., 1979), and because principal component analyses (Donchin & Heffley, 1978) of ERP data indicated that the two deflections load on a single temporal factor (Rohrbaugh et al., 1978, 1979; Squires, Donchin, Herning, & McCarthy, 1977; Squires, Squires, & Hillyard ,1975). In fact, all of the antecedent conditions noted by Nieuwenhuis and colleagues to apply to both P3 amplitude and LC phasic activity also apply to N2 amplitude. For example, both N2 amplitude and P3 amplitude increase with increasing unexpectedness of an event, and both are larger to targets than non-targets (e.g. Courchesne, Hillyard, & Galambos, 1975; Duncan-Johnson & Donchin, 1977; Nieuwenhuis et al., 2003; Ritter et al., 1979; Simson, Vaughan, & Ritter, 1976; Squires, Squires, & Hillyard, 1975; Squires, Wickens, Squires, & Donchin, 1976). Both also scale to the amount of attention paid to a stimulus, with a larger N2 and larger P3 to attended versus unattended stimuli, and to attention-capturing/highly deviant stimuli versus less deviant stimuli (Courchesne, Hillyard, & Galambos, 1975; Daffner, Mesulam, Scinto, Calvo, Faust, & Holcomb, 2000; Daffner, Scinto Calvo, Faust, Mesulam, West, & Holcomb, 2000; Folstein, van Petten, & Rose, 2008; Ford, Roth, & Kopell, 1976; Hillyard, Squires, Bauer, & Lindsay, 1971; Squires et al., 1977; Squires et al., 1975). Thus, it is equally possible that the phasic burst of NE that enhances signal detection in the cortex produces the N2, whereas the subsequent depletion of NE due to the LC refractory period produces the P3 (see Fig. 5).

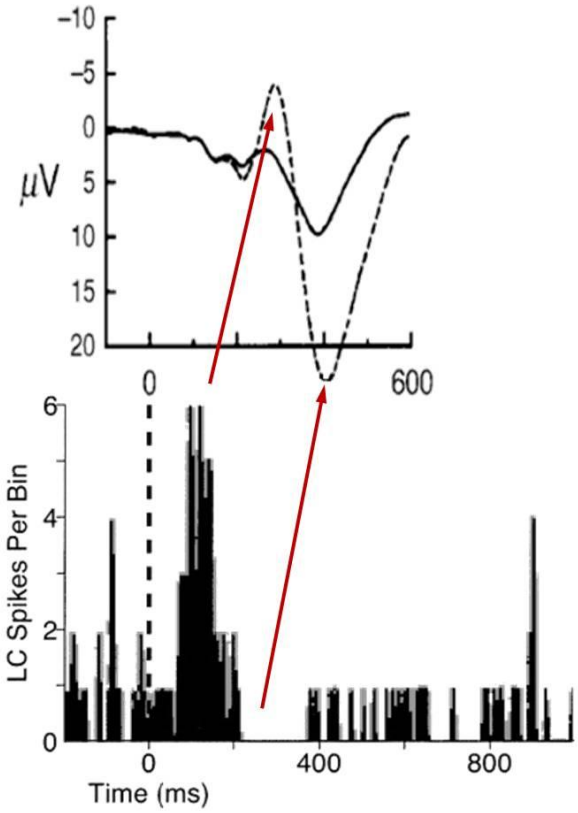

Figure 5. Comparison of the time course of a typical N2-P3 complex (top) amplified to infrequent events (dotted line) relative to frequent events (solid line) (negative is plotted up), and the typical time course of LC-NE system phasic activity (bottom). Note that the peak of the N2 occurs approximately 180 ms after the peak of LC firing, and the peak of the P3 occurs approximately 180 ms after complete inhibition of the LC. LC-NE system figure adapted from Usher, Cohen, Servan-Schreiber, Rajkowski, and Aston-Jones (1999); N2-P3 figure adapted from Nieuwenhuis, Yeung, van den Wildenberg, and Ridderinkhof (2003).

This idea modifies the LC-P3 theory by proposing that the LC refractory period, rather than the LC burst, produces the P3. Note that according to this account the duration of the refractory period of the LC is directly related to the size of the initial NE burst such that the electrophysiological manifestation of the refractory period (the P3) covaries in size with the electrophysiological manifestation of the phasic burst (the N2); thus this modification of the theory naturally accounts for the observed relationship between the N2 and the P3.

The modified LC-P3 theory also eliminates each of the issues I described with the original LC-P3 theory. First, the modified LC-P3 theory provides an ERP correlate of both NE abundance due to a phasic release and NE depletion due to the subsequent refractory-like period of the LC. Second, the modified LC-P3 theory accounts for the close relationship between the temporal profile of the P3 and the temporal profile of the attentional blink (MacArthur et al., 1999). The correspondence reflects the real relationship between the electrophysiological manifestation (the P3) and the cognitive manifestation of the NE depletion (the attentional blink), whereas the timing of the N2 should correspond to the timing of effective target processing that precedes the attentional blink. Examination of Figure 1 (adapted from Usher et al., 1999) suggests that the phasic burst of NE has a shorter duration than the refractory like period, just as the N2 has a shorter duration than the P3, and furthermore the duration of the N2s observed in this work (see Figs. 6, 8-11) is approximately equal to Usher et al.'s assessment of the NE-mediated effective target processing that precedes the attentional blink (180 ms). Thus, the modified LC-P3 theory holds that for a brief moment following the eliciting stimulus, cortical NE is plentiful, target processing is effective and the N2 occurs, followed by a longer period of depleted cortical NE, impaired target processing (the attentional blink) and P3 generation.

Finally, the modified LC-P3 theory better aligns the timing of the putative ERP response to LC-NE activity (the N2, onset at about 180 ms post-stimulus) with the actual timing of the phasic LC-NE signal (at about 170 ms post-stimulus). Furthermore, the N2 more consistently precedes behavioural responding than does the P3 (e.g. Krigolson et al., 2008), and there is evidence N2 peak latency is more tightly correlated to reaction time than P3 peak latency (Ritter et al., 1979).

The reinforcement learning theory of the ERN and FRN

The error related negativity (ERN) was discovered independently by Falkenstein, Hohnsbein, Hoormann, and Blanke (1991) and Gehring, Coles, Meyer, and Donchin, (1990), and given its name

by the Gehring et al. group (see also Gehring, Goss, Coles, Meyer, & Donchin, 1993). It is observed by comparing the ERP for error trials with the ERP for correct trials in speeded response time tasks. Error trials show a negative deflection peaking around 80 ms after commission of the error. Miltner, Braun, and Coles (1997) later discovered that when the ERP to reward feedback is subtracted from the ERP to no-reward feedback, the difference wave reveals a negative deflection approximately 200 ms - 300 ms after onset of the eliciting feedback, termed the feedback-related negativity (FRN). They argued that the ERN and FRN were different instances of the same ERP component elicited by a generic system for error processing.

The error-monitoring theory of the ERN/FRN was substantially extended by Holroyd and Coles (2002) who linked the ERN and FRN to the activity of the DA system. The reinforcement learning theory of the ERN/FRN (RL-ERN theory) holds that the ACC is responsible for developing intentions into actions by selecting from multiple response options on the basis of a reinforcement history implemented by the DA system. The ACC acts as a motor control filter, selecting between different actions that are appropriate for the task at hand according to a history of reinforcement. TD-error DA signals indicating the degree to which an event was better or worse than expected are used by the ACC to adjust the probability of responding with a specific action to a specific event based on reinforcement learning principles. The impact of DA signals on ACC processing are reflected by the ERN/FRN.

The RL-ERN theory was presented with abundant neurophysiological evidence of the role of the ACC in developing intentions into actions (e.g. Bentivoglio, Kultas-Illinsky, & Illinsky, 1993; Morecraft & Van Hoesen, 1998), along with an extensive review of the role of the basal ganglia in reinforcement learning (e.g. Berridge & Robinson, 1998). Holroyd and Coles also provided two computational simulations coupled with behavioural studies examining the ERN and FRN and demonstrated that the ERN and the FRN are the same component. In a probabilistic learning task,

Holroyd and Coles varied the degree to which subjects could be sure an action would result in reward. When subjects were not sure if their response would be rewarded, the ERN at the time of response was smaller than the FRN at the time of the feedback. However, when subjects were sure their response would be rewarded (or not) the response ERN was larger than the FRN, an effect mirrored by the simulation. This result follows from the assertion that the ERN/FRN reflects a reward-prediction error signal carried by the DA system that propagates back in time to the earliest predictor of reward. Thus the ERN is a FRN that has propagated back to the commission of an error, from later (redundant) feedback that an error has been made.

A wealth of research has since been published supporting the RL-ERN theory (e.g. Baker & Holroyd, 2009; Butterfield & Mangels, 2003; Cohen & Ranganath, 2007; Donkers & van Boxtel, 2005; Dunning & Hajcak, 2007; Eppinger, Kray, Mock, & Mecklinger, 2008; Hajcak, Moser, Holroyd, & Simons, 2007; Hewig et al., 2007, 2008; Holroyd & Coles, 2002, 2008; Holroyd & Krigolson, 2007; Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Krigolson & Holroyd, 2007; Krigolson, Pierce, Holroyd, & Tanaka, 2009; Morris, Heerey, Gold, & Holroyd, 2008; Nieuwenhuis, Nielen, Mol, Hajcak, & Veltman, 2005; Nieuwenhuis et al., 2002; Yasuda, Sato, Miyawaki, Kumano, & Kuboki, 2004). One finding that is particularly relevant to the work reported here is that the probability of reward/no-reward impacts ERN/FRN amplitude (just as probability impacts DA system activity as observed in the 2003 paper by Fiorillo and colleagues). When a probability manipulation is added to the standard reward/no-reward FRN paradigm, the FRN constructed by subtracting the ERP to infrequent rewards from the ERP to infrequent no-rewards produces a larger "difference wave FRN" (dFRN) than the dFRN constructed from frequent rewards and no-rewards (e.g., Holroyd & Krigolson, 2007). Note that though the ERN and FRN are typically assessed in the difference wave, the names themselves suggested they are driven by a change in the raw ERP to errors or negative feedback, even though

positive feedback may also contribute to the difference. I therefore use the terms "dFRN" and "dERN" to refer to the difference-wave components, and will reserve the terms ERN and FRN to refer to the negative deflections in the raw ERP to error or negative feedback.

Relationship between the FRN and the N2

Commonalities between the ERN and FRN and the conflict N2 have provoked questions as to whether they are really distinct components. Holroyd (2004) noted the similar timeframe, topology, and polarity of the FRN and the conflict N2 and suggested that they may be manifestations of a common underlying mechanism that is affected by both stimulus probability and valence. Yeung and colleagues (2004) further noted that the ERN and the conflict N2 have reconcilable antecedent events and suggested that they are the same component. These considerations warrant close examination of these ERP components as they relate specifically to reinforcement learning on the one hand, and to conflict detection/resolution on the other.

The RL-ERN theory originally emphasized that unexpected negative events produce the dERN and dFRN, but recent research by Holroyd and colleagues (e.g. Baker & Holroyd, 2011; Eppinger et al., 2008; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Hosseini & Holroyd, in press) suggest the dFRN is driven more by unexpected reward than by unexpected errors. According to this position, a standard N2 deflection is produced in response to the task-relevant event, but rewards also elicit a "reward positivity" that is superimposed over the N2, attenuating its manifestation at the scalp (Baker & Holroyd, 2011; Holroyd et al., 2008). Hence, the dFRN reflects variance in the ERP associated with both the N2 to error feedback and the reward positivity to positive feedback. However, in addition to the reward positivity elicited by positive feedback, the RL-ERN theory proposes that dips in DA in response to negative feedback should amplify the size of the N2. To date there has been little if any experimental work that indicates that N2 amplitude is enhanced by negative feedback.

The role of the conflict N2 in the modified LC-P3 theory

The modified LC-P3 theory benefits from the strengths of the original theory while resolving three apparent discrepancies. At the same time, this view contrasts with a current theory of the origin of the N2 that, based mainly on the results of ERP and fMRI studies, holds that conflict detection by the ACC produces the N2 (ACC-N2 theory) (Botvinick et al., 2001; Yeung, et al., 2004). Conflict is typically elicited in paradigms such as the Eriksen flanker task (Eriksen & Eriksen, 1974) in which incompatible stimuli evoke activity associated with both the correct response to the central character and the incorrect response associated with the distracting flankers. In addition, according to this position infrequently occurring target stimuli in speeded response tasks elicit conflict between the erroneous response associated with the frequently occurring stimulus and the correct response associated with the infrequently occurring stimulus (Nieuwenhuis et al., 2003; Yeung et al., 2004). Prominent projections from the ACC to the LC may be responsible for recruiting the LC-NE system to resolve the conflict (e.g. Gilzenrat et al., 2002).

The original LC-P3 and ACC-N2 theories hold that the N2 and P3 co-vary in amplitude because the conflict signal (N2) dictates -- via the projection from the ACC to the LC -- the amount of NE release (P3). The modified LC-P3 theory also holds that the ACC produces the conflict N2 upon the detection of conflict, which in turn stimulates a phasic burst of LC activity. However, the modified LC-P3 theory holds that the NE burst increases the signal to noise ratio of all active cortical areas, including the ACC, resulting in increased ERP amplitudes generated by those areas. Thus the modified LC-P3 theory describes a positive feedback loop: ACC activity produces the N2 and activates the LC, which in turns boosts the activity of the ACC and increases the N2. Furthermore, according to this account the subsequent LC refractory period and concomitant reduction in cortical NE produces the P3.

Although the ACC-N2 and modified LC-P3 theories overlap, a key difference concerns their predictions about the scalp distribution of the N2. Specifically, the ACC-N2 theory holds that changes in N2 amplitude reflect changes in ACC activity specifically, and thus the N2 is (always) maximal at scalp locations dorsal to the ACC, whereas the modified LC-P3 theory predicts that the scalp distribution of the N2 should vary according to task specifics--even within the same general task paradigm (e.g. the oddball task). This latter position follows from the idea that the modulatory impact of NE is greatest on the cortical regions that are most activated by the task at hand (Nieuwenhuis et al., 2005a; Nieuwenhuis, de Geus, & Cohen, 2011). Thus, if the N2 is modulated by the presence of NE, then the *change* in N2 amplitude due to NE should appear maximal over whatever area is working the hardest.

Summary and Objectives

Neuromodulators impact brain function by modifying neuronal responses to other neurotransmitters, thus changing communication between neurons. Both the LC-NE and the DA neuromodulatory systems have been associated with specific ERP components. The LC-P3 theory holds that the P3 is a manifestation of enhanced cortical activity due to noradrenergic modulation, whereas I propose in this thesis that the P3 is a manifestation of dampened cortical activity due to auto-inhibition of the LC, and the preceding N2 reflects the phasic NE release that precedes LC auto-inhibition. A key prediction of the modified LC-P3 theory is that the N2 should demonstrate a scalp distribution that varies according to task specifics because NE differentially impacts cortical activity dependent on the relative level of activity at the time of NE arrival. The RL-ERN theory holds that the dFRN and dERN reflect phasic DA signals being processed in the ACC that carry information relevant to reinforcement learning. Current thinking on this theory is that phasic bursts of DA in response to rewards attenuate the N2, such that when the dFRN/dERN is constructed by

subtracting reward trials from no-reward trials a negative difference is observed. These theories intersect where they characterize or have implications concerning the N2 ERP component.

In five experiments I tested the prediction of the modified LC-P3 theory that the N2 should demonstrate a variable scalp distribution dependent on the task by conducting oddball experiments using faces as stimuli. Previous research suggests face discrimination highly exercises the fusiform face-processing area (FFA) producing a negative deflection in the ERP at lateral occipital electrode sites (e.g. Tanaka, Curran, Porterfield, & Collins, 2006). In my experiments, the faces were either male or female and were tinted either yellow or blue. Within each block one color and one sex occurred infrequently. Between blocks, subjects were instructed to discriminate faces based on the sex of the face or on the color. This design allowed us to isolate the putative LC effects by subtracting the frequent category ERPs from the infrequent category ERPs. Thus, comparison of the difference-wave N2 due to frequency/conflict (hereafter called the "dN2") between the attend-color and attend-sex conditions reveals the impact of NE modulation in each condition. I predicted that discriminating faces based on distinctive colors should be relatively easy, resulting in a dN2 distributed over frontal areas of the scalp as commonly observed in oddball tasks (but see Folstein & Van Petten, 2007, for a different characterization of the *visual* dN2), whereas discriminating faces on the basis of sex should be more difficult and require greater involvement of the fusiform face-processing area (FFA), resulting in a dN2 distributed over lateral-occipital regions. By contrast, other theories of the N2, in particular the ACC-N2 theory, predict no difference in the topology of the dN2 between these two conditions as the frequency manipulation should impact any specific neural generator (i.e. the ACC) identically in both.

In addition, in four of these experiments, I deviated from the typical oddball paradigm by having the target stimuli carry reward information. That is, within each block the category of the tinted face stimuli indicated whether or not a subject had received a reward; depending on condition

either the male faces, the female faces, the blue color or the yellow color indicated that participants received 5 cents for that trial. The inclusion of reward values allowed me to compare simultaneously the way the dN2 and FRN scalp distributions are impacted by attending to either the color or the sex of the stimulus. The RL-ERN theory holds that the FRN is specifically generated in the ACC, whereas the modified LC-P3 theory posits the dN2 has no specific cortical generator. Thus, I predicted that the dN2 would exhibit changes in scalp distribution across task but that the dFRN would not. This result addresses the possibility that a variable scalp distribution observed for the dN2 could be due to a methodological flaw or to a confounding third variable. For example, it is conceivable that a variable scalp distribution is a general characteristic of ERP components in the 200 to 400 ms time range; if this were the case then the result would also be true of the dFRN.

Finally, across these four experiments that included rewards, I manipulated the degree to which subjects were challenged/engaged by the task in a manner that encouraged trying to learn the task to maximize rewards. If the RL-ERN theory is correct in positing that rewards produce a positivity in the ERP, increasing the engagement of learning systems across experiments should simultaneously increase the size of the dFRN (by increasing the size of the reward positivity), and decrease the apparent size of the dN2 (because of component overlap with the reward positivity). Furthermore, the RL-ERN theory holds that the amplitude of the raw N2 should be increased due to no-reward feedback, but empirical evidence for this prediction remains to be demonstrated. I predicted that across Experiments 2 to 5, comparison of the individual ERPs to infrequent no-reward and frequent no-reward conditions will reveal enhanced negativities to no-rewards as subjects are increasingly more challenged/engaged by the task.

Chapter One: Experiment One[2]

Abstract

A prominent theory of the N2 event-related potential component holds that the "oddball" N2 is generated in the anterior cingulate cortex. However, observations of oddball N2s with posterior scalp distributions are inconsistent with this hypothesis. I suggest that variability in the topology of the oddball N2 is a key characteristic of the component that can inform theories of its neural basis. I propose that the oddball N2 reflects cortex-wide noradrenergic modulation of ongoing cortical activity and thus should have a topology that varies systematically according to task specifics. Subjects engaged in an oddball task with male and female faces tinted either yellow or blue, counting targets according to color or sex. Between blocks, targets were frequent or infrequent, counterbalanced across task (attend-color, attend-sex) and category (blue male, yellow male, blue female, yellow female). I created difference waves by subtracting frequent from infrequent category trials to isolate the oddball-N2. When subjects attended to color the oddball N2 was maximal over frontal-central areas and when they attended to sex it was maximal over lateral-occipital areas. Thus, the oddball N2 has a variable scalp distribution that depends on the relative engagement of cortical areas, consistent with noradrenergic modulation having the greatest impact in those areas most engaged by the task at hand.

[2] This experiment has been published: Warren, C. M., Tanaka, J. W., & Holroyd, C. B. (2011). What can topology changes in the oddball N2 reveal about underlying processes? *NeuroReport, 22(17):* 870-874.

Introduction

The N2 is a negative deflection in the human event-related brain potential (ERP) occurring approximately 200-300 ms after an eliciting stimulus. Two prominent reviews have classified the visual N2 into three subcomponents based on antecedent conditions (Folstein & Van Petten, 2007; Pritchard, Shappell, & Brandt, 1991). The "novelty" N2 exhibits an anterior scalp distribution and is sensitive to highly deviant or attention-capturing but task-irrelevant stimuli. The "conflict" N2 exhibits an anterior distribution and is exercised by the exertion of cognitive control, for example, when inhibiting a pre-potent response or when responding to incongruent stimuli as in the Eriksen Flanker Task (Eriksen & Eriksen, 1974). Finally, the "oddball" N2 is modulated by the probability of task-relevant stimulus categories, being largest for infrequently occurring target stimuli; it is said to exhibit an anterior scalp distribution in the auditory modality and a posterior scalp distribution in the visual modality (Folstein & Van Petten, 2007; Pritchard et al., 1991; Ritter, Simson, Vaughan, & Friedman, 1979). However, several research groups have also reported a visual oddball N2 that is maximal at frontal-central sites (e.g., Holroyd, 2001; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Nieuwenhuis, Yeung, Van den Wildenberg, & Ridderinkhof, 2003).

A prominent theory proposes that both the oddball and conflict N2s are manifestations of conflict monitoring/detection in the anterior cingulate cortex and constitute the same ERP component (Nieuwenhuis et al., 2003; Yeung, Botvinick, & Cohen, 2004). Neural conflict is defined as the simultaneous activation of incompatible response channels (Botvinick, Braver, Barch, Carter, & Cohen, 2001). On the incongruent trials in the Eriksen Flanker Task, conflict is elicited by having information relevant to two responses appear simultaneously on the screen. Likewise, in oddball tasks, a frequently appearing stimulus cultivates a bias to respond to the frequent stimulus that must be overcome on infrequent trials. The conflict monitoring theory of the N2/anterior cingulate cortex is supported by source localization studies (Nieuwenhuis et al., 2003; Yeung et al., 2004) and

by functional neuroimaging studies that show increased anterior cingulate cortex activation for events that produce the conflict and the oddball N2 (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Braver, Barch, Gray, Molfese, & Snyder, 2001; Mathalon, Whitfield, & Ford, 2003).

We investigated the discrepancy in reports of oddball N2 topology by conducting a typical oddball experiment using faces as stimuli. The faces were either male or female and were tinted either yellow or blue. Subjects discriminated faces based on the sex or the color of the face. Thus, the stimuli were exactly the same across conditions; only the attended dimension changed. If the oddball N2 scalp distribution is produced by specific cortical generators then it should be the same across task conditions, but if it is sensitive to task demands then it should change accordingly. I predicted that, consistent with previous observations from my own lab and others (Holroyd, 2001; Holroyd et al., 2008; Nieuwenhuis et al., 2003), the oddball N2 in the color condition would exhibit an anterior scalp distribution. By contrast, based on previous observation of the N250 -- a negative deflection in the ERP that is maximal over lateral-occipital areas and believed to be generated by the "fusiform face area" (Tanaka, Curran, Porterfield, & Collins, 2006) -- I predicted that the oddball N2 in the sex condition would exhibit a lateral-occipital scalp distribution. I argue that the variability of the topology of the oddball N2 is a key characteristic of the component that suggests the involvement of the locus coeruleus-norepinephrine (NE) system in its production.

Method

Participants

Twenty-six people (19 females, mean age 20.6 years, *SD* = 2.8 years) signed up through the research participation system at the University of Victoria, Canada, and were compensated with extra credit in an undergraduate psychology course. The study was approved by the human subjects review board at the University of Victoria and conducted in accordance with the ethical standards prescribed in the 1964 Declaration of Helsinki.

Apparatus and Procedure

Participants were seated comfortably, approximately 50 cm in front of a computer screen, in an electromagnetically shielded booth. Stimuli consisted of male or female faces (40 examples of each, excluding hair and contour of head) tinted either blue or yellow (~4.4º visual angle). For both stimulus dimensions (color, sex), one stimulus type occurred infrequently (20% of all trials). This manipulation was counterbalanced across task conditions such that the unattended dimension occurred with 20/80% frequency within each level of the attended dimension. The order of stimulus presentation was randomized with replacement, and the rate of presentation was one item every 1500 ms. Subjects were required to count target stimuli silently. The task consisted of 8 blocks of 100 trials each, counterbalanced such that each of the four stimulus types occurred in two blocks as the target, and of those two blocks, once as a frequent target and once as an infrequent target. Subjects reported their count every fifty trials (halfway through the block) using the keyboard.

Data Acquisition

The electroencephalogram (EEG) was recorded from 41 electrode locations arranged in the standard 10-20 layout using Brain Vision Recorder software (Version 1.3, Brainproducts, Munich, Germany). During recording, the EEG data were referenced to the average voltage across channels, sampled at 250 Hz, and amplified (Quick Amp, Brainproducts) and filtered through a passband of 0.017-67.5 Hz (90 dB octave roll off). Impedances were below 12 kΩ.

EEG Data Analysis

The EEG data were filtered off-line through a 0.1-20 Hz passband phase-shift-free Butterworth filter and re-referenced to linked mastoids. Ocular artifacts were removed using the algorithm described in Gratton, Coles, and Donchin (1983). Trials in which the change in voltage at any channel exceed 35 μV per sampling point were removed. In total, .4% of the data were discarded. 1200 ms epochs of data were extracted from the continuous EEG from 200 ms before stimulus

onset to 1000 ms after. The data were baseline-corrected according to the average amplitude of the EEG over the 200 ms preceding stimulus presentation and ERPs were created by averaging the EEG data for each condition, electrode site, and participant. To isolate the effect of frequency while holding the effect of attentional demands constant, I first constructed four ERPs by averaging the data associated with all the infrequent stimulus presentations (infrequent targets and infrequent non-targets) together, and the data associated with all the frequent stimulus presentations (frequent targets and frequent non-targets) together, separately for each task condition. I then created difference waves by subtracting the frequent-stimulus ERP from the infrequent-stimulus ERP, yielding attend-sex and attend-color oddball N2s that were purely sensitive to frequency of occurrence. This method has previously been prescribed for assessing the oddball N2 without confounding it with target status (e.g., Folstein & Van Petten, 2007; Pritchard et al., 1991).

Visual inspection of the grand average difference waves (Fig. 6) suggested that the oddball N2 varied in latency and duration across the attend-sex and attend-color conditions, with the N2 occurring later and lasting longer in the attend-sex condition. For this reason, I assessed mean oddball N2 amplitude in time windows adjusted for each condition according to the approximate period of oddball N2 manifestation, as follows. For both the attend-color and attend-sex conditions, the location of maximum amplitude was obtained from the grand average ERP by averaging the data recorded at each electrode channel within a temporal window post-stimulus (attend color, 220 - 300 ms; attend-sex, 240 - 420 ms) and identifying the channel with the most negative value. Attend-colour and attend-sex oddball N2 amplitudes were then assessed for each subject by averaging the difference-wave values within the corresponding temporal windows at these two channels.

I conducted a secondary analysis of the impact of frequency on "raw" N2 amplitude. For this secondary analysis, I first determined the raw P3 peak (because it is the most prominent component in the ERP) as the most positive point in the ERP occurring between 210 ms and 800

ms in the attend-color condition, and between 310 ms and 800 ms in the attend-sex condition. I then defined the raw N2 peak as the most negative deflection preceding the raw P3 peak, and occurring no earlier than 200 ms in the attend color condition, and 300 ms in the attend-sex condition. To determine the topology of the difference in raw N2 peak amplitude, I subtracted the frequent raw N2 peak amplitude value from the infrequent value for each electrode and task condition.

Results

Behavioural Results

We assessed error rate by dividing the absolute value of the difference between the reported and correct counts by the correct count, giving a percentage of the correct count. I excluded the data of 2 out of 26 participants who exhibited error levels more than two standard deviations above the mean, corresponding to error rates of 39.0% and 40.5%. For the remaining subjects, the mean error rate for categorizing faces was 14.6% whereas for color it was 5.2%. This difference in performance between categorizing faces based on sex versus color was significant, $t(23) = 5.1$, $p < .001$.

EEG Results

The attend-sex oddball N2 was maximal at lateral-occipital channel P8 (mean amplitude = -.89 μV), and the attend-color oddball N2 was maximal at frontal-central channel C4 (mean amplitude = -2.31 μV). Conversely, the mean amplitude of the attend-sex oddball N2 at channel C4 was -.29 μV and the mean amplitude of the attend-color oddball N2 at channel P8 was -1.47 μV. Figure 6 shows the ERPs elicited by the frequent and infrequent target stimuli for the attend-sex and attend-color conditions, recorded at channels P8 and C4, and the difference waves associated with the two tasks. A 2 x 2 ANOVA on oddball N2 amplitude with task (attend-color vs. attend-sex) and electrode channel (P8 vs. C4) as repeated measures revealed a main effect of task such that the

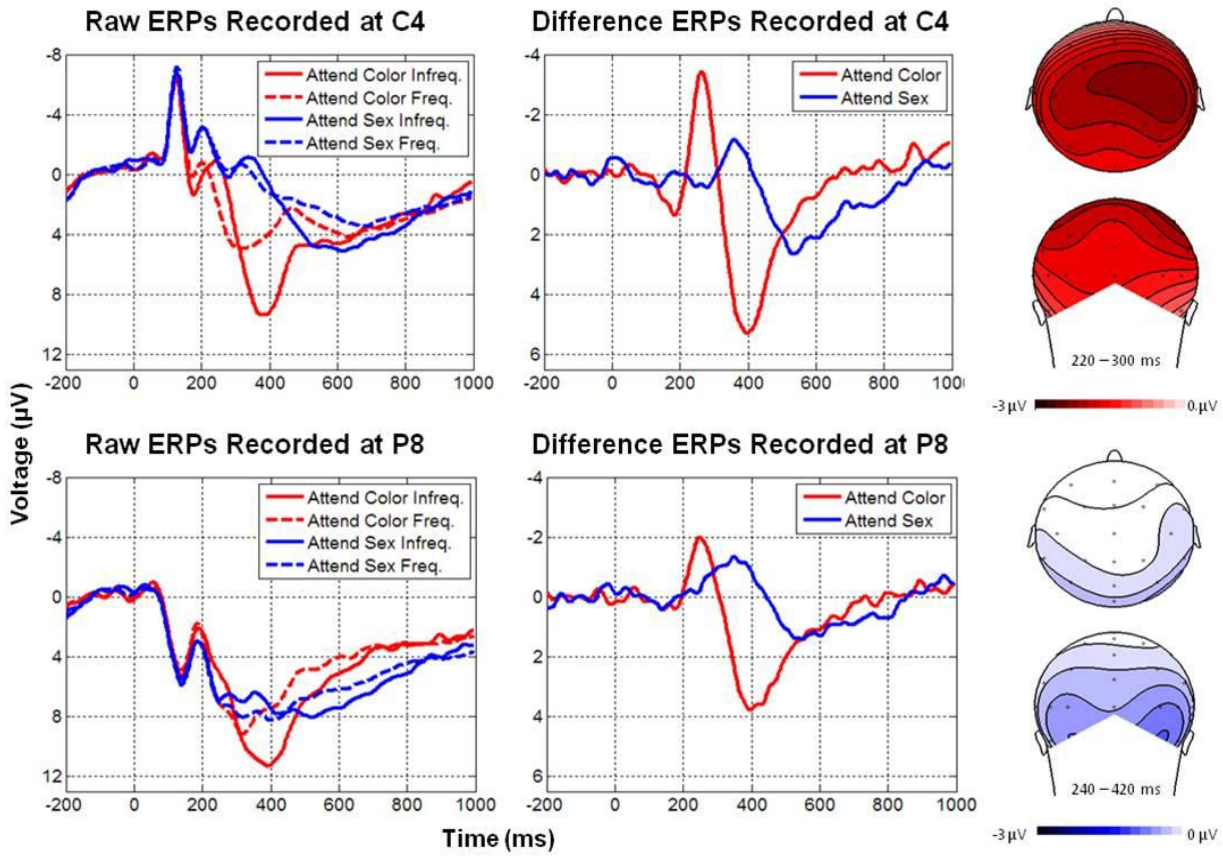

Figure 6. Grand average ERPs recorded from channel C4 (top) and P8 (bottom), and scalp distributions associated with the difference waves. The top distribution (in red) is of the mean amplitude between 220 ms and 300 ms, in the attend-color condition. The bottom distribution (in blue) is the mean amplitude from 240 ms to 420 ms, in the attend-sex condition. Note that negative is plotted up.

oddball N2 was larger to attend-color than to attend-sex, $F(1, 23) = 14.10$, $p < .01$, and no main effect of electrode, $F < 1$. Importantly, there was an interaction between electrode and task such that the oddball N2 was larger at channel P8 in the attend-sex condition whereas the oddball N2 was larger at channel C4 in the attend-color condition, $F(1, 92) = 55.00$, $p < .01$. Paired t-tests confirmed these results, attend-sex: $t(23) = -3.35$, $p < .01$; attend-color: $t(23) = -3.45$, $p < .01$.

We performed a secondary analysis on raw N2 peak amplitude to check that my results were not specific to mean amplitude measures. The difference between frequent and infrequent raw N2 peak amplitude was greatest at frontal-central site FCz (-2.28 µV) in the attend-color condition and

at P8 (-1.74 μV) in the attend-sex condition. Paired t-tests revealed that the peak difference was significantly larger at FCz than P8 in the attend-color condition (-2.28 μV vs. -.39 μV , $t(23) = -2.70$, $p = .01$), but larger at P8 than FCz in the attend-sex condition (-1.74 μV vs. .06 μV , $t(23) = 3.06$, $p < .01$).

Discussion

Subjects were asked to discriminate blue- or yellow-tinted, male or female faces based on the color or on the sex of the face, with the target occurring on either 80% or 20% of the trials. When the participants attended to color, the oddball-N2 was maximal over frontal-central areas, but when they attended to sex, the oddball-N2 was maximal over lateral-occipital sites. my results demonstrate that the effect of stimulus frequency on the scalp distribution of the N2 depends on the attended dimension of the stimulus. Such a dramatic shift in the topology of the N2 within the same task and modality has, to my knowledge, never been shown before.

I also observed a main effect of task such that the size of the oddball N2 was larger on attend-color than attend-sex trials across channel locations. Possibly the longer latency associated with categorizing by sex resulted in greater latency jitter that disrupted the N2 in that condition. Additionally, different levels of accuracy and confidence across conditions may have affected N2/P3 amplitude. Subjects were less accurate in the attend-sex task and, with intermittent feedback indicating their accuracy, were probably less confident as well. Confidence is related to N2/P3 amplitude, with less confident judgements giving rise to smaller N2/P3s (Donchin, 1968; Hillyard, Squires, Bauer, & Lindsay, 1971).

Inspection of the raw ERPs suggest that this effect may have resulted from N2-related latency differences rather than amplitude differences across conditions. That is, if subjects categorized frequent categories faster than infrequent categories, perhaps the N2 occurred earlier for frequent categories and the difference between the N2 to frequent and infrequent categories actually

reflects a latency difference rather than an amplitude difference. However, I observed the same effect of task on topology (frontal-central vs. lateral-occipital across task conditions) when I analyzed the frequency-related difference in raw N2 peak amplitudes, a measure independent of the difference in latency.

One may question why my attend-color oddball N2 was maximal over frontal central regions, rather than regions specific to color processing. In both task conditions, I expected engagement of the anterior cingulate cortex because the anterior cingulate cortex is sensitive to conflict/infrequent events. However, I used the easy, color-discrimination task as a baseline with which to compare different levels of fusiform gyrus engagement, anticipating that the anterior cinglulate cortex would be the most active brain area in the attend-color task, and that the difficulty of the attend-sex task would increase fusiform gyrus activity above anterior cingulate cortex activity.

Some ERP components, such as the C1 and the Readiness Potential, exhibit multiple scalp distributions (Luck, 2005). Of relevance to the present case, release of norepinephrine (NE) by the locus coeruleus, which has broadly distributed afferent projections capable of impacting neural processing throughout the neocortex (Berridge & Waterhouse, 2003), increases the responsivity of individual neurons and improves the signal-to-noise ratio of associated neural networks (Servan-Schreiber, Prinz, & Cohen, 1990) As a consequence, the impact of NE is greatest in whatever area of the brain is most activated by the task at hand (Nieuwenhuis, Aston-Jones, & Cohen, 2005; Nieuwenhuis, de Geus, & Aston-Jones, 2011). In view of this, I suggest that the N2 reflects the impact of NE on active brain areas. This position develops a recent theory that the P3 ERP component reflects the impact of a phasic NE release on cortical processing (Nieuwenhuis et al., 2005). Notably, phasic activity of the locus coeruleus-NE system is characterized by two stages, the phasic burst in firing and a subsequent refractory-like period of quiescence due to auto-inhibition. Thus, large releases of NE are followed immediately by a depletion of NE dependent on the amount

of the initial release. In contrast to the original theory, which holds that the P3 is elicited directly by the phasic increase in NE, I suggest that the N2 reflects the initial NE release and the P3 reflects the subsequent depletion of NE. This hypothesis accounts naturally for the close relationship between the N2 and the P3, which together have historically been termed the "N2/P3 complex" (Duncan-Johnson & Donchin, 1977) and is consistent with the evidence supporting involvement of the locus coeruleus-NE system in P3 production.

Conclusion

The different oddball N2 scalp distributions observed here indicate that they are associated with different neural computations. On the one hand, the N250 has previously been linked to face processing and to familiarity for the target face (Tanaka et al., 2006), and on other hand, the frontal-central N2 has been associated with conflict monitoring by the anterior cingulate cortex. Nevertheless, the oddball N2 – as measured as a difference wave – indicates a common function underlying these very different brain processes: both brain areas respond comparably to changes in stimulus frequency. By this definition, the oddball N2 is the same ERP component despite the different scalp distributions. I propose that this variability reflects the task-dependent impact of phasic NE release on face processing activity in the fusiform gyrus, conflict-related activity in the anterior cingulate cortex, and processing by other cortical areas (e.g., Folstein & Van Petten, 2007; Pritchard et al., 1991).

Chapter Two: Experiments Two through Five[3]

Abstract

I applied the event-related brain potential (ERP) technique to investigate the involvement of two neuromodulatory systems in learning and decision making: The locus coeruleus-norepinephrine (LC-NE) system and the mesencephalic dopamine system (DA system). I have previously presented evidence that the N2, a negative deflection in the ERP elicited by task-relevant events that begins approximately 200 ms after onset of the eliciting stimulus and that is sensitive to low-probability events, is a manifestation of *cortex-wide* noradrenergic modulation recruited to facilitate the processing of unexpected stimuli (Warren, Tanaka, & Holroyd, 2011). Further, Holroyd and Coles (2002) proposed that the impact of DA reinforcement learning signals on the anterior cingulate cortex (ACC) produces a component of the ERP called the feedback-related negativity (FRN). The N2 and the FRN share a similar time range, a similar topography, and similar antecedent conditions. I varied factors related to the degree of cognitive deliberation across a series of experiments to dissociate these two ERP components. Across four experiments I varied the demand for a deliberative strategy, from passively watching feedback, to more complex/challenging decision tasks. Consistent with my predictions, the FRN was largest in the Active Learning experiment and smallest in the Passive Learning experiment whereas the N2 exhibited the opposite effect. Within each experiment, when subjects attended to color, the N2 was maximal at frontal-central sites, and when they attended to sex it was maximal over lateral-occipital areas, whereas the topology of the FRN was frontal-central in both task conditions. I conclude that both the DA system and the LC-NE system act in concert when learning from rewards that vary in expectedness, but that the DA system is relatively more exercised when subjects are relatively more challenged/engaged by the learning task.

[3] These experiments have been invited for submission to a special issue of *Frontiers in Decision Neuroscience.*

Introduction

Adaptive decision making depends on both fast and efficient processing of stimulus events for effective responding (e.g. Servan-Schreiber, Prinz, & Cohen, 1990) and slow trial-to-trial learning of action values for optimizing the selection process (e.g. Schultz, Montague, & Dayan, 1997). The catecholinergic neuromodulatory systems that distribute norepinephrine (NE) and dopamine (DA) have been implicated in these two groups of processes, respectively (Schultz et al., 1997; Servan-Schreiber et al., 1990). Further, putative manifestations of these systems have been identified in the human electroencephalogram (Holroyd, & Coles, 2002; Nieuwenhuis, Aston-Jones, & Cohen, 2005a; Nieuwenhuis Gilzenrat, Holmes, & Cohen, 2005b; Warren, Tanaka, & Holroyd, 2011). However, the way these two systems interact has yet to be explored.

The locus coeruleus-norepinephrine system (LC-NE system) is believed to play a key role in facilitating fast and effective processing of task-relevant stimuli (Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999). The locus coeruleus (LC) is a neuromodulatory nucleus in the midbrain that briefly enhances cortical processing in reaction to motivationally salient or conflict-inducing events (Gilzenrat, Holmes, Rajkowski, Aston-Jones, & Cohen, 2002; Usher et al., 1999). The LC is the primary source of NE to the cortex and other regions (Berridge, & Waterhouse, 2003), where NE release increases the responsivity of individual neurons and improves the signal-to-noise ratio of associated neural networks (Servan-Schreiber et al., 1990). Single cell recordings from the LC in monkeys show that the LC releases NE in phasic bursts to motivationally salient events, and periods of greater phasic release of NE are associated with better performance in target discrimination tasks (Usher et al., 1999). The LC-NE system is also auto-inhibitory, such that phasic

bursts of NE are followed by a refractory-like[4] period lasting approximately 500 ms characterized by reduced or arrested NE supply to the cortex.

In a previous paper (Warren et al., 2011), I proposed that the impact of phasic bursts of NE on cortical processing manifests in the human electroencephalogram as an increase in amplitude of the N2, a negative deflection of the human event-related brain potential (ERP) occurring between about 200 ms and 300 ms after the onset of the eliciting stimulus, the amplitude of which is exercised by unexpected or conflict-inducing events (e.g. Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003). This theory is a modification of a previous "LC-P3 theory" that holds that the phasic bursts of NE produce the P3 – a prominent, positive deflection in the ERP that immediately follows the N2 – rather than the N2 itself (Nieuwenhuis et al., 2005a). Thus, my "modified LC-P3 theory" develops this account by proposing that the LC burst impacts cortical activity somewhat earlier than originally proposed, during the time period of the N2 (~ 250 ms post-stimulus), whereas the LC refractory period coincides with P3 generation.

The timing of the N2 makes it an excellent candidate to reflect the impact of phasic NE release. Aston-Jones and Cohen (2005) estimate NE should begin to arrive in the cortex about 170 ms after onset of the eliciting event, and Usher and colleagues (1999) assess the duration of phasic LC firing in monkeys to be about 180 ms. Thus the impact of NE should appear between 170 ms and 350 ms after the eliciting stimulus, roughly equivalent to the typical onset and offset latencies of the N2.

A key prediction of my proposal is that any change in the ERP due to noradrenergic modulation should exhibit a variable scalp distribution dependent on relative engagement of the different cortical areas giving rise to the ERP. This position follows from two key characteristics of the LC-NE system. First, the broadly dispersed efferent projection system of the LC distributes NE

[4] As opposed to the potassium-mediated refractory period common to individual neurons throughout the brain.

to all regions of the cortex, so any given phasic release can modulate neural activity (and the associated N2) anywhere in cortex (Berridge & Waterhouse, 2003; Nieuwenhuis, Nieuwpoort, Veltman, & Drent, 2007). Second, NE-mediated *changes* in activity should be greatest in cortical areas that are most engaged by the task at hand because increasing the signal-to-noise ratio in the entire cortex will have the greatest impact in those areas (Nieuwenhuis et al. 2005; Nieuwenhuis, de Geus, & Cohen, 2011). This position contrasts with theories of the N2 which posit that the N2 is produced specifically by the anterior cingulate cortex (ACC) and should therefore exhibit a relatively fixed topology, maximal at frontal-central regions of the scalp (e.g. Van Veen, & Carter, 2002a, 2002b; Yeung, Botvinick, & Cohen, 2004).

In previous work, I supported the modified LC-P3 theory by demonstrating that the scalp distribution of the N2 varies widely according to task changes that relatively engage different cortical areas (Warren et al., 2011). I presented subjects with pictures of male and female faces that were tinted either blue or yellow. Subjects attended to either the sex or the color of the faces and counted targets in an oddball task. The impact of frequency was isolated by subtracting frequent stimulus trials from infrequent stimulus trials, yielding a difference wave representative of the change in neural activity specifically caused by differences in stimulus probability (and putatively due to differences in NE recruitment). When subjects attended to the color of the face, the N2 in the difference wave was maximal over frontal-central regions as is often observed in simple oddball tasks (e.g. Holroyd Pakzad-Vaezi, & Krigolson, 2008; Nieuwenhuis et al., 2003; but see Folstein and Van Petten, 2007), consistent with arguments that the N2 is generated in the ACC (Van Veen, & Carter, 2002a, 2002b; Yeung et al., 2004). By contrast, when subjects attended to the sex of the faces the N2 in the difference wave was maximal over lateral occipital regions, consistent with a relatively large change in activity within the fusiform face-processing area (FFA). This study demonstrated that identical task stimuli (colored faces) presented with identical task designs (standards and

deviants) can nevertheless radically alter the topology of the N2 depending on which aspect of the stimuli participants are instructed to attend.

An interesting special case of the N2 occurs when the eliciting stimulus is a feedback stimulus in a reward/no-reward paradigm. A negative feedback stimulus (e.g., that indicates a potential reward was not received) elicits a frontal-central negative deflection in the same time range as the N2 called the feedback error-related negativity (FRN), but positive feedback does not (Miltner, Braun, & Coles, 1997); the FRN is usually measured with a difference wave approach whereby the ERP to reward feedback is subtracted from the ERP to error feedback. The negative deflection observed in this difference wave is characterized by variance in the ERP associated with both negative and positive feedback, and therefore I distinguish it from the FRN by calling it the difference-wave feedback-related negativity (dFRN). Source localization studies suggest that the dFRN is generated in, or very close to, the ACC (Gehring, & Willoughby, 2002; Hewig, Trippe, Hecht, Coles, Holroyd, & Miltner 2007; Miltner, Lemke, Weiss, Holroyd, Scheffers, & Coles, 2003). Additionally, a neurocomputational theory of this ERP component is based on the seminal observation that rewarding events elicit phasic bursts of dopamine (DA) activity that are utilized by the targets of the DA system for the purpose of adaptive decision making (Schultz et al., 1997). In particular, single-cell recordings from primates show increased phasic DA activity in response to unexpected rewards or reward predictors, and shallow dips from baseline DA activity in response to punishment or to the absence of expected rewards (e.g. Schultz, 2002). Holroyd and Coles (2002) proposed the reinforcement learning theory of the dFRN, which holds that the dFRN reflects the impact of these phasic DA signals on the ACC such that motor neurons in the ACC are inhibited and disinhibited by phasic increases and decreases of DA, respectively.

Recent evidence suggests that these phasic DA signals specifically modify the amplitude of the N2. According to this position, the ACC produces a negative deflection to unexpected task

relevant events (the N2), including unexpected negative feedback and unexpected reward feedback. However, unexpected reward feedback also elicits a dopamine-induced positive deflection ("the reward positivity") that is superimposed over the N2 and cancels it out (Holroyd et al., 2008). In other words, unexpected error and reward feedback elicit the N2, but unexpected reward feedback *also* elicits a reward positivity that obscures the N2, creating the difference observed between the ERPs to positive and negative feedback (the dFRN).

To dissociate the reward positivity from the N2, a recent multi-experiment study presented subjects with complicated reward-feedback that indicated not only whether a subject had won or lost money, but also what response was required of them for the subsequent trial (Baker & Holroyd, 2011). In one experiment, a stimulus-induced delay in reward processing caused the reward positivity to appear about 100 ms later than usual (peaking at about 350 ms), thereby exposing the N2 on those trials. When the reward-feedback stimulus was simplified in further experiments, the reward positivity appeared earlier and attenuated the N2. Furthermore, factors related to response conflict impacted N2 amplitude and reduced the reward positivity on high-conflict reward trials. Additional research has demonstrated that the reward positivity to monetary rewards is alternatively larger in problem gamblers (Hewig et al., 2010) and smaller in substance abusers, who evidence a large N2 to reward feedback (Baker, Stockwell, Barnes, & Holroyd, 2011).

The ACC has been posited to be the neural generator of both the N2 (Van Veen & Carter, 2002a, 2002b; Yeung et al., 2004) and the dFRN (Holroyd & Coles, 2002). Furthermore, here I have proposed that noradrenergic modulation enhances activity in the ACC and all across the cortex, amplifying the N2 in target areas. Thus, there are three factors that push the amplitude of the N2 at frontal-central scalp locations up and down; ACC activity, noradrenergic modulation, and dopaminergic modulation. If we have any chance of understanding how the frontal-central N2

provides insight into ACC function, I need to understand how these systems interact - otherwise I will be at a loss to interpret N2 data.

To investigate this issue, I employed the same paradigm used in my previous study (Warren et al., 2011), presenting subjects with male or female faces tinted either blue or yellow, with frequent or infrequent category presentations based on either the sex or the color of the faces. But here the stimuli also indicated reward or no-reward, allowing me to simultaneously examine the N2 and the dFRN. I manipulated the amplitudes of the reward positivity and the N2 along two independent dimensions. Along one dimension, I varied (across subjects) the degree of participant engagement in a feedback task, which is known to affect dFRN amplitude. For example, Yeung, Holroyd, and Cohen (2005) manipulated degree to which a deliberative strategy was required of subjects, from passively observing reward/no-reward outcomes, to actively making a decision that would result in either reward or no-reward. The dFRN was significantly larger when subjects utilized the feedback to optimize their decisions, as opposed to passively collecting rewards (see also Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Li, Han, Lei, Holroyd, & Li, 2011; Peterson, Lotsz, Halfren, Sejnowski, & Poizner, 2011; Zhou, Yu, & Zhou, 2010). I implemented this manipulation across three experiments wherein subjects passively collected rewards in Experiment 1 (Passive Learning), made a decision based on multiple stimulus feature-response combinations in Experiment 2 (Active Learning), and intermediate to these, made a decision based on relatively simple response-reward contingencies in Experiment 3 (Moderate Learning). I predicted that the dFRN would be largest in the Active Experiment and reduced or absent in the others. By contrast, I predicted that the N2 would be smaller with increasing task engagement because of component overlap with the reward positivity elicited by infrequent rewards.

Along the second dimension I varied N2 amplitude by manipulating (within subjects) the attended dimension of the feedback: As in my previous study, subjects were required to attend to

either the color or the sex of the feedback stimuli (male or female faces tinted either blue or yellow). I predicted that switching from color to sex would move the N2 from frontal-central to lateral-occipital regions of the scalp. By contrast, I predicted that the dFRN would remain frontal-central irrespective of the attended dimension of the feedback. Further, I predicted that I would observe maximal interference between the two components in the color condition of the Active Experiment, where both the N2 and the reward positivity are frontal central. These results would validate my claim that the N2 and dFRN are produced by distinct neural mechanisms, one that produces a negativity to infrequent events that has a variable scalp distribution consistent with a noradrenergic origin, and one that produces a positivity to rewards and a negativity to no-rewards that has a frontal-central scalp distribution consistent with genesis in the ACC.

Experiment 1: Passive Learning

In the Passive Experiment I sought to replicate the results of my previous study by engaging the LC-NE system and the N2 in an oddball task with minimal involvement of reinforcement learning systems and therefore minimal interference from the dFRN. I employed the exact same paradigm as reported in my previous work (Warren et al., 2011) except that instead of counting stimuli associated with a target category (e.g., male faces), subjects counted earnings accrued with each stimulus presentation (e.g., if subjects were told that they would be given 5 cents for each male face); they were asked to report the sum once during the block and a second time at the end of the block. Importantly, because participants were not required to make an overt response on each trial, I expected this task to elicit only a small dFRN, if any (Holroyd et al., 2009; Li et al., 2011; Yeung et al., 2005). Further, as I observed previously, I predicted that relative engagement of the fusiform face-processing area in the attend-sex condition would enhance the N2 over lateral-occipital sites, whereas relative engagement of the ACC in the attend-color condition would enhance the N2 over

frontal-central sites. Finally, I predicted that the dFRN – to the extent that it was present -- would not exhibit any changes in scalp topography.

Method

Methods were identical across all four experiments except where indicated.

Participants. 21 people (3 males) completed this experiment. For all experiments reported in this paper, participants signed up through the research participation system at the University of Victoria, Canada, and were compensated with extra credit in an undergraduate psychology course or were paid $20.00 Canadian for their time. This project (experiments 1 through 4) was approved by the human subjects review board at the University of Victoria and conducted in accordance with the ethical standards prescribed in the 1964 Declaration of Helsinki.

Apparatus and Procedure. Participants were seated comfortably, approximately 50 cm in front of a computer screen, in an electromagnetically shielded booth. Stimuli consisted of male or female faces (30 examples of each, lifted from black and white photos, excluding hair and contour of head) tinted either blue or yellow (~4.4º visual angle). In a previous experiment (Warren, et al., 2011), I used a larger set of the same stimuli (40 males and 40 females), but because the error rates in discriminating between male and female faces were high, here I selected a subset of those stimuli: The 75% that were most accurately discriminated previously. For both stimulus dimensions (color, sex), one stimulus type occurred infrequently (20% of all trials). The order of stimulus presentation was randomized with replacement. At the beginning of each block, subjects were instructed by the computer program to keep track of presentations of a specific target stimulus (blue faces, yellow faces, male faces or female faces). The task consisted of 8 blocks of 75 trials each (600 total trials), counterbalanced such that each of the four stimulus types (blue males, yellow males, blue females, yellow females) occurred in two blocks as the target, and of those two blocks, once as a frequent target and once as an infrequent target. Stimuli were presented for 1200 ms and were separated by a

fixation cross displayed for 300 ms (see Fig. 7, Passive Learning, for a graphic representation of the task).

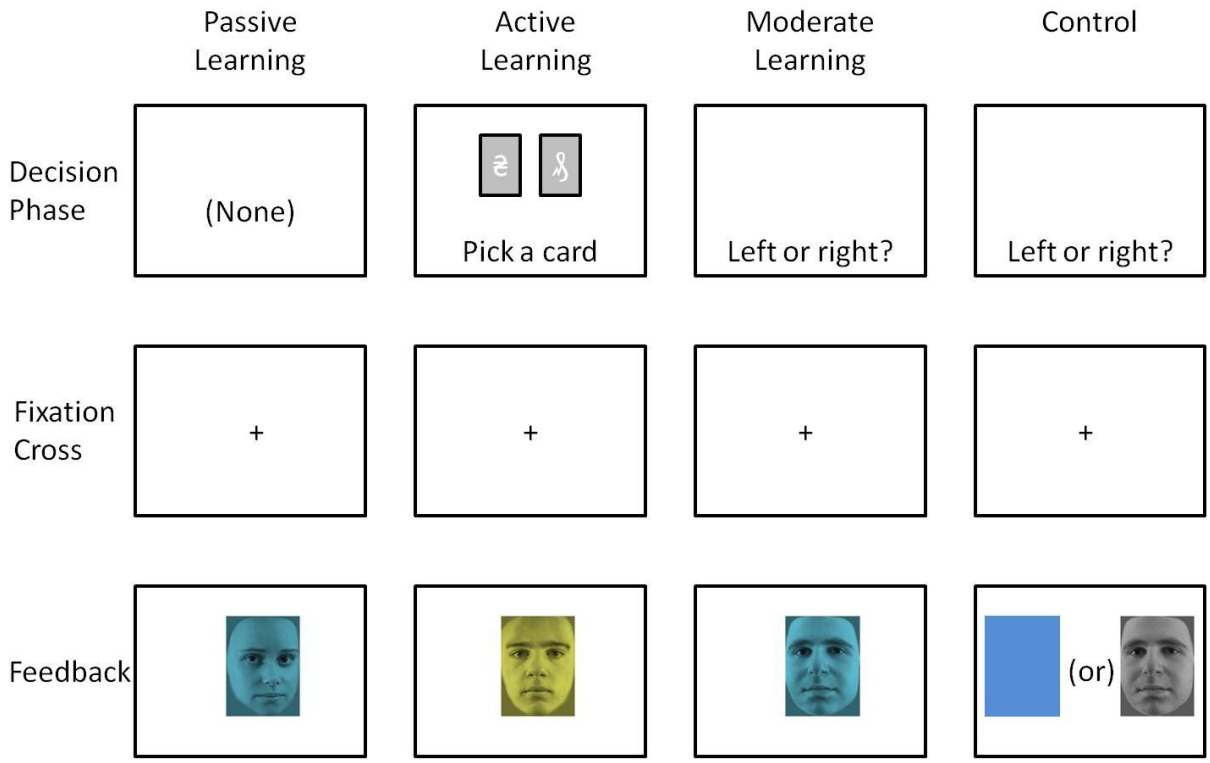

Figure 7. Graphic representation of the procedure for each of the four experiments in Chapter 2.

Each presentation of the target stimulus category indicated that the subject won $0.05. Subjects were instructed to keep track of the money won and were required to report their count twice per block (at a random trial number about halfway through each block, and at the end of each block). This method yielded 16 reports of the subject's money count. Subjects reported their count by answering an eight-choice multiple choice question, choosing from several ranges within which the correct count fell (e.g. between $0.30 and $0.50, or between$ 0.55 and $0.75, etc...). I assessed accuracy by dividing the number of correct reports by the number of total reports.

Data Acquisition. The electroencephalogram (EEG) was recorded from 41 electrode locations arranged in the standard 10-20 layout using Brain Vision Recorder software (Version 1.3, Brainproducts, Munich, Germany). During recording, the EEG data were referenced to the average voltage across channels, sampled at 250 Hz, and amplified (Quick Amp, Brainproducts) and filtered through a passband of 0.017-67.5 Hz (90 dB octave roll off). Impedances were below 12 kΩ.

EEG Data Analysis. The EEG data were filtered off-line through a 0.1-20 Hz passband phase-shift-free Butterworth filter and re-referenced to linked mastoids. Ocular artifacts were removed using the algorithm described by Gratton, Coles, and Donchin (1983). Trials in which the change in voltage at any channel exceed 35 μV per sampling point were removed. In total, .02% of the data were discarded. 1000 ms epochs of data were extracted from the continuous EEG from 200 ms before stimulus onset to 800 ms after. The data were baseline-corrected according to the average amplitude of the EEG over the 200 ms preceding stimulus presentation and ERPs were created by averaging the EEG data for each condition, electrode site, and participant.

To isolate the effect of reward independent of frequency, I subtracted the ERPs associated with reward from the ERPs associated with no-reward yielding an attend-color dFRN and attend-sex dFRN that were equated for the effect of stimulus probability. This method maximized the signal-to-noise ratio in the ERPs, as opposed to averaging the ERPs separately for the infrequent reward trials, frequent reward trials, infrequent no-reward trials, and frequent no-reward trials. Similarly, to isolate the effect of frequency independent of reward feedback, I subtracted the ERP associated with the frequently occurring stimuli from the ERP associated with the infrequently occurring stimuli, collapsed across reward condition, yielding a difference-wave N2 for each task condition (attend-color, attend-sex). Thus, each of the infrequent and frequent ERPs contained equal numbers of reward and no-reward trials such that the difference between these ERPs were equated for the effects of reward. The interaction of the "raw" N2 (as opposed to difference-wave

N2) and the reward positivity to the four individual conditions was examined separately in an across-group comparison (below).

Note that because LC-NE system activity causes a change in the relative activation of the underlying cortical systems (i.e. making ERP components larger), the impact of NE on the ERP is most appropriately measured in a difference wave that isolates that change. In the oddball paradigm, the infrequent stimulus category putatively recruits a larger NE release than the frequent stimulus category (Nieuwenhuis et al., 2005a). Thus to assess the impact of NE, I subtract the frequent-stimulus ERP from the infrequent-stimulus ERP. When this is done, the difference wave typically exhibits a prominent negativity between about 200 ms and 300 ms post stimulus (sometimes later if the stimulus is more difficult to categorize) that hereafter I refer to as the difference-wave N2 (dN2).

The amplitudes of the dN2 and dFRN (in the difference waves) were assessed using a base-to-peak measure as follows: For each subject in each condition, the most negative peak between 200 ms and 280 ms in the attend color condition, or 300 ms vs. 380 ms in the attend-sex condition was identified and recorded as the dN2/dFRN peak amplitude. The base amplitude of the dN2/dFRN was then taken as the most positive voltage prior to the dN2/dFRN, and these values were subtracted from the dN2/dFRN peak amplitude, yielding my base-to-peak measures. This procedure controls for overlap with the P2, a positive deflection that typically immediately precedes the dN2, that can push the dN2 into positive peak values. Note that because the dFRN is not typically preceded by any notable deflection in the ERP, the base measure is approximately 0 μVs; for this reason the base-to-peak measure of the dFRN is equivalent to a peak amplitude measure. However, I chose to assess dFRN base-to-peak for consistency with my method for assessing dN2 amplitude.

In assessing the change in component topology across task conditions, I focused on two electrode sites representative of frontal-central and lateral-occipital scalp regions as I did in my

previous study, specifically at channel locations FCz and P8. Both the dFRN and the dN2 are typically maximal at channel FCz (e.g. Holroyd et al., 2008) and the dN2 was maximal at channel P8 in the attend-sex condition of my previous study (Warren et al. in press). Single-tailed t-tests were applied to assess the amplitudes of these ERP components at these channels because of my a priori hypotheses of the direction of each difference. For example, I predicted that the dN2 would be larger at channel P8 than at channel FCz for the attend-sex condition; a dN2 that was larger at FCz than at P8 would run contrary to my hypothesis.

Results

Behavioral Results. Mean accuracy was 79.2% (*SD* = 14.4%) for the attend-color condition and 68.5% (*SD* = 21.2%) for the attend-sex condition. The data of one subject were eliminated from further analysis because the accuracy score was more than two standard deviations below the mean in the attend-color condition. For the remaining 20 subjects, mean accuracy was 80.6% (*SD* = 13.1%) for the attend-color condition and 70% (*SD* = 20.4%) for the attend-sex condition. This difference approached significance using a two-tailed t-test, $t(19) = -2.0$, $p < .10$.

EEG Results. The raw ERPs, difference waves and scalp distributions are shown in Figure 8. Inspection of the scalp distributions suggests that the attend-color dN2 was maximal over frontal-central sites (FCz, -4.5 μV) whereas the attend-sex dN2 is maximal at lateral-occipital regions (PO8, -3.5 μV). This impression was confirmed with a 2x2 ANOVA on dN2 amplitude with electrode (FCz vs. P8) and task (attend-color vs. attend-sex) as repeated factors. There was an effect of task such that the dN2 was larger in the attend-color condition (-4.0 μV) than the attend-sex condition (-2.8 μV), $F(1,19) = 10.8$, $p < .01$, $\eta^2 = .36$. There was also an interaction of electrode and task, $F(1,19) = 6.8$, $p < .05$, $\eta^2 = .26$, and one tailed paired samples t-tests revealed that in the attend-color condition, the dN2 was larger at FCz than P8 (-4.5 μV vs. -3.4 μV), $t(19) = -2.0$, $p < .05$, whereas in

the attend-sex condition the dN2 was larger at P8 than at FCz (-3.2 μV vs. -2.5 μV), $t(19) = 2.0$, $p < .05$.

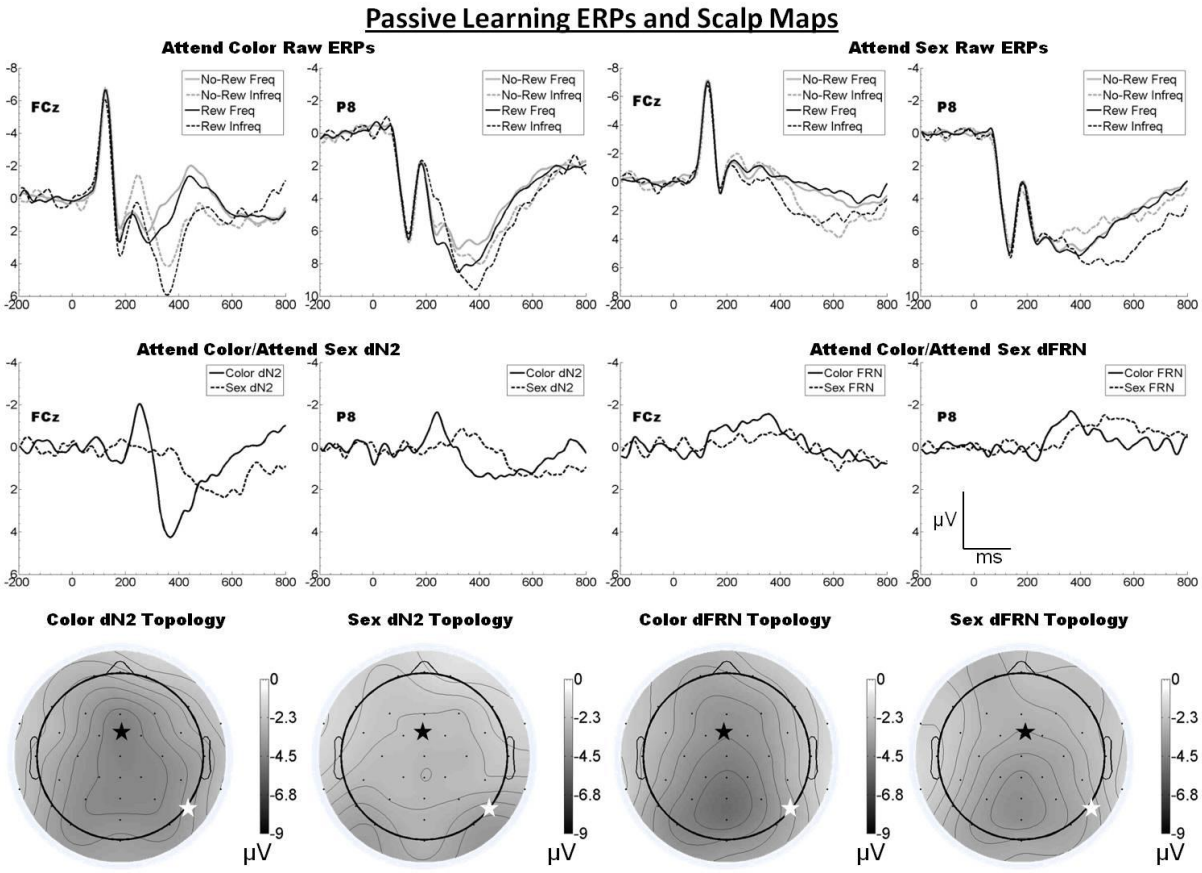

Figure 8. Grand average ERPs for experiment 1, recorded from channel FCz and P8 (see labels) and scalp distributions associated with the difference waves. The top row shows the "raw" ERPs for each of the frequency by reward conditions across tasks and electrodes. The middle row shows the dN2 and dFRN difference waves across task and electrodes. The bottom row shows the scalp distributions of the dN2 and dFRN across tasks. The scalp distributions reflect the base-to-peak measure of each of the dN2 and dFRN. The black star on the scalp map denotes channel FCz, and the white star denotes P8. Note that negative is plotted up.

Inspection of the scalp distributions in Figure 8 further indicates that the dFRN was distributed over posterior, rather than frontal, regions of the head in both the attend-color (Pz, -5.2 μV) and attend-sex (POz, -4.0 μV, followed by Pz, -4.0 μV) conditions. A 2x2 ANOVA on dFRN amplitude with electrode and task as repeated factors revealed an effect of electrode such that the

dFRN was larger at FCz than at P8 (-3.5 μV vs. -2.9 μV), $F(1,19) = 4.8$, $p < .05$, $\eta^2 = .20$. There was a trend toward a main effect of task such that the attend-color task yielded a larger dFRN than the attend-sex task (-3.5 μV vs. -2.9 μV), $F(1,19) = 4.0$, $p < .10$, $\eta^2 = .18$). There was also a trend toward an interaction of electrode and task, $F(1,19) = 3.2$, $p < .10$, $\eta^2 = .14$, and one-tailed paired samples t-tests revealed that in the attend-color condition, the dFRN was larger at FCz than P8 (-4.1 μV vs. -3.0 μV), $t(19) = -2.3$, $p < .05$, whereas there was no significant difference in the attend-sex condition (FCz: -3.0 μV; P8: -2.9 μV, $p > .05$). For an additional check of the dFRN scalp distribution, I ran two t-tests to examine the difference between channel Pz and FCz. These tests indicated that the dFRN was larger at Pz than FCz in the attend-sex condition, $t(19) = 2.9$, $p < .01$, but not in the attend-color condition, $t(19) = 1.4$, $p > .05$.

Discussion

We proposed that the dN2 is a manifestation of cortex-wide NE neuromodulation, and predicted that the impact of NE modulation on cortex and therefore the topology of the dN2 should vary according to task demands. By contrast, a standard theory of the dFRN holds that it is produced by the impact of DA signals on ACC activity, and therefore that the dFRN should appear with a consistent frontal-central scalp topology across task conditions. Here, I replicated my previous finding that the dN2 changes from exhibiting a primarily central scalp distribution when subjects categorize tinted faces based on color to a more lateral-occipital distribution when subjects categorize the same face stimuli based on the sex of the face. Further, although the dFRN was larger at frontal central regions in both the attend-color and attend-sex conditions, it was not significantly larger at FCz than P8 in the attend-sex condition, it was relatively small overall (ranging from -2.9 μV to -4.1 μV), and it exhibited a scalp distribution that was mostly posterior (see Fig. 8). These results are inconsistent with the identification of this component with the dFRN (Miltner et al.,

1997) and indicate that (as predicted) this task did not produce a robust dFRN. I conclude that, with minimal interference from the dFRN, the dN2 exhibits a prominent yet variable scalp distribution.

Experiment 2: Active Learning

The Active Experiment maximized engagement of the system underlying the dFRN by presenting subjects with an apparently complex decision task that encouraged deliberation. Subjects were asked to choose between two elaborate images of tarot cards presented side-by-side on a computer screen by pressing either a left or right key on a keyboard. Six different cards were paired a total of 15 different ways. The subjects were told that with each pairing one card had a better chance of winning than the other, and that they were required to learn which card to pick in any specific pairing (as opposed to finding which of the six cards had the best chance of winning overall). The complexity of the stimulus displays was intended to cultivate a sense that the task was challenging yet learnable (when in fact it was not). In so doing I expected the feedback stimuli to elicit a relatively large dFRN with a frontal-central scalp topography for both the attend-sex and attend-color conditions. I further predicted that the dFRN would interfere with the production of the dN2 in both the attend-color and attend-sex conditions.

Method

Participants. 20 people (6 males) participated in this study.

Apparatus and Procedure. Stimuli and procedure were the same as in the Passive Experiment except where indicated. Each trial began with presentation of two tarot cards appearing on a computer display side by side (see Fig. 7, Active Learning). Instead of passively counting their accumulated winnings as in the Passive Experiment, subjects were required to choose between the two tarot cards by pressing the appropriate key on the keyboard. The choice screen was displayed until the participant made their decision. When a selection was made the cards were replaced by a fixation cross for 600 ms and then the face feedback stimulus was presented for 1200 ms.

Tarot cards were detailed images (six images in total) presented in random pairs. Subjects were instructed to try and learn which cards had a better chance of "paying-off" in any given pairing, and to maximize their winnings by consistently making the best choice. In addition, subjects were told there would be "hard" blocks in which the pay-off chances for making the right choice were only 10% and 30%, and "easy" blocks in which the pay-off chances were 70% and 90%. Because of the length of the task, the number of trials was reduced from 600 in the Passive Experiment to 400 for the Active Experiment. The task consisted of 8 blocks of 50 trials each and the conditions were counterbalanced across blocks as in the Passive Experiment. Additionally, I included only 8 (rather than 16) money-count reports (one per block). As in the Passive Experiment, I assessed accuracy in reporting the money count for the Active Experiment by dividing the number of correct reports by the number total reports.

EEG Data Acquisition and Analysis

The EEG data were acquired and analyzed and the dN2 and dFRN were assessed in the same way as in the Passive Experiment.

Results

Behavioral Results.. The mean accuracy was 82.5% (*SD* = 21.6%) for the attend-color condition and 86.2% (*SD* = 15.1%) for the attend-sex condition. This effect of task condition on accuracy was not significant (p>.05).

EEG Results. The raw ERPs, difference waves and scalp maps are shown in Figure 9. Inspection of the scalp distributions suggests that the attend-color dN2 was mostly flat across the scalp, but with a maximum over lateral-occipital sites (PO7, -3.8 μV). This was also true for the attend-sex dN2 (PO7, -3.7 μV). A 2x2 ANOVA on dN2 amplitude with electrode (FCz vs. P8) and task (attend-color vs. attend-sex) as repeated factors yielded no significant effects (all ps >.05). The

mean dN2 amplitudes were as follows: Attend-color: FCz , -3.1 μV, P8, -3.1 μV; attend-sex: FCz, -3.1 μV, P8, -3.1 μV.

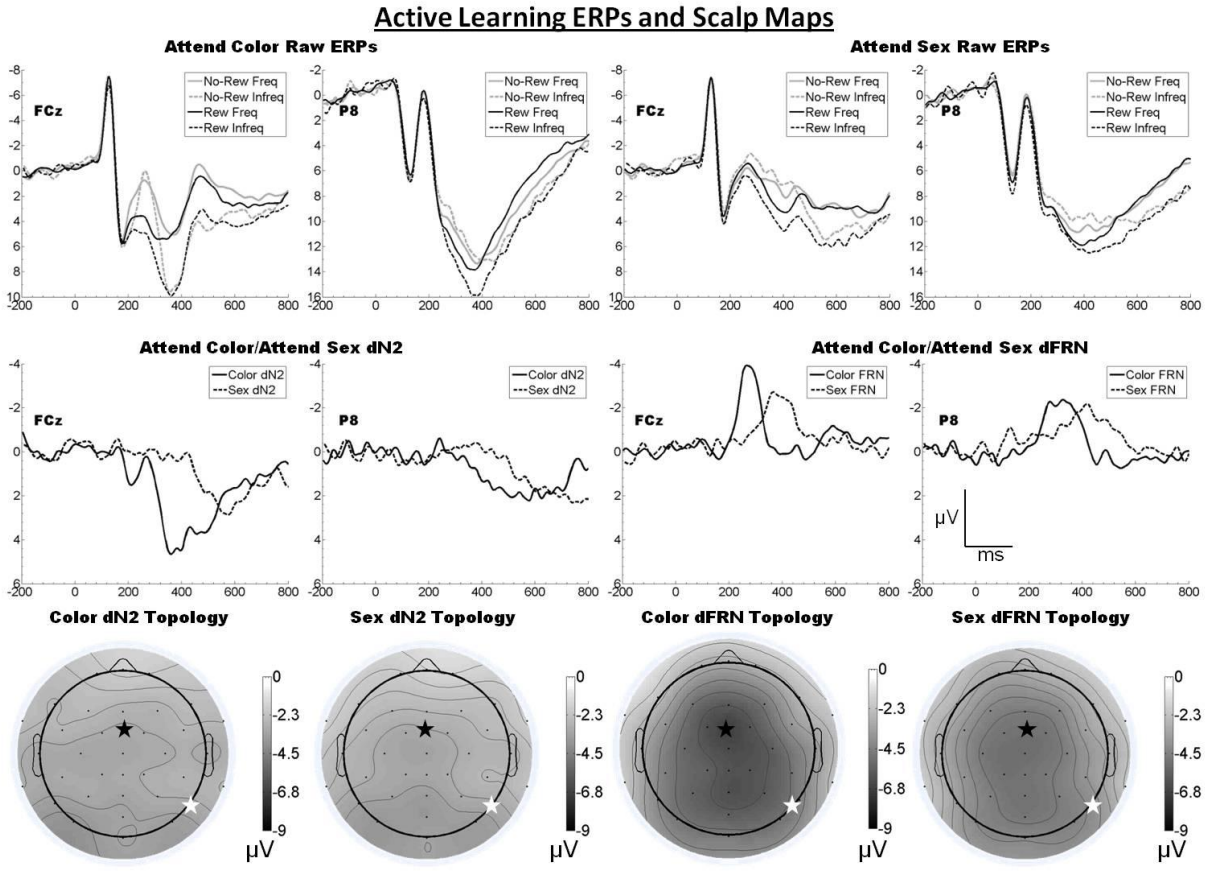

Figure 9. Grand average ERPs for experiment 2, recorded from channel FCz and P8 (see labels) and scalp distributions associated with the difference waves. The top row shows the "raw" ERPs for each of the frequency by reward conditions across tasks and electrodes. The middle row shows the dN2 and dFRN difference waves across task and electrodes. The bottom row shows the scalp distributions of the dN2 and dFRN across tasks. The scalp distributions reflect the base-to-peak measure of each of the dN2 and dFRN. The black star on the scalp map denotes channel FCz, and the white star denotes P8. Note that negative is plotted up.

Inspection of the scalp distributions of the dFRN indicates the dFRN was strongly frontal-central in both conditions (attend-color: FCz, -6.7μV; attend-sex: Cz -5.2 μV followed by CP1, -5.1 μV, and FCz, -5.0 μV). A 2x2 ANOVA on dFRN amplitude with electrode and task as repeated factors confirmed this impression, revealing an effect of electrode such that the dFRN was larger at

FCz than at P8 (-5.9μV vs. -4.3 μV), $F(1,19) = 14.3$, $p < .01$, $\eta^2 = .43$. There was also an effect of task such that the attend-color dFRN was larger than the attend-sex dFRN (-5.6 μV vs. -4.5 μV), $F(1,19) = 5.9$, $p < .05$, $\eta^2 = .24$. The interaction of electrode and condition was not significant ($p > .05$). Further analysis revealed that in the attend-color condition the dFRN was significantly larger at FCz than at P8 (-6.7 μV vs. -4.6 μV), $t(19) = -3.2$, $p < .01$, and in the attend-sex condition the dFRN was also larger at FCz than at P8 (-5.0 μV vs. -4.0 μV), $t(19) = -2.2$, $p < .05$. An additional t-test examining the scalp distribution of the attend-sex dFRN revealed it was not significantly larger at Cz than FCz ($p > .05$).

Discussion

As predicted, increasing the complexity of the stimulus display resulted in a larger dFRN for both the attend-color and attend-sex conditions, evidently because these task elements were better able to engage the system that produces the dFRN. Further, the dFRN appeared frontal central in both conditions, as predicted. By contrast, the dN2 was relatively small and noisy, and its topology was relatively flat in contrast to the results of the Passive Experiment and my previous work (Warren et al., 2011). I suggest that component overlap with the dFRN reduced dN2 amplitude in this task in both the attend-color and attend-sex conditions, a question to which I will return in my across-experiments analysis (see below).

Experiment 3: Moderate Learning

The probability manipulation in the Passive Experiment elicited a strong dN2 but the passive nature of the task failed to strongly engage the systems that produce the dFRN. By contrast, the challenging learning task utilized in the Active Experiment produced a large dFRN that strongly attenuated the dN2. In the Moderate Experiment, I sought to utilize a task that would produce both a dN2 and dFRN to compare the two components within a single experiment. I therefore simplified the decision task in the Active Experiment such that it would engage (putatively) the DA system

sufficiently to produce a dFRN, but not so strongly that the dFRN would obscure the dN2. I predicted that in this task the dN2 would exhibit a variable scalp distribution across the attend-color and attend-sex conditions, whereas the dFRN would not.

Method

Participants. 22 people (5 males) participated in this study.

Apparatus and Procedure. Stimuli and procedure were the same as in the Passive Experiment and the Active Experiment except where indicated. Instead of passively watching faces or choosing between two tarot cards, on each trial participants made a choice between a left or right key press. The decision screen consisted only of the words "left or right?" (see Fig. 7, Moderate Learning). Participants chose between a left or right key press and were subsequently presented with the face-feedback stimulus. Subjects were told that for a random number of consecutive trials, each key had a set probability of "paying off," and the underlying probabilities would change randomly approximately every 20 trials. Subjects were instructed to try and maximize their winnings by finding and choosing the "better" key during any given set of trials, and to switch their choice whenever they suspected the underlying probabilities had changed. Subjects were told there would be easy blocks of trials with high probabilities of pay-off, and hard blocks of trials with low probabilities of pay-off, just as in the Active Experiment. The decision screen was presented until subjects made a choice whereupon a fixation cross was presented for 500 ms, followed by the face feedback for 1000 ms. Trial numbers were increased to 100 trials per block over 8 blocks, counterbalanced across blocks in the same manner as in the Passive Experiment. Subjects were told in the debriefing that in fact the pay-off chances were set at 20% and 80% throughout the experiment, independent of their key presses.

Subjects were required to report their exact reward earnings count for each set of trials (starting at zero from the last accuracy test), twice per block, for a total of 16 reports. Responses

within $0.25 of the correct count were coded as correct; total accuracy was defined as the number of correct reports divided by the number of total reports (16).

EEG Data Acquisition and Analysis

The EEG data were acquired and analyzed, and the dN2 and dFRN were assessed as in the Passive and Active experiments.

Behavioral Results. Mean accuracy was 90.3% (*SD* = 11.5%) in the attend-color condition and 84.7% (*SD* = 18.5%) in the attend-sex condition. I eliminated the data from three subjects for having either attend-color or attend-sex accuracy scores more than two standard deviations below the mean. With these subjects eliminated, attend-color accuracy was 92.8% (*SD* = 7.6%) and attend-sex accuracy was 90.8% (*SD* = 8.2%). This difference in accuracy was not significant ($p > .05$).

EEG Results. The raw ERPs, difference waves and scalp maps are shown in Figure 10. Inspection of the scalp distributions suggest the attend-color dN2 was mostly frontal central (FCz, 4.5 μV), but with a left-lateral-occipital maxima (PO7, -4.6 μV). A two-tailed t-test indicated these channels were not significantly different ($p>.05$). The attend-sex dN2 was maximal at PO8 (-3.2 μV). The impression of a mostly frontal-central attend-color dN2 was supported by the results of 2x2 ANOVA on dN2 amplitude with electrode and task as repeated factors. There was an effect of task such that the attend-color dN2 was larger than the attend-sex dN2 (-4.1 μV vs. -2.5 μV), $F(1,18) = 13.9$, $p < .01$, $\eta^2 = .44$, and there was an interaction of task and electrode such that the attend-color dN2 was larger at FCz than P8 (-4.5 μV vs. -3.7), whereas the attend-sex dN2 was larger at P8 than FCz (-2.9 μV vs. -2.1 μV), $F(1,18) = 10.0$, $p < .01$, $\eta^2 = .36$. One-tailed paired t-tests indicated these differences were significant, attend-color: $t(18) = -1.9$, $p < .05$; attend-sex: $t(18) = 2.7$, $p < .01$.

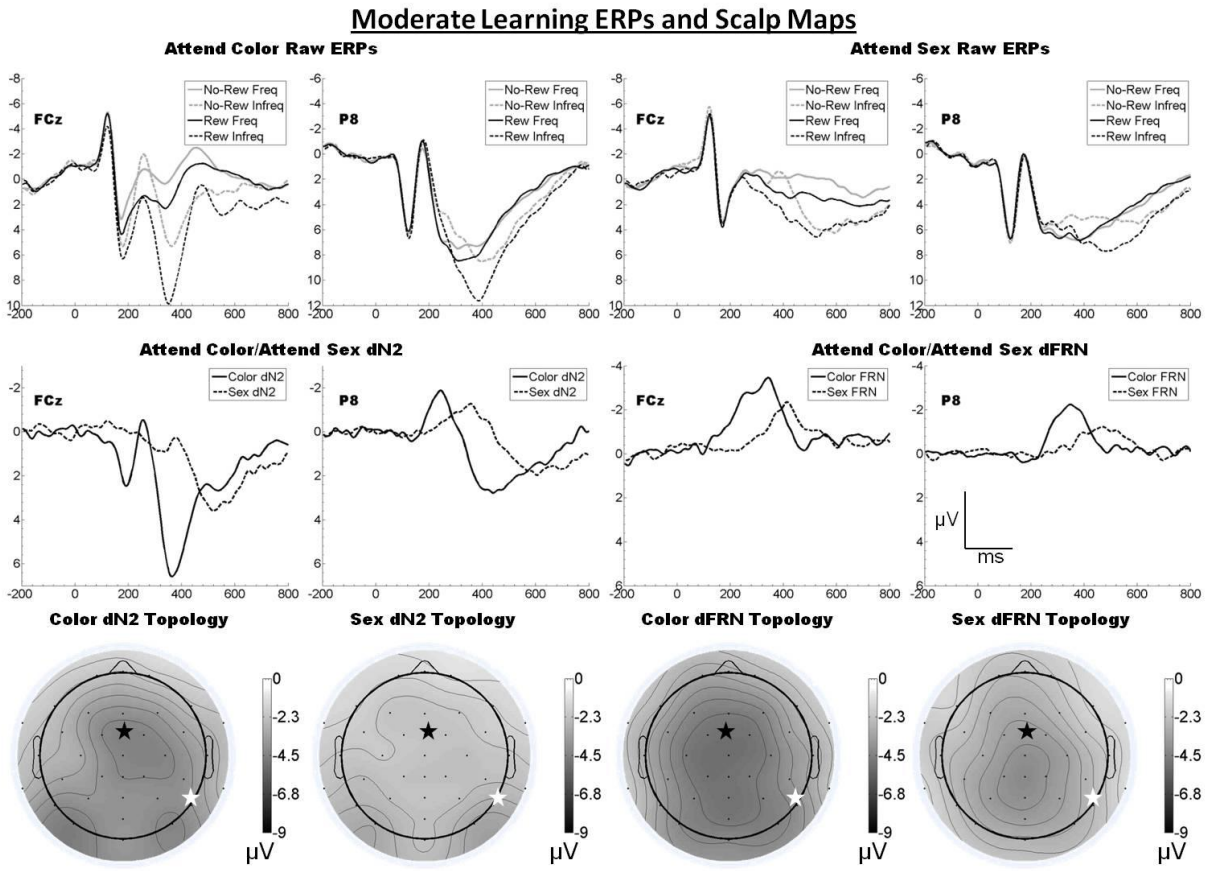

Figure 10. Grand average ERPs for experiment 3, recorded from channel FCz and P8 (see labels) and scalp distributions associated with the difference waves. The top row shows the "raw" ERPs for each of the frequency by reward conditions across tasks and electrodes. The middle row shows the dN2 and dFRN difference waves across task and electrodes. The bottom row shows the scalp distributions of the dN2 and dFRN across tasks. The scalp distributions reflect the base-to-peak measure of each of the dN2 and dFRN. The black star on the scalp map denotes channel FCz, and the white star denotes P8.Note that negative is plotted up.

By contrast, inspection of the scalp maps in Figure 10 suggests that the attend-color dFRN was shallow, and maximal at central channels (CPz, -4.6 μV, followed by Pz, -4.5 μV, Cz, -4.5 μV, CP1, -4.5 μV, and FCz, -4.4 μV). The attend-sex dFRN was maximal at CPz (-4.2), followed by Cz (-4.0 μV). FCz was the seventh most negative channel (-3.6 μV). A two-tailed t-test indicated the amplitude of the dFRN at CPz and FCz did not differ significantly in the attend-color condition (p>.05), but the attend-sex dFRN was larger at CPz than FCz, $t(18) = 2.6$, $p < .05$. A 2x2 ANOVA

on dFRN amplitude revealed an effect of electrode such that the dFRN was larger at FCz than at P8 (-4.0 μV vs. -2.7 μV), $F(1,18) = 10.4$, $p < .01$, $\eta^2 = .37$. The effect of task was also significant such that the attend-color task yielded a larger dFRN than the attend-sex task (-3.7 μV vs. -3.0 μV), $F(1,18) = 5.8$ $p < .05$, $\eta^2 = .24$. The interaction of electrode and task was not significant ($p > .05$). As in experiments 1 and 2, I used one-tailed paired samples t-tests comparing dFRN amplitude at FCz and P8 in the attend-color and attend-sex conditions. In the attend-color condition, the dFRN was significantly larger at FCz than at P8 (-4.4 μV vs. -2.9 μV), $t(18) = -2.9$, $p < .01$. In the attend-sex condition, the dFRN was also larger at FCz (-3.6 μV vs. -2.5 μV), $t(18) = -2.7$, $p < .01$.

Discussion

We found that in a task designed to engage the learning system only moderately, a dFRN was elicited over central scalp sites irrespective of whether participants attended to the faces or tint of the feedback stimuli, but there was a slight change in the scalp distribution in that the attend-sex dFRN was more parietal than the attend-color dFRN suggesting that its amplitude overlapped with the enhanced P3 to reward feedback. This slight change in the scalp distribution likely reflects a less prominent dFRN in this experiment because the reinforcement learning systems that putatively generate the dFRN should be only moderately engaged. This moderate dFRN seems to be more susceptible to overlap from the P3 to rewards, and more susceptible to being obscured by increased noise in the attend-sex condition. The scalp distribution of the dN2 varied between frontal-central and lateral-occipital locations depending on which of these stimulus attributes they attended, however, there was enhanced processing at PO7 comparable to FCz in the attend-color condition. The dN2 at FCz should have been attenuated because of interference with the dFRN, possibly pushing activity down to levels comparable to activity in PO7. Furthermore, the posterior activity may have developed from processing of the irrelevant dimension. This question is addressed with my Control Experiment.

Experiment 4: Control Task

As a control, I ran an additional experiment that followed a more standard approach for eliciting the dN2 and dFRN. Namely, I dissociated face processing from color processing entirely by employing the same task as in the Moderate Experiment, but in one condition the stimuli consisted only of yellow and blue colors (without faces), and in a second condition the stimuli consisted of male and female faces (without colors) (see Fig. 7, Control). In principle, in the previous experiments the mere presence of the information on the unattended dimension could have influenced processing along the attended dimension, thereby disrupting the dN2 or dFRN. Thus the control experiment allowed for a pure assessment of these ERP components in a relatively standard oddball task. The Control Experiment was identical to the Moderate Experiment except for this change.

Method

Participants. 19 people (5 males) participated in this study.

Apparatus and Procedure. Stimuli and procedure were exactly the same as in the Moderate Experiment except that in the attend-sex task, monochromatic faces were presented as feedback, and in the attend-color task, blue and yellow rectangles (exact same size as the face stimuli) were presented as feedback.

EEG Data Acquisition and Analysis

The EEG data were acquired and analyzed, and the dN2 and dFRN were assessed as in the Passive, Active, and Moderate Experiments.

Results

Behavioral Results. Mean accuracy was 94.1% (*SD* = 9.7%) in the attend-color condition and 88.2% (*SD* = 16.9%) in the attend-sex condition. I eliminated the data from one subject whose accuracy scores on both the attend-sex and attend-color trials was more than 2 standard deviations

below the mean. For the remaining subjects, attend-color accuracy was 95.8% (*SD* = 6.1%) and attend-sex accuracy was 91.0% (*SD* = 12.0%). This difference was not significant ($p > .05$).

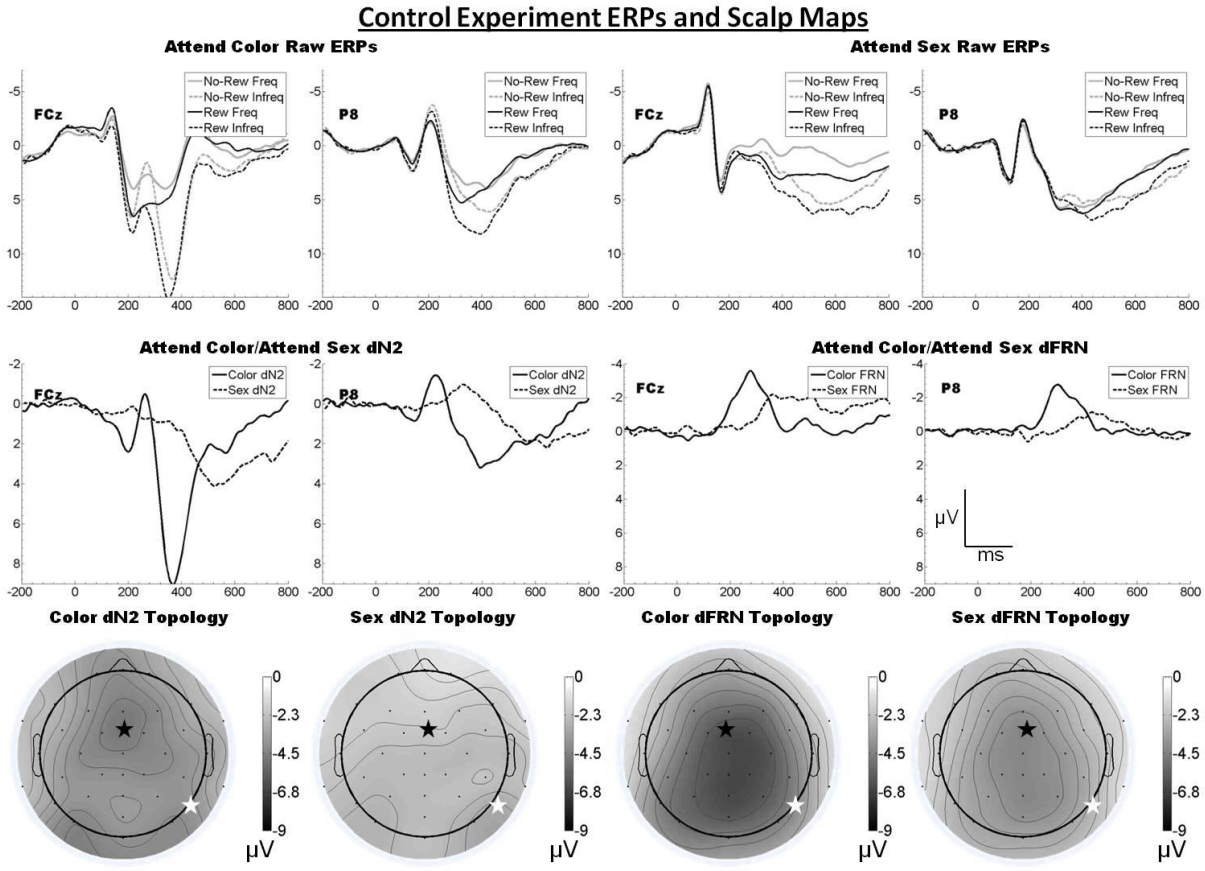

Figure 11. Grand average ERPs for experiment 4, recorded from channel FCz and P8 (see labels) and scalp distributions associated with the difference waves. The top row shows the "raw" ERPs for each of the frequency by reward conditions across tasks and electrodes. The middle row shows the dN2 and dFRN difference waves across task and electrodes. The bottom row shows the scalp distributions of the dN2 and dFRN across tasks. The scalp distributions reflect the base-to-peak measure of each of the dN2 and dFRN. The black star on the scalp map denotes channel FCz, and the white star denotes P8.Note that negative is plotted up.

EEG Results. The raw ERPs, difference waves and scalp maps are shown in Figure 11. Inspection of the scalp maps suggests the attend-color dN2 was maximal at frontal-central sites (FCz, -4.8 μV) and the attend-sex dN2 was maximal lateral-occipital sites (PO8, -3.0 μV). This impression was confirmed by A 2x2 ANOVA on dN2 amplitude with electrode and task as repeated

factors, indicating an effect of task such that the attend-color dN2 was larger than the attend-sex dN2 (-4.3 μV vs. -2.5 μV), $F(1,17) = 34.0$, $p < .001$, $\eta^2 = .67$, and an interaction of task and electrode such that the attend-color dN2 was larger at FCz than P8 (-4.8 μV vs. -3.8), whereas the attend-sex dN2 was larger at P8 than FCz (-3.0 μV vs. -2.0 μV), $F(1,17) = 18.0$, $p < .01$, $\eta^2 = .51$. One-tailed paired t-tests indicated that these differences were significant, attend-color: $t(17) = -2.5$, $p < .05$; attend-sex: $t(17) = 3.0$, $p < .01$.

By contrast, inspection of the scalp maps in Figure 11 suggests that the attend-color dFRN was shallowly distributed over central channels (CPz, -6.5 μV, FCz was the fifth most negative electrode, -6.3 μV). A paired t-test indicated no significant difference between these channels ($p>.05$). The attend-sex dFRN also appeared shallowly distributed over central channels (FCz, -4.3 μV, followed by CPz, -4.3 μV). Another paired t-test indicated no significant difference between these channels ($p> .05$). A 2x2 ANOVA on dFRN amplitude revealed an effect of electrode such that the dFRN was larger at FCz than at P8 (-5.3 μV vs. -3.5 μV), $F(1,17) = 12.8$, $p < .01$, $\eta^2 = .43$. The effect of task was also significant (-5.2 μV vs. -3.6 μV), $F(1,17) = 15.3$, $p < .01$, $\eta^2 = .47$, and there was no interaction of electrode and task ($p >.05$). One-tailed paired samples t-tests indicated that the dFRN was significantly larger at FCz than P8 in both the attend-color and attend-sex conditions, attend-color: -6.3 μV vs. -4.1 μV, $t(17) = -3.3$, $p < .01$; attend-sex: . -4.3 μV vs. -2.8 μV, $t(17) = -3.0$, $p < .01$.

Discussion

The Control Experiment confirmed the variability vs. stability of the dN2 and dFRN scalp distributions. Further, this experiment eliminated the potentially confounding influence of stimulus information along the unattended dimension. The left-lateral-occipital maxima observed in the Moderate Experiment was not reproduced in the Control Experiment, suggesting that some

processing of the irrelevant dimension in the Moderate Experiment contributed to exercising the FFA in the attend-color condition.

Across four experiments I have demonstrated that the topology of the dN2 changes reliably from a frontal-central to a lateral occipital distribution when task requirements increasingly engage lateral occipital face-processing areas. By contrast, the dFRN, when present, remained frontal-centrally or centrally distributed across these two conditions. This indicates that the observed variability of the dN2 scalp distribution does not result from a methodological flaw, nor is it associated with a general characteristic of ERP components in the 200 to 400 ms time range; if this were the case then the result would also be true of the dFRN. Furthermore, the dissociation in the behavior of the dN2 and dFRN supports the position that these ERP components manifest distinct neural processes, despite their considerable spatial and temporal overlap. The reinforcement learning theory of the dFRN holds that the dFRN is produced by DA signals being processed in the ACC, and as such, the dFRN should always exhibit a frontal-central scalp distribution consistent with this neural generator. In contrast, the modified LC-P3 theory holds that the dN2 is a manifestation of cortex-wide noradrenergic modulation, an index of the change in neural activity due to NE phasic release by the LC; it follows that the dN2 should exhibit a variable scalp distribution contingent on relative cortical activation at the arrival of NE. My results are consistent with these predictions.

Between-Subjects Analysis (Across Experiments)

Experiments 1-4 confirmed my prediction that the scalp distribution of the dFRN would remain frontal central whereas that of the dN2 would change according to task demands. Further, I found that the dFRN interfered with the dN2 in conditions where the dFRN was large. However, the specific nature of the interference remains to be investigated. To do so, I compared how these components varied across (rather than within) experiments to examine systematically the effects of increasing dFRN amplitude on the dN2. For this purpose I focused on the attend-color condition

where the effects of the interaction were greatest (because both components in this condition are frontal-central). Furthermore, I compared the results of the Passive, Active, and Moderate Learning Experiments, but not the Control Experiment, as the stimuli in the last experiment deviated from the first three and thus are not fully comparable. Finally, to investigate the specific mechanism driving changes in the dFRN and dN2 across experiments, I assessed the base-to-peak amplitude of "raw" N2 in each of the four conditions separately: frequent reward, frequent non-reward, infrequent reward, and infrequent non-reward. I predicted that, all other things being equal, the raw N2 would be larger to infrequent relative to frequent events (due to the LC), but that this increase would be attenuated in the case of infrequent rewards (due to overlap with the DA-driven reward positivity).

Method

We began with an across-experiment comparison of dN2 and dFRN amplitudes. Then, to analyze the raw N2s for each of the four reward by frequency conditions for each experiment, I quantified the size of the raw N2 base-to-peak as the change in voltage between the peak of the raw P2 and the peak of the raw N2. The N2 peak was assessed as the maximum negative amplitude in the ERP between 200 ms and 300 ms after onset of the feedback stimulus, and the raw P2 peak was assessed as the maximum positive voltage between 100 ms after onset of the feedback stimulus and the latency of the N2 peak for each subject and condition.

Lastly, I normalized N2 amplitude across subjects to assess within-subject variance in raw N2 amplitude across conditions. To do so, I converted the raw N2 values to z-scores as follows: For each subject, I determined the mean and standard deviation of the raw-N2 values across the infrequent no-reward, infrequent reward, frequent no-reward and frequent reward conditions. I then divided the difference between each raw-N2 value and the mean raw-N2 value by the standard deviation of the raw-N2 values (see Fig. 13 for raw (top) and normalized (bottom) means).

Results

A 3 x 2 mixed ANOVA with component (dN2 vs. dFRN) as a repeated factor and experiment (1-3) as a between-subjects factor revealed a significant main effect of component such that the dFRN was larger than the dN2 (-5.1 μV vs. -4.0 μV), $F(1,56) = 6.4$, $p < .05$, $\eta^2 = .10$. There was also an interaction of experiment and component indicating that the dFRN and dN2 changed in different ways across experiments, $F(3,56) = 10.2$, $p < .001$, $\eta^2 = .27$ (see Fig. 12). The between-subjects effect of experiment was not significant ($p > .05$).

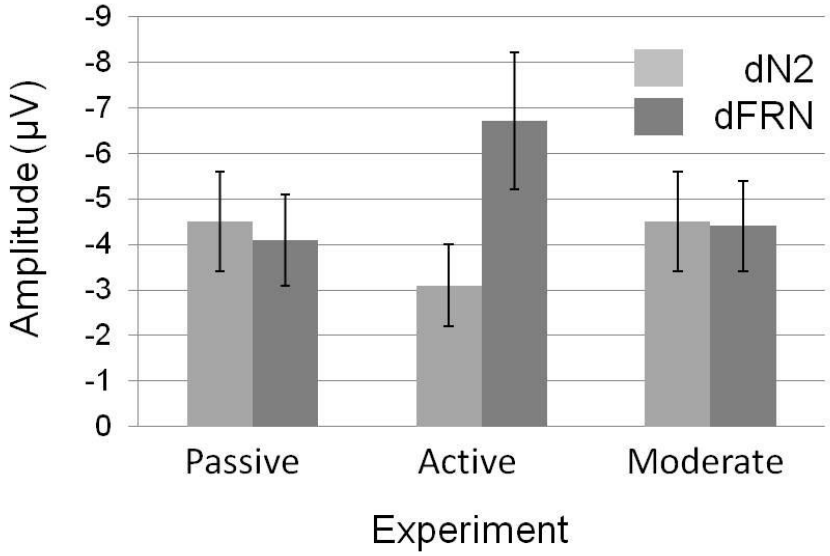

Figure 12. Mean dN2 and dFRN base-to-peak amplitudes across experiments. Note that negative is plotted up. Error bars represent 95% within-subjects confidence intervals.

We decomposed the interaction of component and experiment with a set of three two-tailed independent-samples t-tests (exp. 1 vs. exp. 2, exp. 1 vs. exp. 3, exp. 2 vs. exp. 3) for each component. The dN2 was larger in the Passive Experiment than in the Active Experiment (-4.5 μV vs. -3.1 μV), $t(38) = 2.0$, $p = .05$. The dN2 in the Moderate Experiment (-4.5 μV) was not significantly different than in the Passive Experiment, $p > .05$). The dN2 in the Moderate

Experiment trended toward being significantly larger than the dN2 in Active Experiment, $t(37) = 1.9$, $p < .10$. By contrast, the dFRN exhibited a different pattern across experiments. The dFRN in the Active Experiment was significantly larger than in the Passive Experiment (-6.7 μV vs. -4.1 μV), $t(38) = 3.0$, $p < .01$, and also significantly larger than the dFRN in the Moderate Experiment (-4.4 μV), $t(37) = -2.5$, $p < .05$, whereas the dFRN between the Passive and Moderate Experiments did not differ significantly ($p > .05$). Levene's test for equality of variances was significant for the dFRN comparison between the Passive and Active, and Active and Moderate Experiments, but the difference remained significant when the correction was applied ($p < .01$, $p < .05$, respectively). These results suggest that the Active Experiment was the critical experiment for demonstrating a change in the dN2 and dFRN across experiments: The dN2 was smallest in the Active Experiment and similar between the Passive and Moderate Experiments, whereas the dFRN was largest in Active Experiment and similar between the Passive and Moderate Experiments.

To investigate what caused the changes in the dN2 and dFRN amplitudes across experiments, I examined the normalized "raw" N2 values (see methods). I subjected raw N2 z-scores (Fig. 13, bottom) to a 2 x 2 x 3 mixed ANOVA with reward condition (reward vs. no-reward), frequency condition (infrequent vs. frequent) as repeated factors, and experiment (1- 3) as a between subjects factor. This analysis revealed a main effect of reward condition such that rewards produced a less negative raw N2 than no-rewards, $F(1,56) = 19.5$, $p < .001$, $\eta^2 = .26$, and a main effect of frequency condition such that the infrequent raw N2 was more negative than the frequent raw N2, $F(1,56) = 27.3$, $p < .001$, $\eta^2 = .33$. There was also an interaction of reward condition and frequency condition such that the effect of reward was larger in the infrequent condition than in the frequent condition, $F(1,56) = 6.1$, $p < .05$, $\eta^2 = .10$. In addition, Experiment interacted with both reward condition, $F(2,56) = 3.6$, $p < .05$, $\eta^2 = .11$, and frequency condition, $F(2,56) = 4.7$, $p < .05$, $\eta^2 = .14$, such that the difference between frequency and infrequent normalized raw N2s was smallest,

and the difference between reward and no-reward normalized raw N2s was largest in the Active Experiment.

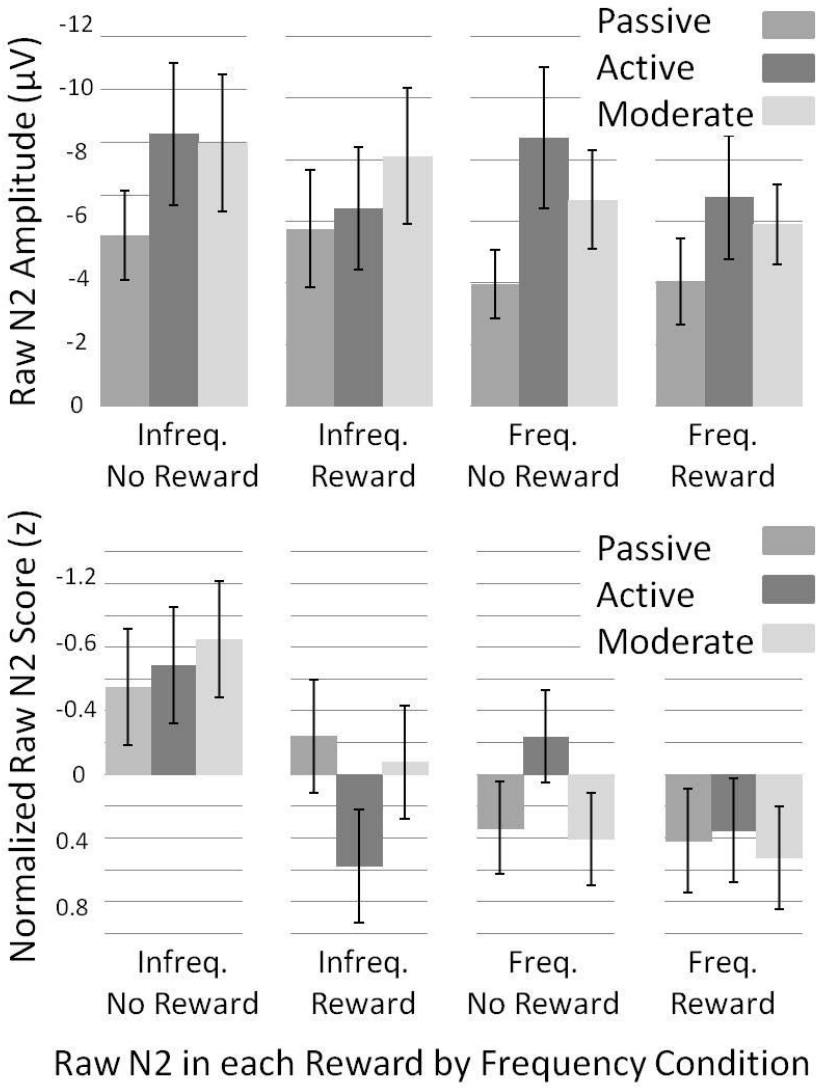

Figure 13. Normalized (bottom) and un-normalized (top) raw N2 *z* scores and amplitudes across experiments for each of the frequency by reward conditions. Note that negative is plotted up. Error bars represent 95% confidence intervals for the mean of each condition by experiment.

We used independent samples t-tests on normalized raw N2 values to uncover which of the four normalized raw N2s (infrequent no-reward, infrequent reward, frequent no-reward and frequent reward) best accounted for the change in the dFRN and dN2 in across experiments (Figure

13, bottom). The infrequent, no-reward raw N2 was similar across experiments (all *p*s > .05), as was the frequent, reward raw N2 (all *p*s >.05). The frequent, no-reward raw N2 was significantly larger in the Active Experiment than in the Passive Experiment, $t(38) = -3.0$, $p < .01$, whereas the infrequent, reward raw N2 was significantly smaller in the Active Learning than Passive Experiment, $t(38) = -3.1$, $p < .005$. Results were the same in comparing the Active Experiment with the Moderate Experiment: The frequent, no-reward raw N2 was larger in the Active Experiment, $t(37) = -3.1$, $p <.005$, whereas the infrequent reward raw N2 was smaller in Active Experiment, $t(37) = 2.9$, $p < .01$. There were no significant differences between the Passive Learning and Moderate Experiments (all *p*s>.05). These results suggest that the apparent need for greater deliberative strategy in the Active Experiment produced a larger negativity to frequent no-reward trials, and a greater attenuation of the raw N2 on infrequent reward trials.

Discussion

We examined how challenging subjects with an involving learning and decision making task impacted dN2 and dFRN amplitude across experiments. I demonstrated that across three experiments, the task that most engaged a deliberative learning strategy enhanced the dFRN and simultaneously attenuated the dN2, albeit the latter finding only trended toward statistical significance for the comparison between the Active versus Moderate Experiments. I also examined the raw N2s that underlie the dN2 and dFRN as a function of frequency, reward, and learning engagement. I converted the raw N2 values to z-scores to control for between-subjects variability in the overall size of the raw N2, and then analyzed how the normalized N2s for each of the frequent reward, frequent no-reward, infrequent reward, and infrequent no-reward conditions changed across experiments according to task demands. The interaction of reward and frequency on raw N2 amplitude indicates that these factors both contribute to the dN2 and dFRN amplitudes. That is, the effect of infrequent no-rewards versus rewards is larger than the effect of frequent rewards versus

no-rewards[5]. This is consistent with the prediction of the reinforcement learning theory of the FRN that infrequent rewards should elicit a greater positivity than frequent rewards, and that the infrequent no-rewards should elicit a greater negativity than frequent no-rewards. Finally, independent samples t-tests on raw N2 z-scores indicated that the larger dFRN and smaller dN2 in the Active Learning experiment were driven both by a larger raw N2 to frequent no-rewards and greater attenuation of the raw N2 to infrequent rewards compared to the Passive Learning and Moderate Experiments. These contrasting changes work synergistically to increase the amplitude of the dFRN but against each other to decrease the amplitude of the dN2.

General Discussion

The modified LC-P3 theory holds that the dN2 is produced by the impact of a brief, cortex-wide increase in cortical NE due to phasic LC firing in response to infrequent, task-relevant events. In support of this, I demonstrated that the dN2 exhibits a scalp distribution that changes according to task specifics in a manner consistent with a noradrenergic origin. Furthermore, the modified LC-P3 theory and the reinforcement learning theory of the dFRN together hold that the dN2 and dFRN are driven independently by modulation of the raw N2 at frontal-central channels by both the LC-NE system and the DA system. Whereas NE amplifies the raw N2, DA depresses it, such that these influences interfere with one another in producing scalp potentials over anterior regions of the scalp. I demonstrated that factors that exercise learning and decision making systems enhance the dFRN but attenuate the dN2. Furthermore, I provided evidence that these changes in the dFRN and dN2 are driven most strongly by an enhanced negativity to frequent no-rewards and an enhanced positivity to infrequent rewards.

[5] Note that for the purpose of brevity I did not include analyses of the effect of reward on the dN2 difference wave and of the effect of frequency on the FRN difference wave, although both of these effects were statistically significant.

As the name indicates, my account of the relationship between LC-NE system activity and the dN2 is a modification of the original LC-P3 theory (Nieuwenhuis et al., 2005a). Below I review the LC-P3 theory in detail and provide the motivation for my modification to it.

The Original LC-P3 Theory

The P3 is a positive deflection in the ERP typically peaking approximately 300 ms to 500 ms after the eliciting stimulus. It has a broad, parietal scalp distribution that is thought to represent the summation of activity in multiple, dispersed neural generators (e.g. Johnson, 1993). Nieuwenhuis and colleagues (2005a) characterize four main categories of conditions that influence P3 amplitude: Subjective probability (unexpected events elicit a larger P3 than expected events), motivational salience (targets elicit a larger P3 than distracters), applied attention (attended stimuli elicit a larger P3 than ignored stimuli, and targets elicit a larger P3 under conditions that demand full attention compared to dual-task conditions), and attention-capturing stimuli (task-irrelevant stimuli that are highly deviant from the stimulus context elicit a larger P3 than less deviant stimuli).

The LC-P3 theory (Nieuwenhuis et al., 2005a) proposes that the P3 is an electrophysiological manifestation of cortex-wide noradrenergic modulation through the LC efferent projection system. In support of the LC-P3 theory, Nieuwenhuis and colleagues presented a comprehensive review of the literature, marshalling abundant evidence that conditions antecedent to phasic LC firing are the same as those conditions that exercise the P3. Additionally, Nieuwenhuis and colleagues (2005a) refer to psychopharmacological and animal lesion studies for support for the link between the P3 and noradrenergic modulation. By and large, noradrenergic agonists such as clonidine and direct lesions of the LC have been reported to reduce the amplitude of a P3-like potential observed in monkeys (e.g. Pineda, Foote, & Neville, 1989; Pineda & Westerfield, 1993; Swick, Pineda, & Foote, 1994). Further, in one psychopharmacological study involving human participants, Halliday and colleagues (1994) found that clonidine reduced the amplitude of the P3.

Issues with the Original LC-P3 Theory

The LC-P3 theory possesses considerable explanatory power and accounts for a wide range of existing data. However, three issues warrant further examination. First, neurophysiological evidence indicates that the NE phasic burst arrives in cortex too early to produce the P3 directly. Aston-Jones and Cohen (2005) suggest NE should reach the cortex within approximately 170 ms of target onset, but the P3 typically doesn't begin for another 50 ms until about 220 ms following target onset (and reaches maximum amplitude from about 300 – 600 ms post-stimulus). Thus there is greater than a 50 ms discrepancy between the time of NE arrival in cortex and the onset of the P3. Although this estimate of the timing of NE arrival is based on single-cell recordings in monkeys, Aston-Jones, Foote, and Segal (1985) demonstrated that conduction speeds in NE-releasing neurons vary across species such that the actual timing of NE arrival in cortex is relatively preserved despite varying axonal distances. Further, P3 onset sometimes occurs after motor response initiation suggesting that the underlying mechanism does not directly implement the stimulus-response mapping (as would be expected if it reflected a signal detection process mediated by the LC) but rather is involved in a subsequent, related process (Duncan-Johnson & Donchin, 1982; Krigolson, Holroyd, Van Gyn, & Heath, 2008; Ritter, Simson, Vaughan, & Friedman, 1979). For example, Krigolson and colleagues (2008) found that when a target changed location in a continuous tracking task, participants adjusted their motor behavior accordingly even before the change in target location elicited the P3. In contrast, studies in monkeys indicate that (unlike the P3) phasic LC activity consistently precedes behavioral responding and has been strongly associated with processes that lead to the response (e.g. Clayton, Rajkowski, Cohen, & Aston-Jones, 2004; Rajkowski, Majczynski, Clayton, & Aston-Jones, 2004).

A second issue with the original LC-P3 theory is that the auto-inhibitory nature of the LC-NE system produces a biphasic change in cortical NE levels: A flood of NE followed by a depletion

of NE due to the refractory-like period of the LC. If enhanced signal detection produces electrophysiological changes detectable at the scalp (the P3), then degraded signal detection associated with reduced NE supply should presumably likewise be visible in the ERP. However, no such deflection is visible in the ERP following the P3.

A third issue with the original LC-P3 theory is related to the "attentional blink," a deficit in stimulus processing attributed to the LC refractory period (e.g. Nieuwenhuis et al., 2005a, 2005b; Usher et al., 1999; Warren et al., 2009). When two targets are embedded within a rapid serial visual presentation, the first target can be reported with high accuracy, but the second target is reported with significantly worse accuracy if it is presented within a window 200 ms to 600 ms after onset of the first target (Raymond, Shapiro, & Arnell, 1992). The spared accuracy for the second target if it appears within 200 ms of onset of the first is termed "lag-1 sparing." Nieuwenhuis and colleagues (2005a, 2005b) argued that the properties of the LC-NE system could account for the attentional blink: They proposed that the onset of the first target elicits LC-NE system phasic response, with the subsequent flood of NE to the cortex benefitting processing of the first target, and also the second target if the second target is presented within ~ 100 ms of the first. After this initial period of effective target processing (~200 ms from onset of the first target to offset of a second target presented 100 ms later), the LC is inhibited and cortical levels of NE are not sufficient to process effectively any targets presented 200 ms to 600 ms after onset of the first target. Consistent with this proposal, MacArthur, Budd, and Mitchie (1999) showed a significant negative correlation between the amplitude of the first-target P3 and second-target accuracy, such that the size and temporal profile of any subject's attentional blink mirrored the size and temporal profile of that subject's first-target P3; according to the LC-P3/LC-AB theories, larger bursts of NE to the first target result in larger P3s, followed by a relatively deep or extended refractory periods that produces a larger attentional blink.

Nevertheless, this proposal raises the question: If the P3 reflects the LC-NE burst, which should facilitate stimulus processing, then why does the P3 peak during the period of the attentional blink (between 300 ms and 500 ms after the first target), which by definition is a period of impaired stimulus processing? Instead, one might predict that the electrophysiological manifestation of LC-NE activity would precede the attentional blink, during the time of effective of stimulus processing known as lag 1 sparing.

The Modified LC-P3 Theory

Nieuwenhuis and colleagues provide alternative explanations for these apparent discrepancies within the LC-P3 account (see Nieuwenhuis et al., 2005b; Nieuwenhuis, & Jepma, 2010). However, the assumption that the dN2 (rather than the P3) reflects the LC-induced enhancement of cortical processing eliminates these issues outright. According to this position, the subsequent P3 reflects the cortex-wide depletion of NE due to the refractory-like period of the LC-NE system. This modification accounts for the three issues above as follows. First, it aligns the timing of the putative ERP response to LC-NE activity (the dN2, occurring at about 200 ms post-stimulus, as opposed to the P3, which occurs about 300 ms – 600 ms post-stimulus) with the actual timing of the phasic LC-NE signal (about 170 ms post-stimulus). Furthermore, unlike P3 onset, the onset of the N2 consistently precedes the overt behavior (e.g. Krigolson et al., 2008); in fact, detailed analyses of RT data to compatible stimuli in a speeded response compatibility task suggests that stimulus information begins to impact the response selection mechanism at about 170 ms post-stimulus and exerts the maximal impact on the response selection process about 250 ms post-stimulus (Holroyd, Yeung, Coles, & Cohen, 2005). Second, the modified theory provides an ERP correlate of both NE abundance (the dN2) *and* NE depletion (the P3) in the cortex. And third, the modified theory associates the P3 with the period of impaired cortical processing due to NE depletion (rather than abundance), which corresponds naturally to the time profile of the attentional

blink. This position also naturally accounts for the evidently close relationship between the N2 and the P3, because the duration of the refractory period of the LC (P3) is directly related to the size of the initial NE burst (N2).

Critically, the evidence reviewed by Nieuwenhuis and colleagues (2005a) as support for the original LC-P3 theory applies equally well to the modified LC theory. For example, the amplitudes of both the P3 and the N2 are sensitive to the same factors: The P3 is typically preceded by the N2, and in early studies of the impact of stimulus probability on the ERP, these two ERP components were collectively termed the N2/P3 complex because of their tendency to co-vary in amplitude and latency (e.g. Duncan-Johnson & Donchin, 1977; see also Ritter et al., 1979). In fact, all of the antecedent conditions noted by Nieuwenhuis and colleagues to apply to both P3 amplitude and LC phasic activity also apply to N2 amplitude. Thus, both N2 amplitude and P3 amplitude increase with increasing unexpectedness of a task-relevant event, and both are larger to targets than non-targets (e.g. Courchesne, Hillyard and Galambos, 1975; Duncan-Johnson & Donchin, 1977; Nieuwenhuis et al., 2003; Ritter et al., 1979; Simson, Vaughan, & Ritter, 1976; Squires, Squires, & Hillyard, 1975; Squires, Wickens, Squires, & Donchin, 1976). Both also scale to the amount of attention paid to a stimulus, with a larger N2 and larger P3 to attended versus unattended stimuli, and to attention-capturing/highly deviant stimuli versus less deviant stimuli (Courchesne, Hillyard and Galambos, 1975; Daffner, Mesulam, Scinto, Calvo, Faust, & Holcomb, 2000; Daffner, Scinto Calvo, Faust, Mesulam, West, & Holcomb, 2000; Folstein, van Petten, & Rose, 2008; Ford, Roth, & Kopell, 1976; Hillyard, Squires, Bauer, & Lindsay, 1971; Squires, Donchin, Herning, & McCarthy, 1977; Squires, Squires, & Hillyard, 1975).

To summarize, evidence supporting the original LC-P3 theory is equally supportive of the modified LC-P3 theory. In addition, the modified LC-P3 theory reconciles three discrepancies

associated with the original theory while simultaneously providing an explanation of why the N2 and P3 ERP components co-vary in amplitude and latency.

Interaction of the LC-NE and DA Systems

The reinforcement learning theory of the FRN holds that DA dips and bursts modulate ongoing activity in the ACC. Specifically, reward feedback elicits a phasic burst of DA that produces a positivity in the ERP typically between 200 ms and 300 ms of the eliciting stimulus (Holroyd et al., 2008), whereas no-reward feedback elicits a dip in DA that produces a negative deflection in the same time range (Holroyd & Coles, 2002). Furthermore, the theory holds that DA signals scale according to the degree of expectedness of the feedback, such that infrequent rewards elicit a larger DA burst and reward positivity than frequent rewards, and infrequent no-rewards elicit larger DA dips and negative deflections than frequent no-rewards. Critically, the theory proposes that the dFRN reflects DA-dependent modulation of ACC activity but does not specify exactly what neural process is being modulated. However, empirical evidence suggests that the ACC produces a negative deflection (the N2) that perhaps reflects response conflict or a related stimulus/response decision-making process (e.g. Botvinick et al., 2001). Thus it has been argued that the reward positivity elicited by phasic DA activity attenuates the N2 produced in the ACC (Holroyd, 2004; Holroyd et al., 2008). Prior to this work, it had yet to be determined whether dopamine dips increase N2 amplitude by disinhibiting ACC activity as originally proposed (Holroyd & Coles, 2002); this may be because the phasic decreases from baseline activity of the DA system are relatively shallow when compared to large increases in DA activity associated with phasic bursts.

According to the modified LC-P3 theory, infrequent events elicit a phasic release of NE that enhances cortical processing and produces an amplified negative deflection in the ERP between about 200 ms and 300 ms after onset of the eliciting stimulus. Critically, NE modifies activity in the same time range as the putative DA signals, including ACC activity when it is present. According to

this position, increased NE gives rise to a larger N2 produced in the ACC. Thus, the two factors push and pull the frontal-central N2 associated with ACC activity up and down.

Here I examined the interaction of the LC-NE and DA systems by including both a frequency and reward manipulation within the same experiment. I hold that frequency insofar as it relates to the expectedness of reward or no-reward has an effect on the DA system independent of its effect on the LC-NE system. For the LC-NE system, infrequent events consistently increase NE release and the associated negativity, whereas for the DA system infrequent rewards produce a relatively large burst in DA and associated positivity and infrequent no-rewards elicit a relatively large dip in DA and an associated negativity. Consistent with this, I observed a significant interaction of frequency and reward such that the difference in N2 amplitude between reward and no-reward was larger when rewards and no-rewards were infrequent relative to when they were frequent. This replicates previous work on the effect of reward-expectedness on the dFRN (Baker & Holroyd, 2009; Hajcak, Moser, Holroyd, & Simons, 2007; Holroyd et al., 2009; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003).

We further examined the interaction of the LC-NE and DA systems by systematically manipulating the degree of participant engagement in reward tasks—and by extension putatively the degree of DA system engagement—across three experiments. In doing so, I observed an interaction of experiment and ERP component, such that the dFRN was largest and the dN2 was smallest in the Active Experiment. Further, I decomposed the dFRN and dN2 across the three experiments by examining the raw N2s to frequent versus infrequent rewards and non-rewards. In keeping with the reinforcement learning theory of the dFRN, I predicted that the enhanced dFRN in the Active Experiment would be driven by both a greater attenuation of the raw N2 to reward feedback and by a greater enhancement of the raw N2 on no-reward trials. Consistent with this, independent-samples t-tests on normalized raw N2 amplitude indicated that the raw N2 to infrequent reward feedback

was significantly smaller (DA-associated positivity attenuates the raw N2) in the Active Experiment than in both the Passive and Moderate Experiments, and the raw N2 to frequent no-reward feedback was significantly larger (DA dip enhances negativity) than in both the Active Learning than Passive and Moderate Experiments. These differences cannot be attributed to greater engagement of the LC-NE system in the Active Experiment to both infrequent and frequent events, because although this would explain the greater negativity observed to frequent no-reward feedback, it is inconsistent with the smaller negativity to infrequent reward feedback.

The finding of a decrease in the amplitude of the raw-N2 to infrequent reward feedback in the Active Experiment relative to the other experiments is expected in light of previous work associating reward processing with a positive deflection in the ERP that attenuates the N2 (e.g. Baker, & Holroyd, 2011; Holroyd et al., 2008). However, the finding of an increased raw N2 to frequent no-reward feedback in the Active Experiment relative to the other experiments to my knowledge constitutes the first evidence of an increased negative deflection elicited by no-rewards feedback. Although the reinforcement learning theory of the dFRN holds that brief decreases in DA activity in response to unexpected no-reward feedback increase the amplitude of a negative deflection in the ERP (Holroyd & Coles, 2002), the dFRN difference wave approach cannot determine whether the difference between the ERPS is due to a positivity to rewards, a negativity to no-rewards, or both (Holroyd et al., 2008). My results indicate that the raw N2 is in fact increased to no-reward feedback as predicted by the reinforcement learning theory of the FRN (Holroyd & Coles, 2002), especially under conditions that demand high task engagement.

The raw N2s to infrequent no-rewards and frequent rewards were not statistically different between the Active Experiment and either of the Passive or Moderate Experiments. This raises the question: Why are frequent no-rewards and infrequent rewards particularly sensitive to changes in deliberative strategy, whereas infrequent no-rewards and frequent rewards are relatively insensitive?

The answer may have to do with the fact that both the infrequent no-reward feedback and the frequent reward feedback always occurred in the same blocks of trials. In this context rewards accumulated frequently (and therefore non-rewards were infrequent). Perhaps this condition of the Active Experiment is much like the Passive and Moderate experiments, in the sense that subjects could disengage from the task because it was apparently easy. By contrast, subjects would have remained engaged in the blocks where rewards were infrequent and non-rewards frequent. Hence with increasing engagement of learning and decision systems across experiments, subjects may have been similarly "unfazed" by rewards and no-rewards in easy blocks but differentially reactive to rewards and no-rewards in difficult blocks.

It is also possible the ACC activity may be characterized by floor and ceiling effects in terms the range of activation elicited by each experimental context. Consider what DA and NE appears to do in each of the four frequency by reward conditions. In the infrequent no-reward condition, a large dip in DA disinhibits the ACC producing greater activity, and simultaneously a larger NE release further increases that activity. It is reasonable to imagine that two processes acting in concert to drive ACC activity upwards could cause the amplitude of the N2 to reach a maximum, even when the DA system is only minimally or moderately engaged by the task. By contrast, in the frequent reward condition, a small burst of DA inhibits the ACC, while simultaneously minimal NE release does little to enhance ACC activity. In this case the ACC is quiescent and varying the engagement of the DA system cannot quiet it any further, resulting in a consistently small N2. The variability across conditions occurs when the DA and NE systems drive ACC activity in opposite directions. Infrequent rewards produce a large burst in DA that inhibits the ACC, but at the same time a larger release of NE enhances ACC activity; together these competing processes maintain N2 amplitude in a range between the floor and ceiling, revealing changes in the relative impact of each of the systems. Similarly, frequent no-rewards produce a small dip in DA that disinhibits the ACC and a small

release of NE that has little impact on ACC activity, producing an N2 that is small but consistently above floor.

Other Issues

These experiments were intended to exercise the system that produces the dFRN differentially -- most in the Active Experiment, least in the Passive Experiment, and to an intermediate degree in the Moderate Experiment[6]. Whereas the Active Experiment elicited the largest dFRN as expected, the dN2 and dFRN results from the Passive Experiment and Moderate Experiment did not differ significantly. In the Active Experiment, subjects were required to pick one of two cards (out of six possible cards; 15 possible combinations), and were told the cards had specific relationships with each other such that in any combination one card was associated with a better chance of winning than the other. Specifically, they were told to try to figure out the right card to pick in each situation (pairing) as opposed to which card was most "powerful" and would win in any given situation. In passive learning, there was no decision to be made at all, so any learning was a consequence of a passive evaluation of the reward frequency in a given block of trials. In the Moderate Experiment, subjects chose between a left or right key press with the cover story that one key was associated with a better chance of winning on any given trial, and that the "winning" key was stable for approximately 20 trials whereupon it would be reassigned randomly (sometimes remaining where it was, sometimes changing). Note that the Active Experiment was characterized by 15 potentially learnable relationships, whereas the Moderate Experiment was characterized by only two such relationships (left button vs. right button). Thus although the degree of engagement in the Moderate Experiment was likely larger than in the Passive Experiment, this difference may have been small relative to the Active Experiment.

[6] In fact, a linear regression analysis demonstrated dFRN amplitude was significantly predicted by task engagement across experiments, while the dN2 exhibited a trend toward the inverse relationship, see Appendix.

A second notable issue is the fact that the base-to-peak measure of the raw N2s can fail to capture some variability in the ERP due specifically to the reward positivity. That is, the base-to-peak method is insensitive to positive deflections that go beyond attenuating the N2 to create a positive deflection in the same time range: The most positive value for the raw N2 that can be assessed is 0 μV, because the algorithm finds the most negative value in the N2 time window, and subtracts from that the most positive value preceding it. Despite this limitation, the method nevertheless yielded significant differences in the raw N2 across experiments that confirmed my hypotheses.

Conclusion

Both the LC-NE system and the DA system modulate processing in the ACC. However, whereas the LC-NE system includes the ACC among many cortical targets, innervation by the DA system of frontal midline cortex is especially great. Consistent with this distinction, I demonstrated that the dN2, an ERP component that I propose reflects noradrenergic modulation of cortical activity, exhibits a scalp distribution that is maximal at varying locations dependent on the relative engagement of specific cortical areas. By contrast, the dFRN, which has been associated with DA system activity, is consistently maximal at scalp locations over the ACC. Furthermore, I demonstrated that under conditions in which the DA system should be highly engaged – specifically, in an apparently complex learning and decision making task – the neural processes underlying the dFRN and dN2 appeared to interfere with one another such that the dFRN was enhanced and the dN2 was attenuated. Finally, I demonstrated that negative feedback stimuli (i.e., feedback associated with the absence of a potential reward) were associated with a negative deflection in the ERP that was larger than the raw N2 typically elicited by motivationally salient events. Taken together, these results paint a picture of two neuromodulatory systems that have relatively independent effects on the ERP despite considerable overlap in the space and time domains as well as shared antecedent

conditions. The ACC seems to be at the center of this overlap, with its putative role in recruiting the LC-NE system to facilitate processing, with its processing subsequently facilitated by NE, and with its activity putatively inhibited by DA bursts and excited by DA dips. These considerations suggest the ACC plays a crucial role in both fast and efficient processing of task relevant events, and adaptive decision making based on a reinforcement history implemented by the DA system.

General Discussion

The work reported here contributes to the current understanding of the LC-NE system, the DA system, and associated ERPs with several novel ideas and findings. I propose that the N2 ERP component is a manifestation of noradrenergic modulation, and demonstrate that the scalp distribution of the dN2 changes according to task specifics in a manner predicted by this position. When subjects attended to the color of tinted male and female faces, the negativity associated with infrequent events was maximal over frontal-central regions, whereas when they attended to the sex of the faces it was maximal over lateral-occipital regions. This result was obtained despite the fact that the stimuli and the probability manipulation were exactly the same across task conditions. To my knowledge, this is the first example that the scalp distribution of the oddball N2 can move from a frontal-central to a lateral-occipital distribution simply by directing attention to a different aspect of the stimulus, even when the task demands otherwise remain constant.

I also dissociate the dN2 from the dFRN. The modified LC-P3 theory and the RL-ERN theory are distinct theories of two ERP components that in typical circumstances occur in the same time range and in the same scalp area, raising concerns that the components themselves may not be distinct. In order to understand both theories, it is essential to explore how the associated ERP components interact. I demonstrate that these components are independent by showing that conditions which relatively exercise the dFRN attenuate the dN2, and that whereas the dN2 has a variable scalp distribution, the dFRN does not. Furthermore, I present the first experimental evidence that a failure to obtain reward is associated with an increase in N2 amplitude, i.e., a negative deflection above and beyond the negativity associated with infrequent or unexpected events. When subjects were presented with frequent feedback indicating no-reward, increasing the engagement of learning across experiments increased the associated negativity. According to the RL-ERN theory, a dip in dopamine following negative feedback caused this increase in the N2. This

result should not be interpreted as greater engagement of the LC-NE system, because the negativity associated with infrequent rewards was not similarly enhanced.

General theories of LC-NE system function

The modified LC-P3 theory is consistent with the original LC-P3 theory and with Nieuwenhuis et al.'s LC-AB theory in terms of the effect of noraderenergic modulation on neural processing. That is, the modified LC-P3 theory holds that noradrenergic modulation enhances the signal-to-noise ratio of target neural networks, producing a brief period of especially effective neural processing. This position is also consistent with another more general theory of LC-NE system that describes the function of both LC phasic and tonic activity (Aston-Jones & Cohen, 2005) . However there are other theories of LC-NE system function that posit phasic bursts of NE are directed toward facilitating a change in the organization of target neural networks, as opposed to facilitating signal detection in those networks. I briefly describe all three theories below.

Adaptive gain theory. Usher et al. (1999) demonstrated two modes of LC activity: Low tonic activity that is characterized by pronounced phasic bursts in response to motivationally salient events, and high tonic activity characterized by less pronounced phasic bursts of firing. The low-tonic/high-phasic mode of activity is associated with good performance on a target detection task, whereas the high-tonic/low-phasic mode is associated with poor performance and erratic behaviour. Based on the ideas and results presented in Usher et al. (1999), Aston-Jones and Cohen (2005) proposed the adaptive gain theory of LC-NE system function. Under the adaptive gain theory, low-tonic/high-phasic activity is optimum for the exploitation of a learned/effective pattern of stimulus and response mappings. This mode promotes effective responding to events that elicit phasic bursts, while simultaneously limiting NE release to irrelevant events. In contrast, high-tonic/low-phasic activity is optimum when an organism finds itself in an unknown environment and must explore many possible stimulus-response mappings to discover the most effective pairings (or when a

learned/effective pattern of stimulus and responses mappings has lost utility, such as if the environment has changed so that actions which formally led to reward do so no longer). The high-tonic/low-phasic mode promotes exploratory responding to events that have not yet been deemed motivationally salient, without giving preference to events which may have been motivationally salient in the past.

The unexpected uncertainty theory of LC-NE system function. Yu and Dayan (2003, 2005, Dayan & Yu, 2006) proposed that increases in LC firing signal unexpected uncertainty. Unexpected uncertainty is defined in relation to expected uncertainty, where expected uncertainty refers to an event that may be uncommon, but does not indicate that an organism should reassess what it knows about its environment. For example, in a probabilistic learning task, if a subject has selected a given box many times and has come to believe such a selection is rewarded about 80% of the time, infrequent no-rewards will be synonymous with expected uncertainty – the subject is not certain any given choice will be rewarded, but does expect a percentage of selections will be rewarded. Yu and Dayan link expected uncertainty to cholinergic activity (2003, 2005).

By contrast, unexpected uncertainty refers to events that suggest the organism's current notions about the environment may no longer apply. For example if the subject in the probabilistic learning task receives several no-rewards in a row, they may begin to suspect that the reward contingencies have changed, and that their favoured selection is no longer paying off 80% of the time. This unexpected uncertainty, according to Yu and Dayan, triggers greater NE release which promotes exploratory behaviour in a manner consistent with adaptive gain theory. Dayan and Yu (2006) have also argued that phasic NE release promotes rapid change from one behavior to another within any given task, such as moving from ignoring distracters to remembering targets in the attentional blink task.

The Network Reset Theory of LC-NE System Function. Bouret and Sara (2005) proposed phasic bursts of NE release serve to interrupt activity in target neural networks to promote reorganization of the elements into new functional networks. Thus, Network Reset Theory holds that the function of phasic NE release is to promote a change in current behaviour or a current pattern of beliefs about the environment/task. This change in behaviour could be as simple is switching from inaction to action when a motivationally salient stimulus is presented.

The modified LC-P3 theory vs. other theories of the LC-NE system. Insofar as the modified LC-P3 theory asserts that noradrenergic modulation produces the N2 ERP component, and the subsequent refractory period produces the P3, the theory is not inconsistent with any of the above theories. It can be true that NE promotes a network reset or a behavioural change and at the same time that NE produces the N2.

Implications of the modified LC-P3 theory for other ERP components

An ERP component is typically defined by its polarity, latency, scalp distribution and the particular experimental manipulation that produces it (Luck, 2005). I have argued that the LC-NE system enhances ERP components, producing its own electrophysiological manifestation (the dN2 followed by a P3, together forming the N2-P3 complex) that is not specific to any cortical generator. However, this position does not take anything away from the importance of spatially-specific ERP components. The cortical areas most engaged by a given task will be most impacted by NE, but insofar as those cortical generators are activated by the task, they can inform us of the timing and location of specific kinds of neural processing. Critically, understanding the impact of noradrenergic modulation on the ERP is essential for understanding any ERP component within the time range of NE-mediated changes in cortical activity, to distinguish between a change in the ERP component due to a manipulation that impacts NE release, and a change in the ERP component due to a manipulation that impacts cortical activity independent of NE. For example, the relative probability

of stimulus events will drive NE activity and produce differences in the ERP in any region of the cortex that is engaged by the task. Tanaka et al. (2006) interpreted an enhanced N250 to target faces as indexing familiarity for the face. However, I have shown how a simple probability effect could explain the result. In this case, a difference in relative levels of cortical NE confounds the conclusion that the FFA is particularly sensitive to familiar faces.

Noradrenergic modulation can also impact the peak latency of an ERP component. When NE amplifies an ERP component, the resulting peak of the component could reflect either the timing of maximal engagement of that cortical area, or the timing of maximal amplification by NE. Therefore, a full understanding of how NE may be impacting the ERP is necessary to interpret peak-latency data.

In addition, the prominence of the N2-P3 complex in ERPs elicited by a multitude of experimental manipulations supports the modified LC-P3 theory in that it provides additional evidence that the N2-P3 complex is produced by a pervasive neural process, such as noradrenergic modualtion. In particular, several negative-going ERP components seem to overlap with the N2-P3 complex, as evidenced by the characteristic negative-positive deflection that often appears with these ERPs.

The Mismatch Negativity. Closely related to the oddball N2 is the mismatch negativity (MMN), a negative deflection typically peaking between 150 ms and 250 ms after presentation of an infrequent tone amidst a stream of frequent, standard tones (Näätänen, Gaillard, & Mäntysalo, 1978; Näätänen, Paavilainen, Rinne, & Alho, 2007; Näätänen, Simpson, & Loveless, 1982; Sams, Alho, & Näätänen, 1983; Sams, Alho, & Näätänen, 1984). The infrequent tones are typically presented on only 10% of the trials, sometimes even more seldom, (e.g. Näätänen et al., 1982; Sams et al., 1984), however the MMN is also observed when the two stimuli are presented with equal probability and the MMN is assessed by subtracting the ERP to a repeated stimulus from the ERP to a mismatch

(Sams et al., 1983). This finding is very similar to the finding from the oddball N2 paradigm, that the N2 and P3 are exercised not only by globally infrequent events (i.e. within a block of 100 trials), but also by events that are infrequent in the context of the preceding few trials (Squires et al., 1976). Indeed, adaptation of the same paradigm from Squires et al. (1976) produced a MMN that was similarly exercised by its relative infrequence in the preceding five presentations (Sams et al., 1983). Despite this commonality, the MMN is believed to be a different ERP component for three reasons. The MMN peaks a little earlier than the typical visual-oddball-N2, has a more frontal scalp distribution than the visual-oddball-N2, and, most notably, the MMN is observed in response to ignored stimuli while the visual oddball N2 is usually not (ignored stimuli in the visual domain do not elicit a N2 or a P3 unless they capture attention by virtue of being emotionally salient or highly deviant in the experimental context, such as produce the novelty N2 as described in the introduction) (Näätänen et al., 2007). Näätänen, Gaillard, and Mantysalo (1978) presented independent series of tones simultaneously to each ear of subjects, with the instructions to attend to and detect rare deviations in only one ear. Deviations in the unattended ear elicited a MMN and a P3, but the P3 was reduced compared to the attended ear. Further, Näätänen and colleagues later showed the MMN appeared to infrequent tones even when subjects ignored all tones altogether and instead read a book, but no P3 was observed in some of these experiments (Näätänen, Simpson, & Loveless, 1982; Sams, Alho, & Näätänen, 1984). Näätänen (2001; 2007; see also Sams et al., 1984) has argued that infrequent tones elicit both a MMN and a N2-P3 complex, but that the MMN is elicited independent of attention whereas that stimuli must be attended to in order to fully exercise the N2-P3 complex. Importantly, the MMN is maximal over frontal electrode site Fz, and when the MMN overlaps with the dN2, it is enhanced at Fz (Näätänen, et al., 1978; Näätänen et al., 1982; Sams et al., 1984), suggesting amplification by NE.

The N250. When subjects are required to discriminate target faces from a sequence of distracter faces, a negativity peaking approximately 250 ms after stimulus onset is observed over occipital-temporal sites, followed by an enhance P3 at central-parietal sites (Tanaka, Curran, Porterfield, & Collins, 2006). This finding is interpreted as a variation of the repetition priming N250, an ERP elicited by a repeated face that is also maximal over occipital-temporal sites (e.g. Schweinberger, Pfütze, & Sommer, 1995; Schweinberger, Pickering, Jentzsch, Burton, & Kaufmann, 2002). Both N250 variants are thought to be generated in the FFA, an area known to be especially active and important in face recognition. In the Tanaka et al. experiment subjects were required to respond to a query on every trial indicating if the preceding face had been the target. Critically, the target face only occurred on 17% of the trials. Thus, this experiment is another example of an oddball paradigm, which in this case produces a lateral-occipital N2 to targets because the targets are face stimuli that highly engage face-processing areas. In contrast, the standard repetition priming N250 may not be related to the N2, as it does not demonstrate a straightforward negative-positive deflection in succession (Schweinberger et al., 1995; 2002).

N2pc. In visual search tasks subjects usually are presented with an array of stimuli and must identify the presence of a distinguishable target. A well-researched "pop-out" effect occurs when all the distracter stimuli are the same and the target stands out by virtue of a color, orientation, or size difference (e.g. Luck, & Hillyard, 1994), such that the presence of the target registers with the subject almost immediately, and reaction time is minimally impacted by increasing the number of distracters in the stimulus array (e.g. Trick, & Enns, 1998). When subjects are required to fixate on a central point and determine whether any presented stimulus array contains a target, an enhanced P2, N2, and P3 appear to target-present displays (e.g. Luck, & Hillyard, 1994). Additionally, the enhanced N2 demonstrates a hemispheric difference such that it is largest contralateral to the target, with the difference peaking between 200 ms and 300 ms after display onset (Eimer, & Mazza, 2005;

Heinze, Luck, Mangun & Hillyard, 1990; Luck, & Hillyard, 1994; Suwanzono, Machado, &Knight, 2000; Woods, Alho, & Algazi, 1992). This component has been dubbed the N2pc, because the hemispheric difference is maximal over posterior sites (and contralateral to the target). The modified LC-P3 theory would predict such a finding, given that one hemisphere should be working harder than the other when a target is off center, thus the impact of NE should be different across hemispheres. However the N2pc is not followed by a P3, which seems problematic at first until one takes into account that the N2pc is observed by creating a difference wave of the ERP from each hemisphere. Since the P3 is not lateralized (because NE depletion has less of a spatially specific impact on the EEG, and later stages of processing are less lateralized than earlier stages), it disappears in the difference wave. Of course, in the raw waveforms, the P3 appears to the target in each hemisphere (Luck & Hillyard, 1994).

The N400. The N400 is a negative deflection in the ERP that peaks around 400 ms post stimulus onset, is maximal at central-parietal channels, and that is elicited by violations of semantic expectancy. Kutas and Hillyard (1980) first demonstrated this effect. They presented subjects with sentences, one word at time, that could either end with a word that was expected due to the context of the sentence (e.g., 'he spread the warm bread with butter') or unexpected (e.g., he spread the warm bread with socks."). When the final word was unexpected the word elicited an N400.

In a 1991 review directed at the relationship between the N2 and the N400, Pritchard, Shappell and Brandt wrote:

> "On the component level, I believe that the evidence extant in the literature points to the conclusion that N200 [N2] and N400 are best considered to be different components. However […] I believe that N200 and N400 should be linked on a level that is higher in the taxonomy than the component level […]. I call conceptually identical phenomena at this level a *supracomponent*. On this level, then, I assert that the N200 and N400 are the same supracomponent."

Pritchard et al. characterize the N2/N400 supracomponent as a violation of expectancy, with the N2 occurring to physical violations and the N400 occurring to conceptual violations. Under the modified LC-P3 theory, both should elicit an LC response.

Examination of the figures from Kutas and Hillyard (1980) revealed that the N400 was followed by a late positivity that looked very much like a late, enhanced P3 to the unexpected word. This pattern was also seen in a replication of this work done by Polich, Vanasse, and Donchin in 1981, and again by Polich in 1985, who then claimed that the N400 was in fact a generic N2 with its typical compliment, the P3. However, the literature contains some examples of a N400 without a subsequent positive enhancement (Brown, Hagoort, & Chwilla, 2000; Franklin, Dien, Neely, Huber, & Waterson, 2007; Holcomb, 1988). Generally, paradigms that employ full sentences tend to elicit N400s that are followed by complimentary positive deflections (e.g., Dien, Michelson, & Franklin, 2010; Kutas & Hillyard, 1980; Polich, 1985; Polich et al., 1981), whereas paradigms that employ only related or unrelated word pairs often produce smaller N400s that do not exhibit a subsequent positivity (e.g., Brown et al., 2000; Franklin et al., 2007; Holcomb, 1988). Recently, Dien and colleagues (2010) presented evidence that the difference-wave N400 is composed of two distinct components, a negativity (N400) associated with unexpected words, and a positivity (P400) sensitive to semantically congruent words. I interpret this result as support for the view that unexpected words engage the LC-NE system, amplifying the N400 and subsequent positivity, whereas fluent semantic processing in the absence of a NE phasic release produces the P400. Consistent with this interpretation, the appearance of a difference-wave N400 to related versus unrelated word pairs without a subsequent positivity suggest the word-pair paradigm does not generate sufficient expectancy to elicit a NE release when unrelated words are presented, eliciting only the P400 effect that drives the difference-wave N400. Whereas when full sentences are used, the unrelated word

strongly violates expectations and elicits a LC-NE system response, generating a N2-P3 complex to unexpected words and a P400 to fluently-processed words.

Specific Dissociations of the N2 from the P3

The modified LC-P3 theory holds that any manipulation that elicits a phasic burst of NE release should produce both a N2 and a P3 in the ERP, because they each reflect different stages in LC phasic activity (the flood of NE and depletion of NE, respectively). As such, the appearance of a N2 without a subsequent P3, or a P3 without a preceding N2 would cast doubt on the modified LC-P3 theory. However, the modified LC-P3 theory constrains how the N2-P3 complex should be revealed.

First, the N2-P3 should be most apparent in a difference wave ERP, where the impact of noradrenergic can be isolated; i.e. the ERP elicited by an infrequent event may not reveal a clear N2-P3 complex on its own, but when the ERP to a frequent stimulus is subtracted from the ERP to the infrequent stimulus, a N2-P3 complex is revealed.

Second, in some cases, the construction of a difference wave may remove the contribution of the P3 to the ERP if the subtraction is done across channels within the same experimental condition, such as the hemispheric difference discussed in relation to the N2pc above. In this case, the method would artificially suggest a N2 in the absence of a P3.

Third, in some circumstances, one of either the N2 or the P3 may be obscured by overlap with another component, giving the appearance of a N2 without a subsequent P3, or vice versa. Such is the case I have demonstrated, when the reward positivity attenuates the N2 but does not similarly attenuate the P3.

Finally, the N2 and P3 do not have to be present at the same electrode channel to implicate noradrenergic modulation. As I have demonstrated, the N2 has a variable scalp topography and thus may appear at any channel, whereas the P3 is more stable, typically maximal at central-parietal

channels and sometimes more frontal. Hence, the absence of a N2 together with a P3 may be more apparent that real, if the N2 is appearing at a different area of the scalp. Further, amplitude changes in the N2 at a specific electrode without a similar change in the P3 at the same or another *specific* electrode site do not constitute a dissociation of the two because change in the amplitude of the N2 can be variable across the scalp. To illustrate, Folsetin and van Petten (2007) have argued that the N2 and P3 are dissociable in two important ways. Citing Courchesne, Hillyard, and Galambos (1975), Folstein and van Petten claim that making complex, novel stimuli 'targets' has no effect on the anterior N2, but increases the parietal P3. Furthermore, citing an examination of the figures from the work of Comerchero and Polich (1998), they claim "the anterior N2 was more sensitive to the degree of perceptual deviation from the other stimuli and less sensitive to task difficulty than the P3a [frontal, novelty-elicited P3]" (p.4). These putative dissociations are not a problem for the modified LC-P3 theory because they only represent dissociations if the processes giving rise to an N2 and P3 are tied to a specific area, rather than to a global modulation of activity that can be observed anywhere in the neocortex.

Future directions

Time frequency. When ERPs are analyzed using time-frequency analysis, decomposing the waveforms into a set of frequencies of varying power, reward trials are associated with an increase in the lower gamma band (21 Hz to 29 Hz) (Bernat, Nelson, Holroyd, Gehring, & Patrick, 2008; Cohen et al., 2007), whereas error trials produce an increase in the theta band (4 Hz to 9 Hz) (e.g. Cohen et al., 2007; Haji Hosseini, & Holroyd, in press). This raises questions regarding how theta activity is related to the N2-P3 complex. It is very interesting to note that the pattern of P2-N2-P3 changes observed in experiments 3 to 5 of this work looks very much like a manifestation of enhanced theta. If this were true, Nieuwenhuis's position that the P2 is a part of the P3 and reflects the beginning of NE-modulation, and my position that the N2 reflects NE modulation would both

be accurate. As previously noted, the timing issues I raise with regard to the LC-P3 theory, that the P3 begins too late to be an index of phasic NE release, would be ameliorated if the onset of the P2 was taken as the beginning of noradrenergic modulation. The P2 peaks between 180 ms and 200 ms in the experiments reported here, and this timing maps very well onto evidence that NE should arrive in the cortex within about 170 ms of onset of the eliciting target. However, the duration of NE modulation would still be inconsistent with the P2-N2-P3 pattern of activity lasting approximately 300 ms, whereas Usher et al. assert NE modulation should only last approximately 180 ms, a time range consistent with the timing of effective target processing preceding the attentional blink.

Of note is that a cursory examination of theta activity in experiments 2 to 5 in this work revealed that the change in theta activity between frequent and infrequent trials did not shift from having a frontal-central scalp distribution to a lateral occipital scalp distribution when subjects attend to sex versus color. More investigation is needed to determine how closely theta is related to the dN2, but it is plausible that NE enhances power in the time-frequency spectrum, and that an enhancement of theta drives the frontal-central dN2, whereas an enhancement of a different sort drives the face-processing dN2.

LC in learning. The LC-NE system may modulate the speed at which new associations are learned. In monkeys, working memory performance is enhanced by direct infusion of NE agonists into the prefrontal cortex, while infusion of NE antagonists impairs working memory (Berridge, & Waterhouse, 2003). There is also evidence that the LC-NE system regulates task engagement (Usher et al., 1999)]. These effects may have important consequences for learning. Further, research suggests that NE action in the hippocampus (a neural structure critical for the formation of long-term memories) promotes synaptic plasticity by facilitating long-term potentiation (Sara, 2009), a cellular mechanism for Hebbian learning (Hebb, 1949). For example, LC neurons respond to both

rewarding and punishing events; when hippocampal neurons in a rat are subjected to high-frequency stimulation followed by exposure of the animal to either reward or punishment, long term potentiation is enhanced, but this effect is blocked by a NE antagonist (Berridge, & Waterhouse, 2003).

There is scarce work examining the impact of NE on human memory, however a natural fall-out from adaptive gain theory is that in probabilistic learning tasks, higher levels of NE and greater, relevant neural activity at the time of feedback may promote larger adjustments in current probabilistic beliefs, speeding learning (see also Verguts, & Notebaert, 2008;2009). In probabilistic learning tasks, subjects make a decision on each trial and then are rewarded according to a set of probabilities associated with their choice (i.e. a right-button press in response to a red stimulus will be rewarded 80% of the time). Subjects must learn these underlying probabilities in order to maximize reward. Computational models designed to simulate human performance in these tasks often use the delta rule (Rescorla & Wagner, 1972), in which a belief about underlying probabilities is updated by the difference between an expected (according to former beliefs) and actual outcome, modified by a fractional learning rate that limits the degree to which any single event will modify beliefs. While the learning rate is typically modelled as a constant, recent research shows that human subjects elevate their learning rate when unexpected outcomes suggest underlying probabilities have changed (Behrens, Woolrich, Walton, & Rushworth, 2007). Although the delta rule clearly maps well onto the TD theory of the DA system, the properties of the LC-NE system make it an excellent candidate for implementing changes in the learning rate. It follows from the modified LC-P3 theory that the amplitude of the N2 on a given trial should index the degree to which the outcome of that trial influences future decisions.

The attentional blink. Nieuwenhuis et al.'s (2005b) proposed a detailed account of how the dynamics of the LC-NE system produce the attentional blink (LC-AB theory), coupled with a

computational model instantiating the theory. When two targets are embedded within a rapid serial visual presentation, the second target suffers a decrease in reporting accuracy relative to the first. The LC-AB theory holds that this deficit for the second target is due to the refractory-like period in LC phasic activity, such that the first target benefits from a phasic burst of NE, but the second target does not. The LC-AB computational model successfully demonstrated key aspects of the attentional blink, such as the finding that the second target does not suffer the attentional blink deficit if it is presented within 100 ms of onset of the first target, and further, that this "lag-1 sparing" is time-dependent, rather than lag-dependent (the lag-1 target does not avoid the blink by virtue of being the very next character presented after the first, but rather by being presented so close in time to the first).

In previous behavioural work, I supported Nieuwenhuis et al.'s LC-AB theory by showing that lag-1 sparing can be extended in time by putatively provoking a larger LC response that would result in an extended period of NE availability in the cortex (Warren et al., 2009). I demonstrated sparing for a second target occurring 200 ms after onset of the first, with two intervening distracters, by moving the first target and the distracter following the first target closer together in time to increase interference between the two. I reasoned that this interference would be synonymous with Botvinick et al.'s (2001) neural conflict, and would encourage greater recruitment of the LC phasic response.

Interestingly, the LC-AB model was an adaptation of an earlier model (Gilzenrat et al., 2002) that was developed to investigate how changes in electronic coupling moved the LC-NE system between high-tonic/low-phasic and low-tonic/high-phasic modes as previously investigated by Usher et al. (1999), and it included a conflict variable that was not used in the LC-AB model. This conflict variable indexed the sum of the products of activity in each neuron and their connection weight across all connected pairings, in the same way that Botvinick et al. modelled neural conflict.

My unpublished simulations with the LC-AB model show increased conflict following onset of the target stimuli in the AB task, as response units related to both targets and distracters are activated simultaneously.

The LC-AB model simulated motivational salience to elicit LC firing (target present versus target absent), but the model could also simulate conflict detection to elicit LC firing. Adapting the LC-AB model to simulate the findings from my previous study (Warren et al., 2009) and presenting it with a replication of that experiment while simultaneously recording ERP data could potentially illuminate how motivational salience and conflict detection interact to elicit phasic activity in the LC, and to produce the dN2.

Conclusion

I have proposed a new idea of the neural genesis of the N2. My idea borrows heavily from the original LC-P3 theory, expanding and modifying it to include the N2 in its scope. My modification to the LC-P3 theory retains the major insights and strong empirical support for the original theory while simultaneously resolving some apparent discrepancies. An important implication of the modified LC-P3 theory is that relative levels of cortical NE should impact a variety of ERP components. My results highlight the fact that differences in the scalp distribution of a change between experimental conditions do not necessarily indicate that the primary neural process driving the change is specific to either cortical area. Thus, our understanding of the neurocognitive processes underlying many ERP components may be informed by investigating their sensitivity to catecholinergic neuromodulation.

The anatomy of the LC-NE system suggests a pervasive influence on cognitive processes. The DA system clearly plays a major role in decision making and how we learn about the consequences of our actions. Together these two neuromodulatory systems have been associated

with a large range of psychiatric disorders and diseases. This highlights the potential value in developing valid indexes of their ongoing activity in human subjects.

References

Aghajanian, G. K., Cedarbaum, J. M., & Wang, R. Y. (1977). Evidence for norepinephrine-mediated collateral inhibition of locus coeruleus neurons. *Brain Research, 136,* 570 –577.

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

Aston-Jones, G., Foote, S. L., & Bloom, F. E. (1984). Anatomy and physiology of locus coeruleus neurons: Functional implications. In M. Ziegler & C. R. Lake (Eds.), *Norepinephrine: Frontiers of clinical neuroscience* (Vol. 2, pp. 92–116). Baltimore: Williams & Wilkins.

Aston-Jones, G., Foote, S. L., & Segal, M. (1985). Impulse conduction properties of noradrenergic locus coeruleus axons projecting to monkey cerebrocortex. *Neuroscience, 15*(3), 765-777.

Aston-Jones, G., Rajkowski, J., Kubiak, P., & Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *The Journal of Neuroscience, 14*(7), 4467-4480.

Baker, T. E., & Holroyd, C. B. (2009). Which way do I go? Neural activation in response to feedback and spatial processing in a virtual T-maze. *Cerebral Cortex, 19*, 1708-1722.

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology, 87*, 25-34.

Baker, T. E., Stockwell, T., Barnes, G., Holroyd, C. B. (2011). Individual differences in substance dependence: At the intersection of brain, behavior, and cognition. *Addiction Biology, 16*, 458-466.

Behrens, T. E., Woolrich, M. W., Walton, M. E., Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), *1214-1221.*

Bentivoglio, M., Kultas-Ilinsky, K., & Ilinsky, I. (1993). Limbic thalamus:Structure, intrinsic organization, and connections. In B. A. Vogt & M.Gabriel (Eds.), *Neurobiology of cingulate cortex and limbic thalamus: Acomprehensive handbook* (pp. 71–122). Boston: Birkhauser.

Bernat, E. M., Nelson, L. D., Holroyd, C. B., Gehring, W. J., & Patrick, C. J. (2008). Separating cognitive processes with principal components analysis of EEG time-frequency distributions. *Proceedings of the Society of Photo-Optical Instrumentation Engineers, 7074*, 70740S.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28,* 309–369.

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews, 42*(1), 33-84.

Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*(6758), 179.

Bouret, S. & Sara, S. J. (2005). Network reset: A simplified overarching theory of locus coeruleus noreadrenaline function. *Trends in Neurosciences, 28*(11), 574-582.

Braver, T. S. *,* Barch, D. M. *,* Gray, J. R. *,* Molfese, D. L.*,* Snyder, A. *(*2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors . *Cerebral Cortex 11(9):* 825-836.

Briand, L. A., Gritton, H., Howe, W. M., Young, D. A. & Sarter, M. (2007). Modulators in concert for cognition: Modulator interactions in the prefrontal cortex. *Progress in Neurobiology, 83(2*), 69-91.

Brown, C. M., Hagoort, P., & Chwilla, D. J. (2000). An event-related brain potential analysis of visual word priming effects. *Brain and Language, 72*, 158-190.

Butterfield, B., & Mangels, J. A. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Cognitive Brain Research, 17*, 793-817.

Calabresi, P., Pisani, A., Mercuri, N. B., & Bernardi, G. (1996). Thecorticostriatal projection: From synaptic plasticity to dysfunctions of the basal ganglia. *Trends in Neurosciences, 19,* 19–24.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749.

Clark, V. P., Fannon, S., Lai, S., Benson, R., Bauer, L. (2000). Responses to rare visual target and distractor stimuli using event-related fMRI. *Journal of Neurophysiology*, *83*(5), 3133-3139.

Clayton, E. C., Rajkowski, J., Cohen, J. D., & Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. *The Journal of Neuroscience, 24*(44), 9914-9920.

Cohen, M. X, & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience, 27*(2), 371-378.

Cohen, M. X, Elger, C. E., Ranganath, C. (2007). Reward Expectation Modulates Feedback-Related Negativity and EEG Spectra. *NeuroImage, 35*(2), 968-78.

Comerchero, M. D., & Polich, J. (1998). P3a, perceptual distinctiveness, and stimulus modality. *Cognitive Brain Research, 7*, 41–48.

Coulson, S., King, J. W., & Kutas, M. (1998), Expect the unexpected: Event-related braind response to morphosyntactic violations. *Language and Cognitive Processes, 13(1),*. 21-58.

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology, 39*, 131-143.

Crottaz-Herbette, S., & Menon, V. (2006). Where and when the anterior cingulate cortex modulates attentional response: Combined fMRI and ERP evidence. *Journal of Cognitive Neuroscience, 18*(5), 766-780.

Daffner, K. R., Mesulam, M. M., Scinto, L. F. M., Calvo, V., Faust, R., & Holcomb, P. J. (2000). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology, 37(6)*, 737-747.

Daffner, K. R., Scinto, L. F. M., Calvo, V., Faust, R., Mesulam, M. M., West, W. C., & Holcomb, P. J. (2000). The influence of stimulus deviance on electrophysiologic and behavioral responses to novel events. *Journal of Cognitive Neuroscience, 12(3)*, 393-406.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon D.Y. & Engel, A.K. (2005). Trial-by-trial coupling of concurrent EEG and fMRI identifies the dynamics of performance monitoring. *Journal of Neuroscience, 25,* 11730-37.

Dien, J., Michelson, C. A., & Franklin, M. S. (2001). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: cognitive and neuroanatomical implications. *Brain Research, 1355*, 126-140.

Donchin, E. (1968). Average evoked potentials and uncertainty resolution. *Psychonomic Science*, *12(3)*. 103.

Donchin, E. & Coles, M. G. H. (1988).). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(3), 357-427.

Donchin, E., & Heffley, E. (1978). Multivariate analysis of event-related potential data: A tutorial review. In Otto, D. A. (Ed.). *Mulltidisciplinary Perspectives in Event-Related Brain Potental Research*. US Government Printing Office: Washington, D.C., 555-572.

Donkers, F.C.L., & van Boxtel, G.J.M. (2005). Mediofrontal negativities to averted gains and losses in the slot-machine task: a further investigation. *Journal of Psychophysiology, 19*(4), 256-262.

Duncan-Johnson, C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456-467.

Duncan-Johnson, C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology*, *14*(1-2), 1-52.

Dunning, J.P., & Hajcak, G. (2007). Error-related negativities elicited by monetary loss and cues that predict loss. *NeuroReport, 18,* 1875-1878.

Durstewitz, D. & Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biological Psychiatry, 64*, 739-749.

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

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149.

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

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299(5614)*, 1898-1902.

Fischer, C., Luaute, J., & Morlet, D. (2010). Event-related potentials (MMN and novelty P3) in permanent vegetative or minimally conscious states. *Clinical Neurophysiology, 127*(7), 1032-1042.

Folstein, J. R., & Van Petten, C. (2007). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45(1)*, 152-170.

Folstein, J. R., Van Petten, C., & Rose, S. A. (2008). Novelty and conflict in the categorization of complex stimuli. *Psychophysiology, 45*, 467-479.

Foote, S. L. & Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annual Review of Neuroscience, 10*, 67-95.

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C. & Fawcett, J. (2010). Antidepressant drug effects and depression severity. *The Journal of the American Medical Association, 303(1)*, 47-53.

Ford, J. M., Roth, W. T., & Kopell, B. S. (1976). Auditory evoked potentials to unpredictable shifts in pitch. *Psychophysiology, 13*, 32-39.

Franklin, M. S., Dien, J., Neely, J. H., Huber, E., & Waterson, L. D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology*, 118(5), 1053-1068.

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

Gehring, W. J., Gratton, G., Coles, M. G., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 198-216.

Gilzenrat, M., Holmes, B., Rajkowski, J., Aston-Jones, G., & Cohen, J. (2002). Simplified dynamics in a model of noradrenergic modulation of cognitive performance. *Neural Networks*, *15(4-6)*, 647-663.

Gratton, G., Coles, M., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55(4)*, 468-484.

Halliday, R., Naylor, H., Brandeis, D., Callaway, E., Yano, L., & Herzig, K. (1994). The effect of D-amphetamine, clonidine, and yohimbine on human information processing. *Psychophysiology, 31(4),* 331–337.

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

Haji Hosseini, A., & Holroyd, C. B., Human frontocentral theta oscillations are sensitive to improbable task-relevant events, not just errors. *11th International Conference on Cognitive Neuroscience (ICON XI)*, September 25-29, 2011 Mallorca, Spain.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley & Sons.

Hewig, J., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2007). Decision making in blackjack: An electrophysiological analysis. *Cerebral Cortex, 17*, 865-877.

Hewig, J., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2008). An electrophysiological analysis of coaching in blackjack. *Cortex, 44*, 1197-1205.

Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, *172(3990)*, 1357-1360.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*, 66-85.

Holroyd, C. B. (2001). *Reinforcement learning and the error-related negativity: A computational and neurophysiological investigation.* Doctoral dissertation, University of Illinois at Urbana–Champaign.

Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN, In M. Ullsperger & M. Falkenstein (Eds.). *Errors, Conflicts, and the Brain: Current Opinions on Performance Monitoring,* 211-218. Leipzig: MPI of Cognitive Neuroscience.

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.

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

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

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

Holroyd, C. B. Nieuwenhuis, S., Mars, R., & Coles, M. G. H. (2004). Anterior cingulate cortex, selection for action, and error processing. In M. Posner (Ed.). *Cognitive Neuroscience of Attention*, 219-231. New York: Guilford Publishing, Inc.

Holroyd, D. B., Nieuwenhuis, S. Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport, 14*, 2481-2484.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688-697.

Holroyd, C. B., & Yeung, N. (in press). An integrative theory of anterior cingulate cortex function: Option selection in hierarchical reinforcement learning. In R. Mars, J. Sallet, M. Rushworth, & N. Yeung (Eds.), *Neural Basis of Motivational and Cognitive Control.* Cambridge, MA: MIT Press.

Holroyd, C. B., Yeung, N., Coles, M. G. H., & Cohen, J. D. (2005). A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General, 154*(2) 163-191.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, *79*, 2554–2558.

Johnson, R. Jr. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology, 30*, 90-97.

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

Kiehl, K.A., Laurens, K.R., Duty, T.L., Forster, B.B., & Liddle, P.F. (2001). Neural sources involved in auditory target detection and novelty processing: An event-related fMRI study. *Psychophysiology, 38*, 133-142

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*, 557-577.

Krigolson, O. E., & Holroyd, C. B. (2007). Predictive information and error processing: The role of the medial-frontal cortex during motor control. *Psychophysiology, 44*, 586-595.

Krigolson, O. E., Holroyd, C. B., Van Gyn, G., & Heath, M. (2008). Electroencephalic correlates of target and outcome errors. *Experimental Brain Research, 190*, 401-411.

Krigolsoon, O. E., Pierce, L. J., Holroyd, C. B., & Tanaka, J. W. (2009). Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. *Journal of Cognitive Neuroscience, 21*, 1834-1841.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207* (4427), 203-205.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*, 291-308.

Luck, S.J. (2005). An Introduction to the Event-Related Potential Technique. Cambridge, MA: The MIT Press.

McArthur, G., Budd, T., & Michie, P. (1999). The attentional blink and P300. *NeuroReport*, 10(17), 3691-3695.

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.

Miltner, W. H. R., Lemke, U., Weiss, T., Holroyd, C. B., Scheffers, M. K. & Coles, M. G. H. (2003). Implementation of error-processing in the human anterior cingulate cortex: A source analysis of the magnetic equivalent of the error-related negativity. *Biological Psychology, 64*, 157-166.

Morecraft, R. J., & Van Hoesen, G. W. (1991). A comparison of frontal lobe inputs to the primary, supplementary, and cingulate motor areas in the monkey. *Society for Neuroscience Abstracts, 17,* 1019.

Morris, S. E., Heerey, E. A., Gold, J. M., Holroyd, C. B. (2008). Learning-related changes in brain activity following errors and performance feedback in schizophrnia, *Schizophrenia Research, 99,* 274-285.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity(MMN) and its magnetic equivalent (MMNm). Psychophysiology, 38, 1–21..

Näätänen, R., Gaillard, A. W. K., & Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologia, 42*, 313-329.

Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*, 2544-2590.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005a). Decision making, the P3, and the locus coeruleus--norepinephrine system. *Psychological Bulletin, 131*(4), 510-532.

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J. D. (2005b). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational theory. *Journal of Experimental Psychology: General, 134*, 291–307.

Nieuwenhuis, S., de Geus, E.J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology, 48,* 162-175.

Nieuwenhuis, S., & Jepma, M. (2010). Investigating the role of the noradrenergic system in human cognition. In T. Robbins, M. Delgado, & E. Phelps (Eds.), *Decision making, Attention & Performance*, *Vol. XXIII*. Oxford: Oxford University Press.

Nieuwenhuis, S., Nielen, M., Mol, N., Hajcak, G., & Veltman, D. (2005). Performance monitoring in obsessive-compulsive disorder. *Psychiatry Research, 134(2)*, 111-121.

Nieuwenhuis, S., Ridderinkhof, K.R., Talsma, D., Coles, M.G.H., Holroyd, C.B., Kok, A., & Van der Molen (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective & Behavioral Neuroscience, 2,* 19-36.

Nieuwenhuis, S., van Nieuwpoort, I. C., Veltman, D. J., Drent, M. L. (2007). Effects of the noradrenergic agonist clonidine on temporal and spatial attention. *Psychopharmacology, 193*(2):261-269.

Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2004). Stimulus modality, perceptual overlap, and the go/no-go N2. *Psychophysiology, 41*(1), 157-160.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience, 3*(1), 17-26.

Peterson, D.A.; Lotz, D.T.; Halgren, E.; Sejnowski, T. J.; Poizner, H; Choice Modulates the Neural Dynamics of Prediction Error Processing During Rewarded Learning, *NeuroImage, 54,* 1385-1394,

Pineda, J. A., Foote, S. L., & Neville, H. J. (1989). Effects of locus coeruleus lesions on auditory, long-latency, event-related potentials in monkey. *Journal of Neuroscience, 9,* 81–93.

Pineda, J. A., & Westerfield, M. (1993). Monkey P3 in an "oddball" paradigm: Pharmacological support for multiple neural sources. *Brain Research Bulletin, 31,* 689–696.

Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of N200/N400: A review and classification scheme. *Advances in Psychophysiology, 4*, 43-106.

Rajkowski , J., Majczynski, H., Clayton, E., & Aston-Jones, G. (2004). Activation of monkey locus coeruleus neurons varies with difficulty and behavioral performance in a target detection task. *Journal of Neurophysiology, 92*, 361-371.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 18(3)*, 849-860.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning:Variations in the effectiveness of reinforcement and nonreinforcement.In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioningII: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Ritter, W., Simson, R., Vaughan, H. G., & Friedman, D. (1979). A brain event related to the making of a sensory discrimination. *Science, 203(4387)*, 1358-1361.

Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1978). Cortical slow waves following non-paired stimuli: Effects of task factors. *Electroencephalography and Clinical Neurophysiology, 45*, 551-567.

Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1979). Cortical slow negative waves following no-paired stimuli: Effects of modality, intensity, and rate of stimulation. *Electroencephalography and Clinical Neurophysiology, 46*, 416-427.

Romo, R. & Schultz, W. (1990). Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology, 63*, 592-606.

Rozenkrants, B., & Polich, J. (2008). Affective ERP processing in a visual oddball task: Arousal, valence, and gender. *Clinical Neurophysiology*, *119*(10), 2260-2265.

Rushworth, M. F. S., Behrens, T. E. J., Rudebeck, P. H. & Walton, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences, 11(4)*, 168-176.

Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology, 17*, 41-58.

Sams, M., Alho, K., & Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology, 21*(4), 434-441.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience, 10*(3), 211-223.

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

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*, 241-263.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successivesteps of learning a delayed response task. *Journal of Neuroscience, 13,*900–913.

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

Schultz, W. & Romo, R. (1987). Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *Journal of Neurophysiology, 57*, 201-217.

Schweinberger, S. R., Pfütze, E., & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 722-736.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. (2002). Event-related potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*(3), 398-409.

Servan-Schreiber, D., Printz, H., & Cohen, J. D. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science, 249* (4971), 892-895.

Simson, R., Vaughan, H. G., & Ritter, W. (1976). The scalp topography of potentials associated with missing visual or auditory stimuli. *Electroencephalography and Clinical Neurophysiology, 40*, 33-42.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497–537). Cambridge, MA: MIT Press.

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

Squires, K. C., Donchin, E., Herning, R. I., & McCarthy, G. (1977). On the influence of task relevance and stimulus probability on event-related potential components. *Electroencephalography and Clinical Neurophysiology, 42*, 1-14.

Squires, N. K., Squires, K. C., Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology 38* (4), 387–401.

Squires, N. K., Wickens, C., Squires, K. C., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of teh cortical event-related potential. *Science, 193(4258)*, 1142-1146.

Swick, D., Pineda, J. A., Schacher, S., & Foote, S. L. (1994). Locus coeruleus neuronal activity in awake monkeys: Relationship to auditory P300-like potentials and spontaneous EEG. *Experimental Brain Research, 101,* 86–92.

Swick, D., Pineda, J. A., & Foote, S. L. (1994). Effects of systemic clonidine on auditory event-related potentials. In Kertesz, A. (Ed.), *Localization and Neuroimaging in Neuropsychology*,73-121.

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: The N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience, 18*(9), 1488-1497.

Trick, L. M., & Enns, J. T. (1997). Measuring preattentive processes: When is pop-out not enough?. *Visual Cognition*, 4(2), 163-198.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage, 14,* 1387–1401.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science, 283*(5401), 549-554.

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *NeuroImage, 14*(6), 1302-1308.

van Veen, V., & Carter, C. S. (2002a). The anterior cingulate as a conflict monitor: FMRI and ERP studies. *Physiology & Behavior, 77*(4), 477-482.

van Veen, V., & Carter, C. S. (2002b). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14*(4), 593-602.

Verguts, T. & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review, 115*(2), 518-525.

Verguts, T. & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Sciences, 13*(6), 252-257.

Warren, C. M., Breuer, A. T., Kantner, J., Fiset, D., Blais, C., & Masson, M. J. (2009). Target–distractor interference in the attentional blink implicates the locus coeruleus–norepinephrine system. *Psychonomic Bulletin & Review*, 16(6), 1106-1111.

Warren, C. M., Tanaka, J. W., & Holroyd, C. B. (2011). What Can Topology Changes in the Oddball N2 Reveal about Underlying Processes? *NeuroReport, 22*(17): 870-874.

Waterhouse, B. D., Moises, H. C., & Woodward, D. J. (1980). Noradrenergic modulation of somatosensory cortical neuronal responses to iontophoretically applied putative neurotransmitters. *Experimental Neurology, 69(1*), 30-49.

Waterhouse, B. D., Moises, H. C., & Woodward, D. J. (1981). Alpha-receptor-mediated facilitation of somatosensory cortical neuronal responses to excitatory synaptic inputs and iontophoretically applied acetylcholine. *Neuropharmacology, 20(10)*, 907-920.

Wickens, J. R., Begg, A. J., & Arbuthnott, G. W. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normallyfollows high-frequency stimulation of cortex in vitro. *Neuroscience, 70,*1–5

Wickens, J., & Kotter, R. (1995). Cellular models of reinforcement. In J.Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing inthe basal ganglia* (pp. 187–214). Cambridge, MA: MIT Press.

Wijnen, V.J.M., Boxtel, G.J.M. van, Eilander, H.J., de Gelder, B. (2007). Mismatch Negativity predicts recovery from the vegetative state. *Clinical Neurophysiology, 118*, 597 – 605.

Williams, S. M., Goldman-Rakic, P. S. (1993). Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cerebral Cortex, 3*, 199-222.

Woodward, D. J., Moises, H. C., Hoffer, B. J. & Freedman, R. (1979a). INteractions of norepinephrine with Purkinje-cell response to putative amino-acid neurotransmitters applied by microiontophoresis. *Experimental Neurology, 64(3*), 493-515.

Woodward, D. J., Moises, H. C., Waterhouse, B. D., Hoffer, B. J., Freedman, R. (1979b). Modulatory actions of norepinephrine in the central nervous system. *Federation Proceedings, 38(7)*, 2109-2116.

Yasuda, A., Sato, A., Miyawaki, K., Kumano, H., & Kuboki, T. (2004 ） Error-related negativity reflects detection of negative reward prediction error. *NeuroReport, 15*, 2561-2565

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

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111*(4), 931-959.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology & Psychology, 18*, 459-482.

Yu, A. J. & Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. In *Advances in Neural Information Processing Systems 15*. MIT Press, Cambridge, MA.

Yu, A. J., & Dayan, P. (2005). Uncertainty, Neuromodulation, and attention. *Neuron, 46*, 681-692.

Dayan, P. & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems, 17*: 335-50.

Appendix

I conducted a linear regression analysis on dFRN and dN2 amplitude, as predicted by learning engagement across the Passive, Moderate and Active experiments. I assigned each experiment an ordinal value according to the putative relative engagement of learning (Passive, 1; Moderate, 2; Active, 3). Learning engagement significantly predicted dFRN amplitude at FCz, r^2=.12, B = -1.22, p <.01. Learning engagement also showed a trend toward predicting dN2 amplitude, r^2=.05, B = .65, p <.10.