4-2022

THE MENTAL NOISE HYPOTHESIS: A RELATION BETWEEN NEUROTICISM AND P3 LATENCY VARIANCE IN A STROOP-STYLE REACTION TIME TASK

Jeremy Lawrence
jelawren@nmu.edu

THE MENTAL NOISE HYPOTHESIS: A RELATION BETWEEN NEUROTICISM AND P3 LATENCY VARIANCE IN A STROOP-STYLE REACTION TIME TASK

By

Jeremy Michael Lawrence

THESIS

Submitted to
Northern Michigan University
In partial fulfillment of the requirements
For the degree of

MASTER OF SCIENCE

Office of Graduate Education and Research

March 2022

SIGNATURE APPROVAL FORM

THE MENTAL NOISE HYPOTHESIS: A RELATION BETWEEN NEUROTICISM AND P3 LATENCY VARIANCE IN A STROOP-STYLE REACTION TIME TASK

This thesis by Jeremy Michael Lawrence is recommended for approval by the student's Thesis Committee and Department Head in the Department of Psychological Science and by the Dean of Graduate Studies and Research.

Jon Barch — 04/05/2022

Committee Chair: Dr. Jon Barch — Date

04/10/2022

First Reader: Dr. Lin Fang — Date

Christina F. Hartline — 04/11/2022

Second Reader: Dr. Christy Hartline — Date

04/11/2022

Department Head: Dr. Adam Prus — Date

Lisa Schade Eckert — 5/3/22

Dr. Lisa Schade Eckert — Date
Dean of Graduate Studies and Research

ABSTRACT

THE MENTAL NOISE HYPOTHESIS: A RELATION BETWEEN NEUROTICISM AND P3 LATENCY VARIANCE IN A STROOP-STYLE REACTION TIME TASK

By

Jeremy Michael Lawrence

Neuroticism is a relatively stable personality dimension characterized by tendencies to experience negative thoughts and affect. Its empirically related outcome measures range from anxiety and mood disorders to increases in mortality. Traditional theories of neuroticism, link the construct to greater threat sensitivity, however, these conceptions fail to account for certain salient features of neuroticism, such as negative affect in threat benign environments. The mental noise hypothesis posits that neuroticism results from a more variable mental control system, with support coming from behavioral, psychometric, and neuroimaging paradigms. To assess whether this more chaotic mental control system would variably disrupt the stimulus evaluation phase of cognition, single trial latency variance of the P3b event related potential was assessed in individuals who scored high and low on neuroticism inventories. The primary analysis in this paper failed to provide support for P3b latency variance as the neural generator of the more variable mental control system previously observed in those high in neuroticism. Exploratory analyses with electrode Pz, however, found interesting relationships between lower order neuroticism aspects and prolonged P3b latencies, as well as higher levels of interparticipant P3b latency variance in high neuroticism groups.

DEDICATION

This thesis is dedicated to my family, David Lawrence, Caroline Cheng, and Edwin Lawrence, and my loving partner Emma Irwin. Thank you for making life worth living.

ACKNOWLEDGEMENTS

The author sincerely thanks his thesis chair, Dr. Jon Barch, for his counsel throughout the project, Dr. Lin Fang, who provided important training in statistics and methodological input, and Dr. Christy Hartline for her insight into the problems plaguing self-report data. This project would not have been completed without my friends and classmates, Cole Holt, Siraj Lyons, and Caleb Coughtry-Carpenter who provided a well needed break between hours of work and the tireless work of the research assistants who helped collect data for this project. The author also wishes to acknowledge Rachel Miller for her countless hours over the past year as a research assistant and Elijah Niemann for his help in setting up EEG data acquisition.

This thesis follows the format prescribed by the *Publication Manual of the American Psychological Association* (7th ed.) and the Department of Psychological Science.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Neuroticism (N) is a personality dimension that captures relatively stable tendencies to experience negative thoughts and affect. This dimension is included in almost all major personality inventories, such as the linguistically derived Big Five, Eysenck's 3-factor, and Tellegen's Multidimensional Personality model (Digman, 1990). Neuroticism has also been found to be significantly correlated with the negative emotionality and psychoticism scales of the MMPI-2 PSY5 in the general population as well as in clinical samples (Trull et al., 1995). The convergence of these scales provides validity for the neuroticism construct and supports the notion that personality disorders can be conceptualized as extreme ends of underlying dimensions. Negative behavioral and physiological correlates related to and mediated by N include substance abuse, gastrointestinal, eating, and cardiac disorders, which all contribute to an increase in all-cause mortality (Malouff et al., 2005). Despite the prevalence of anxiety and mood disorders related to high levels of N (i.e., depression, generalized anxiety, panic disorders), and severity of schizophrenia spectrum and other psychotic disorders related to high levels of N (i.e., borderline, avoidant, and schizotypal disorders), little is known about the functional regions underlying N's salient aspects (Allen & DeYoung, 2017).

Neuroticism and Major Depression

N's association with a diagnosis of major depression (MD) is demonstrated in a longitudinal population based twin study conducted by Kendler et al. (2002). From 1972 to 1973 a self-questionnaire was sent to twin pairs (*M* [age] = 29.2) living in Sweden assessing their level of N and extraversion, with a total of 20,692 responses. The assessments of MD were taken from 1998 to 2003 using a shortened version of the Composite International Diagnostic Interview.

Using logistic regression with controls for sex and birth year, high N was found to be significantly related to lifetime risk of MD compared to low N. When analyzing N's relationship with first onset of MD, high N, relative to low N, was found to significantly increase the risk for first onset of MD by 31%. Studies examining personality variables and diagnostic status concurrently have found the anxiety and self-consciousness facets of the NEO-PI N scale to significantly correlate with severity of depressive symptoms, with effect sizes in the medium range (Bagby et al., 1996). Designs utilizing concurrent assessment have also found significantly higher N in those diagnosed with MD compared to those diagnosed with obsessive-compulsive disorder (Rector et al., 2002).

Neuroticism and Schizophrenia

The ability of N to predict the onset of schizophrenia later in life is illustrated by a longitudinal study conducted by Van Os & Jones (2001), which analyzed data collected in the Medical Research Council National Survey of Health and Development, a dataset of British people born in March, 1946. Members of this survey completed a six item N inventory from the Maudsley Personality Inventory, developed by Eysenck (1958). Affective symptoms were assessed at age 36 with a shortened version of the Present State Examination (PSE), developed by Rodgers and Mann (1986), and at 43 with the Psychiatric Symptom Frequency Scale, developed by Rodgers (1996). The first stage of schizophrenia identification utilized three sources: a questionnaire and interview contacts with survey members, the Mental Health Inquiry, and the shortened version of the PSE administered at 36. The second stage involved analyzing clinical material from hospital notes, correspondence with general practitioners, and the survey member's description of symptoms for meeting the DSM-III-R's criteria for schizophrenia or schizoaffective disorder. N was then divided into three groups and a logistic regression was

performed to assess associations between schizophrenia and N. A strong positive association was found between high N, assessed at age 16, and adult diagnosis of schizophrenia, with a 1 SD increase in N resulting in a 93% increase in the odds of developing schizophrenia.

Neuroticism and Mortality

Both prospective and retrospective studies have investigated the link between personality variables and mortality. These studies have converged on results that link high levels of N to increases in mortality from all causes. In a prospective study conducted by Shipley et al. (2007), data were analyzed from the UK Health and Lifestyle Survey (HALS), which contains personality data, assessed with the Eysenck Personality Inventory, from 9003 adults ranging from age 18 to 99. Survey respondents were randomly selected in 1984 from 12,254 addresses, with one respondent chosen from each address. Following each response, an interview was conducted in the home followed by a physical administered by a nurse. In 2005, data from HALS were analyzed using Cox proportional hazards regression analysis to calculate hazard ratios for the proportion change in mortality risk for each SD increase in N. For all-cause mortality, a 1 SD increase in N was found to increase mortality risk by 9%. In participants aged 40-59, each SD increase in N was found to increase mortality risk by 12%. In all age groups N was found to be a significant predictor of deaths related to cardiovascular disease (CVD), with each SD increase in N resulting in a 12% increase in mortality from CVD, and coronary heart disease (CHD), with each SD increase in N resulting in a 14% increase in mortality from CHD.

The increased effect size for all-cause mortality related to high levels of N in elderly populations has also been found in a retrospective study conducted by Wilson et al. (2005), which used data from the Chicago Health and Aging Project. This project randomly selected 6,158 individuals over age 65 from three neighborhoods adjacent to Chicago and collected

medical history along with personality variables, using the revised 50-point NEO Five-Factor Inventory (Costa & McCrae, 1992). The predictive effect of personality variables was analyzed using a Weibull accelerated failure-time model, which computed relative risk ratios for various causes of mortality (Wilson et al., 2005). During an average of 6.2 years of observation, 39.5% of the sample passed away and a simple analysis found N to be a predictor of mortality. When controlling for age, sex, ancestry, and education, each point increase in N was found to increase risk of death by 1.6%. In the total distribution of N, those in the 90th percentile of N scores were 33% more likely to pass away during the time observed when compared to those in the 10th percentile.

Economic Impact of Major Depression

While it is difficult to quantify the exact cost of all high N related disorders, an analysis conducted by Greenberg et al. (2015) has estimated the economic cost incurred by MD and its comorbidities in 2005 and 2010. Prevalence data for MD included responses from the National Survey on Drug Use and Health, which were categorized using DSM-IV criteria for a major depressive episode (MDE). Cost data were found by selecting individuals diagnosed with major depressive disorder (MDD) from OptumHealth Reporting and Insights, a private insurance administrative claims database that includes 16 million beneficiaries. Total economic cost was comprised of direct costs (i.e., medical services and pharmaceuticals incurred on the same day and location of a medical claim from MDD), suicide-related costs, which on the basis of prior literature was attributed 50% to MDD, and workplace costs, which was assessed by calculating value lost from missed days of work and diminished productivity while working of individuals with MDD compared to matched controls. From 2005 to 2010, the prevalence of MDD rose 11.5% (from 13.8 to 15.4 million adults), with the most abrupt rise in prevalence occurring in the

50 and older age group. The total economic cost of MDD grew from an estimated $173.2 billion in 2005 to $210.5 billion in 2010, with 45-47% attributed to direct costs, 5% to suicide costs, and 48-50% to workplace costs. Individuals with MDD were also found to experience an increased range of other disease burdens relative to matched controls.

Genetic Architecture of Neuroticism

Effect sizes for N's empirically related outcome measures have been found in meta-analyses to be in the medium-large range, indicating that each individual who met the criteria for a related disorder ranged from 0.5 to 1 SD higher than the general population in their level of N (Malouff et al., 2005). These observations are functionally supported by a recent genome wide association study (GWAS) conducted by Luciano et al. (2018). As opposed to candidate gene studies, which look at the effect of a single genetic variant on a phenotype, GWASs survey the whole genome to find single nucleotide polymorphisms (SNPs) associated with the target phenotype. This method allows for the calculation of genetic correlations, with linkage disequilibrium (LD) score regression, which compares the causal SNPs of one phenotype to another to see if they share a common genetic risk factor. Using LD score regression, genetic correlations have been found between N and depressive symptoms, r_g = .82, MDD, r_g = .69, subjective wellbeing, r_g = -.68, schizophrenia, r_g = .21, and anorexia nervosa, r_g ~ .20. Pathway analysis with MAGMA, which relates significant SNPs to their cognate gene through chromatin interactions in order to reveal their gene-level associations, implicate five ontology pathways related to N: neuron-spine, homophilic cell adhesion via plasma membrane adhesion molecules, neuron differentiation, cell-cell adhesion via plasma membrane adhesion molecules, and neurogenesis. Subsequent gene-set analysis revealed that protein-expressing genes that bind to the molecules of anti-depressant medication were also significantly related to N, indicating that

N and depression share common biological pathways. This result highlights the clinical potential of assessing N after stressful life events, as those high in N can be pre-emptively selected after traumatic experiences for therapeutic intervention before the onset of a depressive episode or major depression.

Assessment of Neuroticism

Currently, the prevailing assessment for neuroticism is a series of self-report items, as a subscale in three, five, or multidimensional personality models. While this method provides some clinical utility, it suffers from issues inherent to self-report measurements utilized as a diagnostic criterion, such as social desirability effects. The issues surrounding the diagnostic efficacy of self-report measurements is illustrated by Jeong et al. (2018), who found a false negative rate of 44% in the diagnosis of internet gaming disorders (IGDs). The rate of false negative diagnoses uncovered by Jeong et al. highlights the importance of external diagnostic measures, such as distinct patterns of event related potentials (ERPs) measured with electroencephalography (EEG). Despite its lower spatial resolution, EEG has key advantages over other neuroimaging techniques, namely fMRI and PET scans, for use as a biomarker, such as its relatively low cost of operation and high temporal resolution. ERPs can also be elicited in tasks that require minimal amounts of attention, which avoids the issue of recording activity in certain clinical samples who may have difficulties with complicated tasks. Additionally, gaining a further understanding of N's neurological properties can allow for the development of novel treatments that have impact on properties of ERPs, such as L-Dopa which has been found to reduce the latency of the P300 wave in patients with Parkinson's disease (Stanzione et al., 1991).

The Mental Noise Hypothesis

Traditional theories of N explain the dimension in terms of greater threat sensitivity, but this conception lacks explanatory power for certain features of N, such as negative affect in threat-benign environments (Robinson & Tamir, 2005). The mental noise hypothesis (MNH) posits that a more "chaotic mental control system" is the mechanism underlying N, which is reflected by more intrusive thoughts and lapses of attention that occur regardless of the presence of external stressors. Converging evidence in support of the MNH comes from behavioral studies as well as neuro-imaging techniques. Behavioral studies have found that N is significantly correlated with lapses of attention and greater cognitive failures due to intrusions of task-irrelevant cognitions (Flehmig et al., 2007). Reaction time (RT) paradigms investigating the personality-performance relationship capture trends in performance variability and provide further support for the MNH (Robinson & Tamir, 2005).

Due to the fact that mean RT and RT standard deviations are positively correlated, a measure of variability independent of raw performance, referred to as residual variability z-scored, was created by regressing RT standard deviations on RT means, which was then used to remove their shared variance (Robinson & Tamir, 2005). The first RT study conducted involved making multiple semantic distinctions of a target word (i.e., chair) between two categories (i.e., animal vs. not animal), which would appear before and after presentation of the target word. Across 368 trials with 16 different category endpoints, N was found to be significantly associated with z-scored residual variability, with a medium effect size. Because some category endpoints in this study involved affective states (i.e., pleasant vs unpleasant) cognitive associations with those states could have mediated the relationship between N and residual variability. To address this limitation, a Stroop-style RT task, which had participants report the

font color of the word “red” or “green” crossed with the corresponding or opposite color, was used in the second study. Across 252 trials, a significant relationship was found between N and residual variability, with a larger effect size than the affective RT task. To assess the effects of varying task complexity, the third RT study included a simple RT task, which involves only stimulus detection, a go/no-go RT task, which involves stimulus detection and discrimination, and a choice RT task, which involves at least stimulus detection, discrimination, and response selection. No main effect was observed for task-type, so residual variability was averaged over the three tasks and a significant relationship was again observed between N and residual variability. Despite a lack of effect for task-type in the third study, the effect sizes of the N-residual variability relationship varied across the three studies, with the largest effect observed in the Stroop-style RT task, which has the greatest cognitive load.

Another line of evidence for the MNH comes from continuous tracking paradigms, which involve continuously tracking a target on a computer screen with a mouse or joystick. Performance on these tasks capture micro-momentary lapses of attention, with worse performance indicating more lapses of attention (Klein & Robinson, 2019). This paradigm also sought to test the prevailing theory of threat-sensitivity by presenting an aversive noise randomly throughout the trials. In the first study, a Saitek Aviator joystick was used to continuously track a white “+” sign eight times for 30 seconds, with joystick and target position sampled every 100 ms. An 80-dB white noise sound was presented at random in four of the eight trials and a repeated measure ANOVA was used to assess the effect of aversive noise. Interestingly, a main effect was observed for noise but in the opposite direction hypothesized, with performance slightly better in the presence of white noise. A second general linear model analysis was then conducted with N used as a z-scored predictor. A main effect was observed for N, but not for the

N x noise interaction. Estimated performance means for N revealed performance 15% worse at 1 SD of N above the mean relative to 1 SD below. To address concerns that individuals may systematically have varying degrees of familiarity with a joystick, study two used a regular computer mouse. The white noise blasts were also made more aversive by increasing the volume to 90 dB and reducing the length to 200 ms. Across 60 eight second trials, an initial analysis found a main effect for sound with performance worse in the presence of sound compared to the absence. In a follow up analysis which added N as a z-scored predictor, a main effect was again observed for N, with estimated performance means revealing 23% worse performance at 1 SD above the mean relative to 1 SD below. No N x sound interaction was observed in the second study. These results were replicated again in a third study, which found main effects for sound, N, and no N x sound interaction.

To provide external validity for the results, the fourth study involved testing whether poor tracking performance could serve as an implicit measure of personality and significantly predict daily experiences of negative affect (Klein & Robinson, 2019). To assess daily experiences of negative affect, 100 participants were asked to follow a daily diary protocol for 14 days. Each day of the protocol participants were sent an email containing two negative affect markers, “distress” and “nervousness”, and two positive affect markers, “excited” and “enthusiastic” from the positive and negative affect scale (Watson et al., 1988) along with a questionnaire that assessed common stressors (i.e., “I have a lot of responsibilities”, “I have a lot of things to do at once”). After 14 days, diary data was analyzed with a multilevel modeling procedure, with z-scored tracking performance as a level 2 predictor and person-centered daily stress as a level 1 predictor (Klein & Robinson, 2019). In the first analysis participants with poor tracking performance were found to experience higher levels of daily negative. A second analysis was

then conducted to determine if the relationship between poor tracking performance and negative affect was contingent on daily experiences of stress. Using a cross-level model, the tracking performance-negative affect relationship remained significant and negative affect levels increased with the number of daily stressors. No cross-level interaction was found between poor tracking performance and number of daily stressors indicating that individuals with poor tracking performance were prone to higher levels of negative affect regardless of the number of daily stressors.

Observing patterns of regional brain activity offer another line of evidence in support of the MNH. The mental noise conferred by high N is illustrated in an fMRI study, which used state transitions as biomarkers for new thoughts (Tseng & Poppenk, 2020). Studies investigating the neuroscience of cinematics have shown that a well-made movie will result in similar brain activity across multiple participants and thus is able to guide brain state transitions, with structures such as the hippocampus and angular gyrus tuned to the boundary of events. Separate studies have also linked these regions to spontaneous thought generation and the segmentation of neural events. It has been proposed that these regions are involved in semantic integration, as both novel thought generation and content segmentation represent integrating new information into existing representations resulting in a shifted focal point which progresses existing narratives. Therefore, it is hypothesized that movie-viewing as well as resting neural state transitions represent high level semantic features, such as the boundaries of events, as opposed to purely perceptual processes. Data in this study were taken from the 7 T Human Connectome Project dataset, which contains fMRI data while movie viewing and during rest. fMRI runs were converted into expressions in 15 brain networks over time using *t*-SNE and were reduced from 15 x time to 2x time, which resulted in similar patterns of network activity lying close in

proximity to one another. It was hypothesized that the observed patterns of spatiotemporal organization reflected series of discrete thoughts, each with their own focal point (i.e., what will be eaten for dinner). Meta-state changes, changes in network activity over time, were identified by using squared Mahalanobis distances between points in time for each fMRI run and labeled as a step distance vector. Peaks in each average step vector were interpreted as substantial reconfigurations in network meta-states and labeled as transitions. N was found to be significantly associated with higher transitions during rest and lower conformity (i.e., idiosyncratic transitions relative to other stable movie-induced states) when movie viewing. Other neuroimaging studies, that utilize an EEG, have found greater resting state activity in the right posterior region in individuals high in N relative to controls (Schmidtke & Heller, 2004).

P300 Event Related Potential

The P300 (P3) wave is defined as parieto-central positive deflection of the EEG waveform thought to be elicited during the stimulus evaluation phase of cognition (Picton, 1982). The components, which are defined by Donchin et al. (1978) as "a source of controlled, observable variability", are P3a and P3b (Picton, 1982). The P3a component is a deflection that peaks around 250 ms in the frontal regions and occurs when a stimulus is noticed but not attended to. The P3b component, peaks 250-500 ms after stimulus presentation and is reliant on an attended stimulus. Greater P3b amplitudes, defined as the most positive peak in the voltage deflection 250-500 ms after an attended stimulus is presented, have been associated with greater cognitive resource allocation and context updating, the updating of working memory (Donchin & Coles, 1988). Longer P3b peak latencies, defined as the difference between stimulus presentation and the positive P3 deflection, have been associated with delays in stimulus evaluation and strongly correlate to RT in choice RT paradigms (Kutas et al., 1977). Other processes that may

occur simultaneous to the P300 are the N2 wave (recognition of mismatch negativity), which is thought to represent perceptual registration of repeating auditory stimulus changes and the Nd wave (processing related negativity), which occurs when the attended stimulus is being processed (Picton, 1982). P3b along with N2 latencies have been found to be potential markers for amnestic mild cognitive impairment, an intermediate stage between age related cognitive decline and dementia (Cid-Fernandez et al. 2019). Other clinical applications of ERPs have found significant associations between RT/P3 correlations and dementia and reduced auditory P3b amplitudes and future psychosis onset (Hamilton et al., 2020; Pfefferbaum et al., 1984). Neuroticism's relation to performance variability and micromomentary lapses of attention in behavioral studies, and higher resting state and idiosyncratic transitions in neuroimaging studies, make it likely that the mental noise (i.e., negative thoughts and task unrelated ruminations) conferred by N will variably disrupt the stimulus evaluation phase of cognitive processing. Thus, P3b latency, which is associated with stimulus evaluation time and strongly correlates to RT (Kutas et al., 1977), is likely to be affected by N's mental noise resulting in a greater latency variance across multiple RT trials. The objective of this project is to determine whether the latency variance of the P3b during a Stroop-style RT task can serve as a biomarker for N. Specifically, it is hypothesized that N will be significantly associated with greater P3b latency variance across multiple trials of a Stroop-style RT task and that this relationship will be greatest in incongruent trials, "red" or "green" text crossed with the opposite font color, where the cognitive resources required to make a correct decision are greatest.

METHODS

Participants

38 undergraduate psychology students at Northern Michigan University in Marquette, Michigan participated in the study. One participant was excluded due to a corrupted data file. Participants ranged from 18 – 40 years of age (*M* = 19.84 years, *SD* = 3.64 years) and were comprised of individuals from European American, 78%, African American, 10%, Asian, 2%, Native American, 3.33%, mixed, 3.33%, and other, 3.33%, ethnicities. 61% of the participants identified as female, 37% identified as male, and 2% identified as non-binary. Groups were chosen based on high and low levels of N, which were measured in a Psychology 100 lab assignment that administered a Big Five domain-level inventory outlined by Goldberg (1992). Selection into the low and high groups were empirically determined based on the quartiles of N scores, with the low group consisting of the bottom 25% of scores and high group consisting of the top 25%. Every person who met low and high group criteria were then invited to participate in the study. After the corrupted data file was excluded, the high N group contained 22 participants and the low group contained 15 participants. The high N group (*M* = 18.81, *SD* = 2.84) consisted of scores ranging from 11 to 23 and the low N group (*M* = 38.47, *SD* = 5.03) consisted of scores ranging from 33 to 55. Participants were incentivized with extra credit and entry into a raffle for a $100 Amazon gift card.

Personality Measurement

The 50 item International Personality Item Pool (IPIP) representation of markers for the structure of Big Five Factors outlined by Goldberg (1992) was used to collect initial personality data in the lab section of each participant's PSY 100 course. This personality inventory is a

public domain set of items that can be used free of charge. Each personality domain was assessed with 10 items using a 5-point Likert scale. Higher scores on this inventory represent lower levels of trait neuroticism. Items in the neuroticism domain included “I get stressed out easily” and “I am easily disturbed”. The Big Five Aspects Scale (BFAS), developed by DeYoung et al. (2007), was completed after initial subject recruitment. This scale contains two lower order aspects in each Big Five domain that have been found to account for the variance of a larger set of facet scales. The N aspects in BFAS, withdrawal (W) and volatility (V), reflect the traditional dichotomy between internalizing and externalizing psychopathology. Withdrawal is measured with items such as “I am filled with doubt about things” and “I worry about things” whereas volatility is measured with items such as “I get angry easily” and “I change my mood a lot”. BFAS is also a public domain set of items that can be used free of charge. Each aspect was measured with 10 items using a 5-point Likert scale. Higher scores on the BFAS indicate higher levels of neuroticism aspects. A follow up analysis with BFAS was used to assess test-retest reliability of N scales and to conduct exploratory analyses for each of N’s lower order aspects.

Brief State Rumination Inventory

Each participant’s level of rumination was assessed with the Brief State Rumination Inventory (BSRI) developed by Marchetti et al. (2018). This scale is a public domain set of items that can be used free of charge. State rumination is measured with items such as “Right now, I am reflecting about my mood” and “Right now, I am rehashing in my mind recent things I’ve said or done”. Each item was measured with a 100-point visual analogue scale, with increments at every 10 points along the scale. Higher scores on the BSRI indicate greater levels of rumination. The BSRI was included to provide support that the mental noise conferred by high levels of N was driving the effect between N and intra-participant variability in P3b latency.

Stroop-Style Reaction Time Task

The Stroop-style RT task, programmed in E-Prime, contains choice reaction processing phases with a more difficult evaluation phase achieved through presentation of incongruent stimuli. The three classes of stimuli presented were congruent, “green” or “red” with its corresponding font color, incongruent, “green” or “red” with the opposite font color, or control, “xxxxx” with either green or red font. After a one second fixation, each stimulus appeared until the participant responded, which resulted in a reappearance of the fixation followed by another word. After stimulus presentation, participants were instructed to quickly and accurately respond with the font color of each word.

Electroencephalography Measurement

A 64 channel Geodesic Sensor Net (Electrical Geodesics, Inc., Eugene, OR) with AgCl electrodes placed in accordance with the 10–20 international system was used to collect EEG readings. EEG data were recorded using Net Station 4.5.4 software (Electrical Geodesics, Inc., Eugene, OR) at a sampling rate of 500 Hz re-referenced to the cap average. After recording, raw segmented EEG readings 200 – 1000 ms post stimulus presentation were bandpass filtered at 0.1-3 Hz in Net Station 4.5.4 resulting in a frequency range that has been optimized for single-trial isolation of the P3b component (Ouyang et al., 2017). After band pass filtering, electrodes in the parietal region (Pz, P1, P2, POz, PO1, PO2, Oz, O1, O2) were averaged during each 800 ms epoch. These electrodes were chosen based on visual inspection of waveform averages, prior research implicating the parietal-central region in P3b generation, and to account for micro-spatial differences in electrode positioning between participants (Picton, 1982). The electrodes from this montage are presented visually in Figure 1. All trials with a positive cross-correlation

coefficient were considered good trials and data retention per participant ranged from 53 – 100% (*M* = 85%, *SD* = 13%).

Adaptive Filtering

Parietal electrodes averaged from each individual's segmented files were passed to a custom-built python 3.10 program, which implemented an Adaptive Filtering algorithm, developed by Woody (1967). The adaptive filtering algorithm involves cross correlating single-trial readings to a waveform template, with the first iteration of this procedure using the waveform average in place of the template. Cross-correlation coefficients were calculated for all time intervals through multiplying the reciprocal of the total time of both waveforms by the integral of the waveform average (or template) cross multiplied by the single trial reading shifted by a lag value. The lag value corresponding to the maximum value of the cross-correlation coefficient was then used to shift each single trial, which was placed into an average bin used to create the template. After all trials were shifted by the lag value associated with the maximum value of the cross-correlation coefficient, a second iteration of this procedure occurred, with the template replacing the waveform average. Each shifted trial's latency was found by locating the positional value of the maximum peak in amplitude, which was then converted into millisecond latency by dividing the positional value by the sampling rate (500 Hz) and multiplying that value by 1000. After each participant's single trial latencies were determined, the variance of their latency was computed by calculating their latency SD over 288 trials.

Procedure

Prior to arriving in the lab, participants completed the 50 item IPIP representation of the structure of Big Five Factors identified by Goldberg (1992) in the lab section of their Psych 100 course. This data was analyzed and individuals in the bottom and top quartiles were invited to

participate in the study. The invitation email included the informed consent form as well as the Big Five Aspects Scale, which was completed prior to scheduling a time in the lab. Upon arrival, participants were provided with a questionnaire collecting basic demographic information. The 64 channel Geodesic Sensor Net was then placed on the participant's head. Participants were instructed to press the leftmost key of a response box if the font color was red or the rightmost key of a response box if the font color was green, with the index fingers of their left and right hands. They completed 60 practice trials followed by 4 experimental blocks, each with 72 trials. After completing the Stroop-style RT task, the Brief State Rumination Inventory was administered.

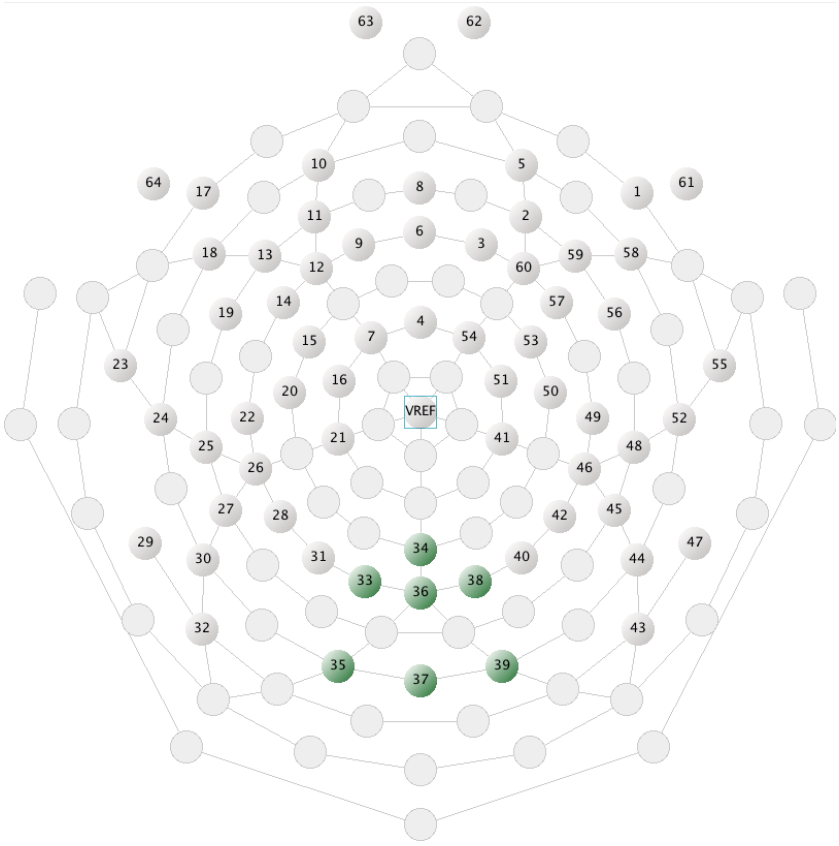
Analysis

Single trial P3b data were obtained with the previously described, custom-built python 3.10 implementation of an Adaptive Filtering algorithm developed by Woody (1967), which resulted in a latency measure for each of the participant's 288 trials. Each participant's single trial P3b latency was then used to compute the latency SD across all trials, the target dependent variable. An example of one participant's P3b deflection shifted one standard deviation above and below their jitter corrected mean is displayed visually in Figure 2. After all data were collected, a 2x3 mixed factorial designs ANOVA, with high and low N as a between subjects variable and stimuli class as a within subjects variable, was used to assess the effects of N and stimulus class on P3b latency variance. All analyses and data visualization were conducted in R 4.1.3 in the R Studio 3.30 IDE. It was hypothesized that a significant main effect for N and stimulus class on P3b latency variance would be observed with greater latency variance in the high N group and when participants were presented with incongruent stimuli. A significant N x

stimulus class interaction was also expected with the greatest latency variance observed when the high N group was presented with incongruent stimuli.

Figure 1

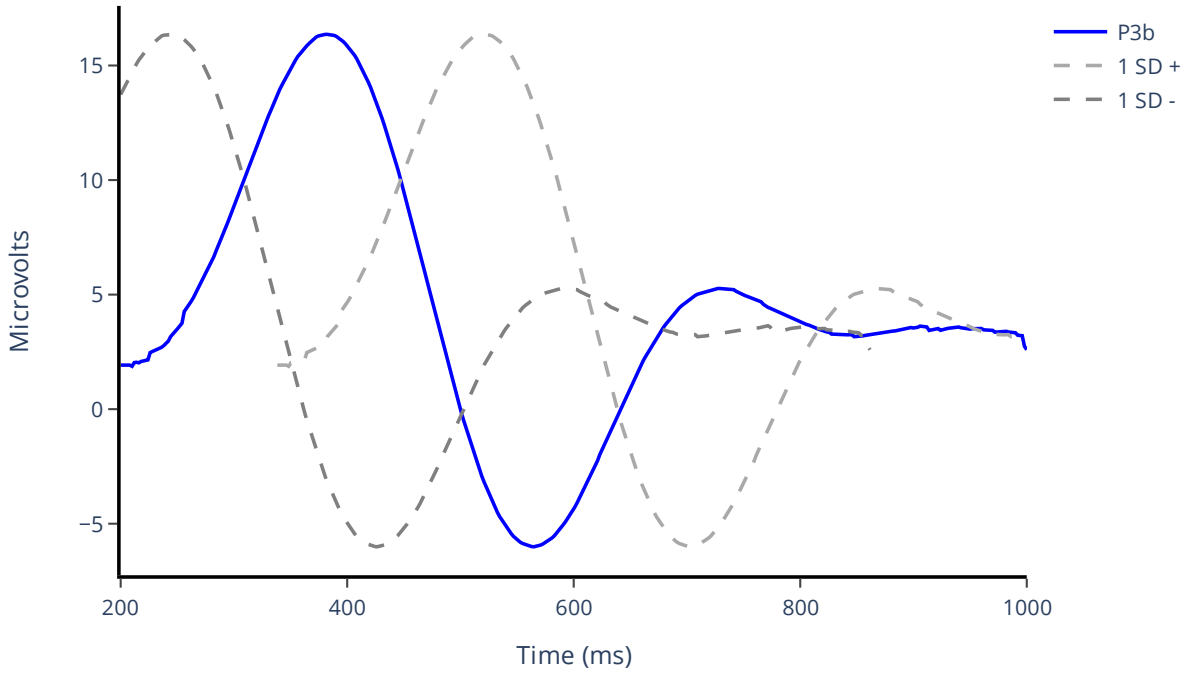
Montage Used in Primary Analysis



Note. Each numbered circle represents an electrode and each circle highlighted in green represents an electrode used in the primary analysis. The exploratory analyses at site Pz are represented with electrode 34.

Figure 2

Example of P3b Latency Variance Using Data from Subject 2



Note. Time represents milliseconds since stimulus presentation, the blue line represents the participant's jitter corrected P3b average, the dark gray line represents the jitter corrected P3b average shifted one standard deviation below the mean, and the light gray line represents the jitter corrected P3b average shifted one standard deviation about the mean

RESULTS

Differences in Intraparticipant Latency Variance by N Group

Initial assessment of neuroticism, using the IPIP representation of Big Five Markers outlined by Goldberg (1992), significantly correlated with the W, $r(35) = -.67$, $p < .01$, 95% CI [-.82, -.43] and V, $r(35) = -.69$, $p < .01$, 95% CI [-.83, -.46], subscales from the BFAS developed by DeYoung et al. (2007). The directionality of the coefficients is due to reversed scoring in the two inventories (i.e., higher scores indicate less neuroticism in Goldberg's (1992) inventory). Summary statistics of composite N scores (combination of scores on the V and W subscales) on the BFAS for the high and low N groups are presented numerically in Table 1. Before fitting the model, outliers were assessed using Tukey's interquartile range method (IQR) for outlier detection, which involves calculating upper (Quartile 3 + 1.5 × Interquartile range) and lower (Quartile 1 – 1.5 × Interquartile range) limits for accepted values. Values that exceed the upper and lower limits were considered outliers and one participant in the low N group exceeded the lower bound limit. Due to the low sample size of the groups, however, it could not confidently be determined that this participant would be an outlier given a larger sample size and thus was retained in the analysis. A Shapiro-Wilk test was used to assess normality and indicated no significant deviations from a Gaussian distribution, *ps* > .13. The assumption of homogeneity of variance was assessed using Levene's test, which indicated homogenous variances across groups for each stimulus type, *ps* > .13. Homogeneity of covariances was assessed using Box's M test, which indicated no heterogeneity in the covariance matrices $p > .05$. Sphericity was assessed using Mauchly's test for sphericity and no significant deviation from sphericity for the within-subjects factor was observed, $p = .54$. Error rates did not exceed 8% and were included in

analysis due to prior literature which found the emergence of the P3b following motor response in error trials, suggesting that in error trials participants are still processing stimuli after a reaction is made (Woody, 1967).

Table 1

Summary Statistics of Composite BFAS N Scores by N Group

#	Group	N	Mean	SD
1.	High	22	72.14	11.89
2.	Low	15	47.93	9.96

Note. The composite BFAS neuroticism scale ranged from 0 – 100 with higher values indicating greater levels of trait neuroticism.

After ensuring that there was no violation of assumptions, a 2x3 mixed designs ANOVA was used to investigate the hypotheses that there would be a significant main effect for N group and stimulus type and that there would be a significant interaction, with the high N group exhibiting the greatest variability when encountering incongruent stimuli. The results from this model indicated that neither N group, $F(1, 35) = 0.018$, $p = .89$, $\eta^2 = .00037$, nor stimulus type (congruent, incongruent, control), $F(2, 70) = 0.16$, $p = .85$, $\eta^2 = .001$ predicted intra-participant latency variability. An interaction between N group and stimulus type also failed to meaningfully predict intraparticipant latency variance, $F(2, 70) = 0.38$, $p = .68$, $\eta^2 = .0030$.

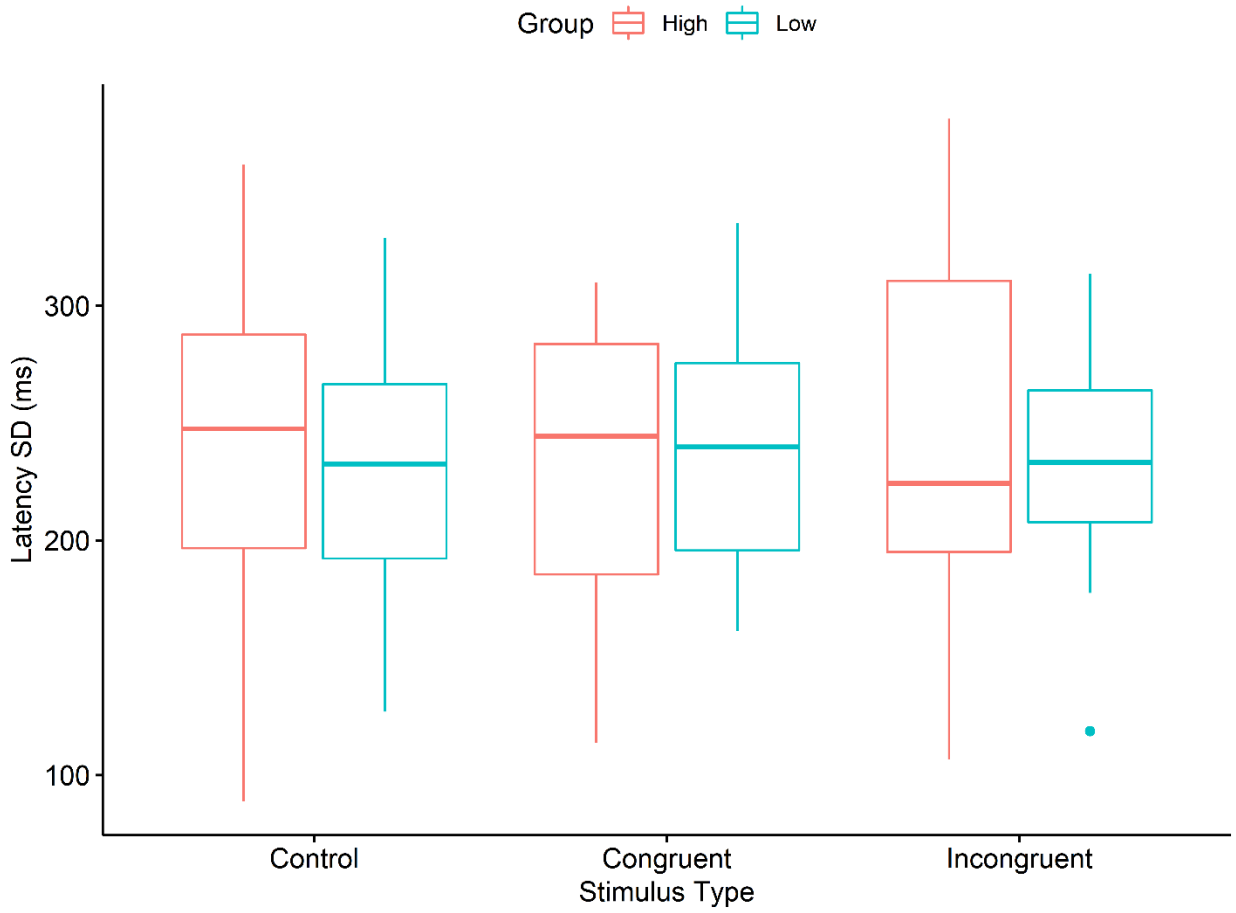
N Aspects, Rumination, and Intraparticipant Variability

Significant correlations were observed between rumination and withdrawal, $r(35) = .57$ $p < .01$, 95% CI [.30, .75], rumination and volatility, $r(35) = .50$, $p < .01$, 95% CI [.21, .71], and volatility and withdrawal, $r(35) = .72$, $p < .01$, 95% CI [.52, .85]. Significant correlations between intraparticipant variability across different stimulus types were also found, $rs(94) = .57$ - $.65$, $ps < .05$. Rumination and N aspects, however, were not associated with levels of intrasubject

variability, *rs*(35) = -.08 - .14, *ps* > .05 in all stimuli classes. Intra-participant latency variance by group and stimulus type is presented visually in Figure 3. The correlations between N aspects, rumination, and latency SDs for each stimulus type are visually presented in Figure 4. Summary statistics of latency SDs by stimulus type and group are listed numerically in Table 2.

Figure 3

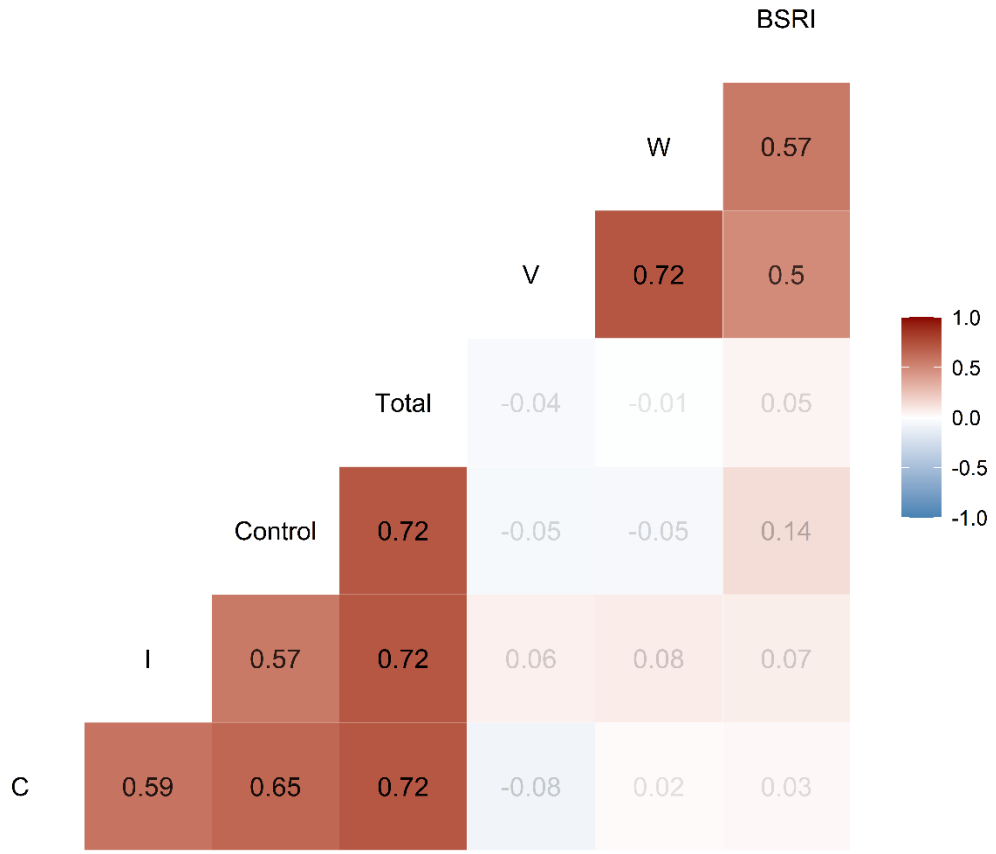
Intraparticipant Latency Variance by N Group and Stimulus Type



Note. Dots represent data that has exceeded the maximum or minimum accepted value outlined by Tukey's IQR method of outlier detection.

Figure 4

Correlations between N Aspects, the BSRI, and Latency SDs by Stimulus Type



Note. Each cell represents the correlation coefficient between two variables. Shaded in cells represent p values less than .05 and faded cells represent p values greater than .05. The color denotes the strength of the correlation coefficient, with darker shades of red representing positive correlations and darker shades of blue representing negative correlations. W denotes withdrawal, V denotes volatility, BSRI denotes the brief state rumination inventory, Total denotes the total intrasubject latency SD across all stimulus types, I denotes the intrasubject latency SD in incongruent trials, and C denotes the intrasubject latency SD in congruent trials, and Control denotes intrasubject latency SD in control trials.

Table 2

Summary Statistics of Latency SD by Group and Stimulus Type

#	Group	Stimulus	N	Mean	SD
1.	High	Control	22	240	67
2.	High	Congruent	22	229	62
3.	High	Incongruent	22	241	74
4.	Low	Control	15	234	57
5.	Low	Congruent	15	236	52
6.	Low	Incongruent	15	233	49

Note. Group denotes high or low neuroticism group, stimulus denotes stimulus type, and N denotes the number of participants per group. Mean and SD values are in millisecond units.

Interparticipant Variance in Intraparticipant Variance

A visual inspection of differences in interparticipant variance of within subject latency variability when processing incongruent stimuli, listed in Table 2, between the high N group (M = 241 ms, SD = 74 ms) and low N group (M = 233 ms, SD = 49 ms) prompted further investigation into whether the observed interparticipant variability in the high N group was meaningfully greater than the low N group. A two-tailed F-Test for equality of variances was conducted between high and low N groups, when processing incongruent stimuli, and indicated that the between-participant variance of within participant variability was not significantly greater in the high N group, *Variance Ratio* [VR] = 2.25, 95% CI [.80, 5.76], p = .12. While not statistically significant, the observed differences in variance gave rise to exploratory analyses into differences in interparticipant P3b latency by N group and stimulus type.

Differences in P3b Latency by N Group

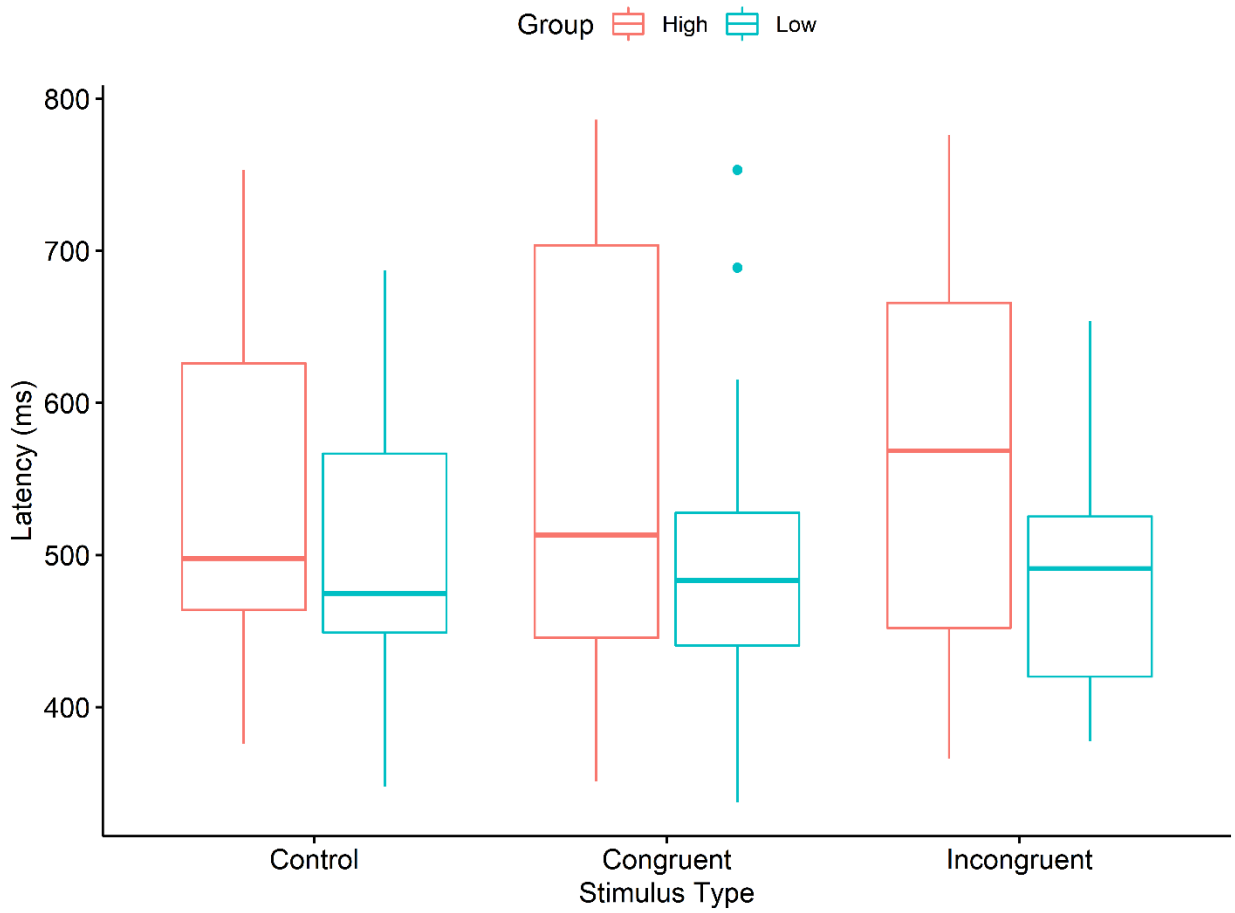
Due to discrepancies in interparticipant variance between N groups, an exploratory 2x3 mixed designs ANOVA was conducted with N group as a between subjects variable, stimulus type as a within subjects variable, and each participant's average P3b latency as the outcome variable. Latency averages for each participant were obtained by averaging their single trial latencies obtained from the Adaptive Filtering algorithm. When assessing outliers with Tukey's IQR method, two participants in the low N group exceeded the upper bound limit. These participants were retained in analysis, however, due to the low sample size of the groups potentially biasing the IQR observed in the sample. After running the model, no main effect was found for group, $F(1, 35) = 1.86$, $p = .18$, $\eta^2 = .044$, or stimulus type, $F(2, 70) = 0.14$, $p = .87$, $\eta^2 = .00048$. Additionally, no significant group by stimulus type interaction was observed, $F(2, 70) = 1.88$, $p = 0.16$, $\eta^2 = .007$. P3b latency by group and stimulus type is presented visually in Figure 5 and numerically in Table 3. In addition to no main effect for N group, Levene's test indicated that the assumption of homogeneity of variances was violated, $p = .02$ and the Shapiro-Wilk test indicated that the assumption of normality was violated, $p = .049$. Because of these violations, the lack of main effect for stimulus type, and prior literature that found no P3b latency differences based on congruence conditions (Ila & Poich, 1999), a one-way Welch's ANOVA was conducted with N group as the between subjects variable and P3b latency over three stimulus types as the outcome.

Welch's ANOVA was chosen due to its ability to test for group mean equality with nonnormal heteroskedastic distributions without a significant loss of power (Cribbie et al., 2007). Outliers were assessed using Tukey's IQR method and no participants exceeded the upper or lower bound limits. The results of this model indicated no main effect of group on P3b latency,

Welch's $F(1, 34.4) = 0.98$, $p = .33$. P3b latency by N group is displayed visually in Figure 6 and numerically in Table 4.

Figure 5

P3b Latency by N Group and Stimulus Type



Note. Dots represent data that has exceeded the maximum or minimum accepted value outlined by Tukey's IQR method of outlier detection.

Table 3

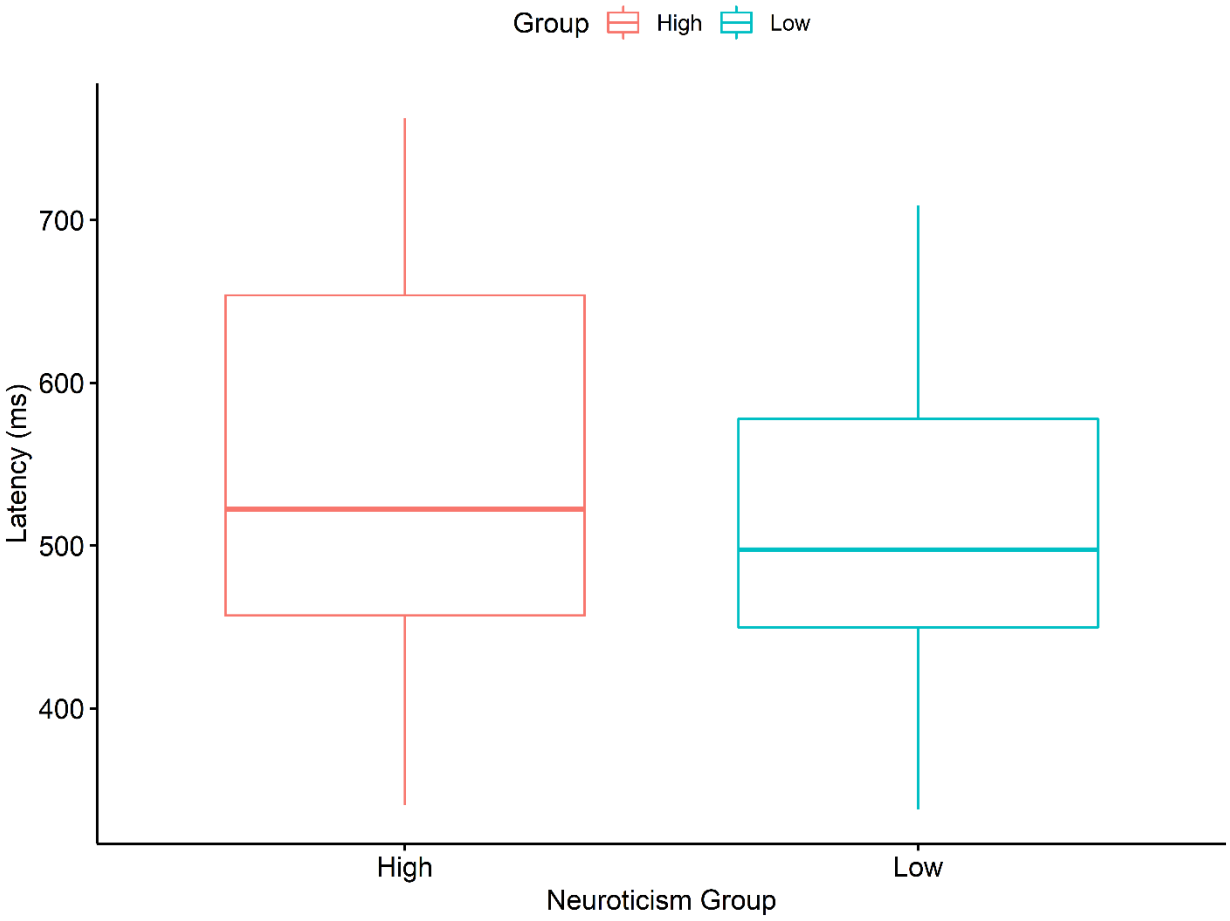
Summary Statistics of Latency by Group and Stimulus Type

#	Group	Stimulus	N	Mean	SD
1.	High	Control	22	536	112
2.	High	Congruent	22	557	148
3.	High	Incongruent	22	563	132
4.	Low	Control	15	510	107
5.	Low	Congruent	15	502	113
6.	Low	Incongruent	15	489	89.9

Note. Group denotes high or low neuroticism group, stimulus denotes stimulus type, and N denotes the number of participants per group. Mean and SD values are in millisecond units.

Figure 6

P3b Latency by N group



Note. No data exceeded the upper or lower bound limits for accepted values outlined by Tukey's IQR method for outlier detection

Table 4

Summary Statistics of Latency by Group

#	Group	N	Mean	SD
1.	High	22	544	135
2.	Low	15	505	103

Note. Group denotes high or low neuroticism group, N denotes the number of participants per group. Mean and SD values are in millisecond units.

Exploratory Latency Analyses at Site Pz

While the relationship between N group and P3b latency using the average signal from a montage of seven parietal electrodes was insignificant, prior research has found temporal differences in P3b emergence between central-parietal electrodes, with electrode Pz (represented with electrode 34 in Figure 1) exhibiting the shortest latency and greatest amplitude (Wei et al., 2020). Due to site Pz potentially exhibiting greater sensitivity to P3b latencies, exploratory analyses were conducted between P3b latency recorded at site Pz and N groups, N aspects, and rumination. One subject in the low N group exceeded the upper quartile limit, but was included in the analysis, again due to the low sample size of the groups potentially biasing the IQR observed in the sample. The results from the Welch's ANOVA, with N group as a between subjects variable and P3b latency at site Pz as the outcome, failed to find a significant relationship between N Group and P3b latency at site Pz, Welch's $F(1, 33.4) = 2.91$, $p = .097$, estimated $\omega^2 = .049$. A follow-up analysis with N aspects and rumination, however, found a significant positive correlation between withdrawal and latency $r = .34$, $p = .038$, 95% CI [.02, .60]. Positive correlations between volatility and latency, $r = .21$, $p = .21$, and rumination and latency, $r = .30$, $p = .069$, were also found but neither surpassed a significance threshold of $p < .05$. P3b latency at site Pz by N group is presented visually in Figure 7 and numerically in Table 5. The correlation between withdrawal and P3b latency is presented visually in Figure 8. Due to the heterogeneity in P3b latency variance at site Pz between high and low N groups, an *F*-Test for equality of variances was conducted between the high and low N groups. The results of this analysis indicated that across all stimulus types, the high N group exhibited significantly greater variance compared to the low N group, *Variance Ratio* [VR] = 3.52, 95% CI [1.24, 9.02], $p = .019$. P3b latency variance by N group, recorded at site Pz is displayed visually in Figure 9.

Table 5

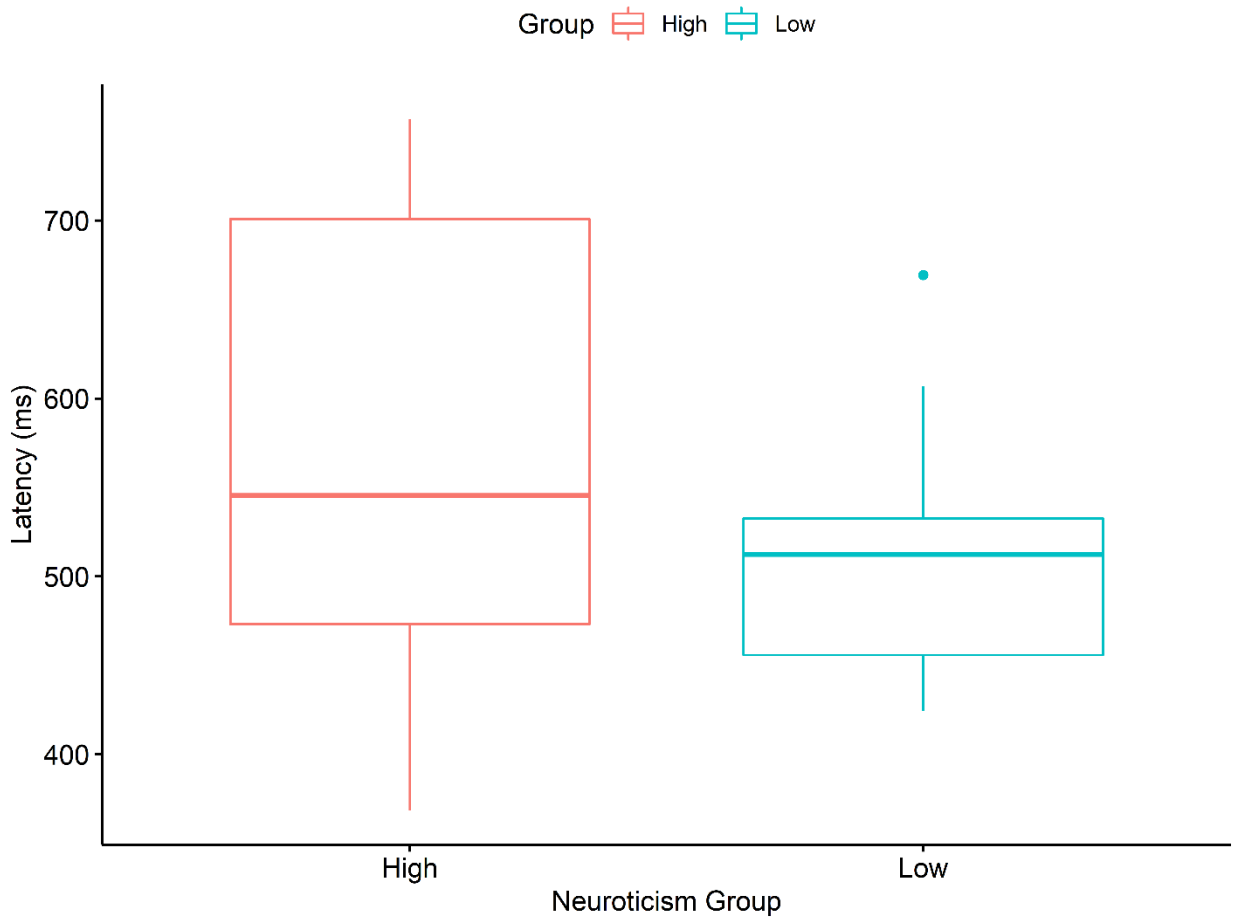
Summary Statistics of Latency at Site Pz by Group

#	Group	N	Mean	SD
1.	High	22	567	130
2.	Low	15	510	70

Note. Group denotes high or low neuroticism group, N denotes the number of participants per group. Mean and SD values are in millisecond units.

Figure 7

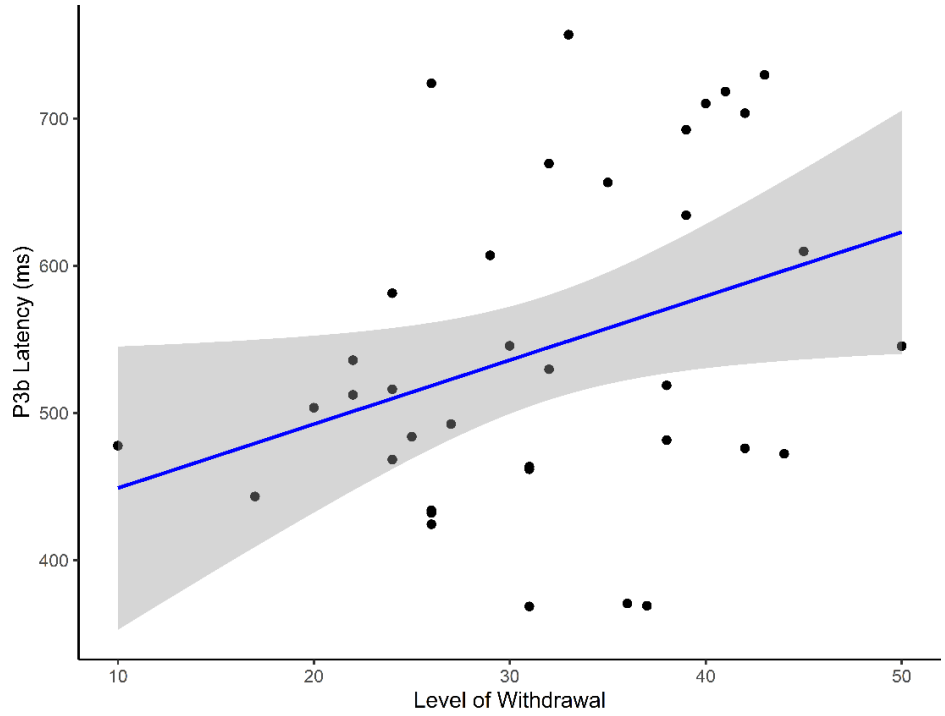
P3b Latency at Site Pz by Group



Note. Dots represent data that has exceeded the maximum or minimum accepted value outlined by Tukey's IQR method of outlier detection.

Figure 8

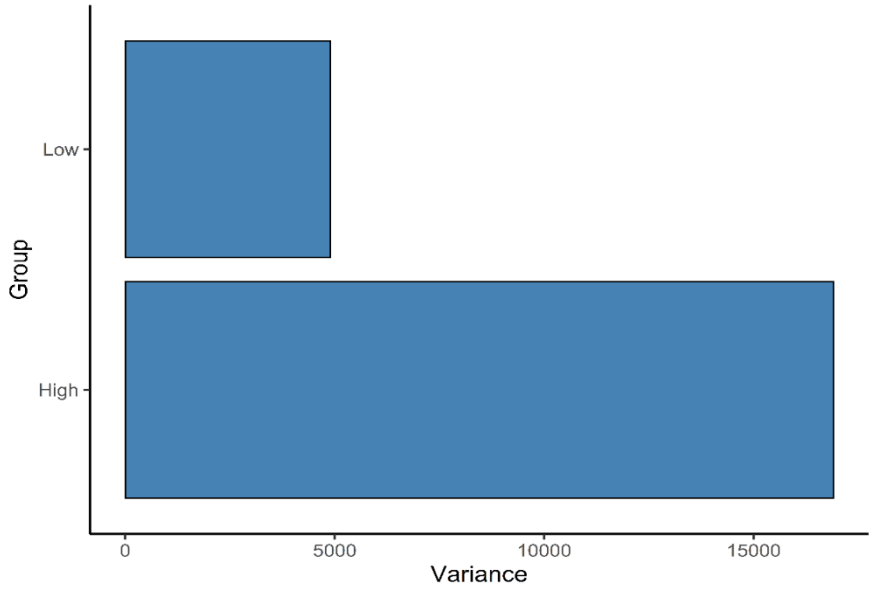
P3b Latency at Site Pz by Level of Withdrawal



Note. Bars surrounding the best fit line indicate the standard error of the mean.

Figure 9

Interparticipant P3b Latency Variance at Site Pz by Group Across All Trials



Note. Variance is presented in square millisecond units

DISCUSSION

The purpose of this study was to evaluate the degree to which mental noise, task-unrelated cognitions and ruminations found to be associated with N in prior literature (Robinson & Tamir (2005), would variably disrupt the onset of stimulus evaluation resulting in a greater P3b latency variance over each participant's trials. To this aim, the project has failed to provide support of intraparticipant variance in P3b latency as the primary neural generator behind the intraparticipant variability in choice RT observed in prior studies. An examination of latency differences between high and low N groups, however, has mirrored the variance in existing literature, where both positive, negative, and no relationships between P3b latency and N have been reported (Fjell et al., 2005; Stelmack et al., 1993; Vorkapic & Tadinack, 2016). Additionally, the exploratory analyses at site Pz support recent findings, linking delays in auditory P3b at site Pz to psychiatric disorders that share common genetic pathways with N, (Luciano et al., 2018; Mathalon, et al., 2000). Lastly, examining interparticipant variability in P3b latency at site Pz between high and low N groups has yielded interesting results that can contribute to the growing personality neuroscience literature.

The MNH and P3b Latency Variance

The theoretical model generating the prediction that individuals with high N would exhibit greater latency variance in the P3b, compared to individuals with low N, is referred to as the mental noise hypothesis and posits that individuals with high N exhibit a more chaotic mental control system (Robinson & Tamir, 2005). Support for this model began with RT paradigms (i.e., simple, Go/No Go, and choice tasks) linking high N to greater RT variability and has subsequently been supported by multiple behavioral (i.e., continuous tracking), psychometric

(i.e., cognitive failure liability), and neuroimaging paradigms (i.e., resting state EEG, fMRI), which all link higher levels of N to greater levels of intrusive thoughts and task-irrelevant cognitions (Flehmig et al., 2007; Klein & Robinson, 2019; Schmidtke & Heller 2004; Tseng & Poppenk, 2020). The primary analysis of this study aimed to establish the P3b as the neural generator of the observed variability reported in prior RT paradigms. The P3b was chosen due to the wealth of existing literature linking its latency and amplitude to delays in cognitive processing and task complexity respectively (Donchin & Coles, 1988; Kutas, 1977; Woody, 1967). Prior literature, which reported a scaling of effect size with task-complexity, resulted in the hypothesis that the observed variability seen in those with high N, was due to variability in cognitive processes underlying stimulus evaluation (Robinson & Tamir, 2005). The close coupling of P3b latency and RT reported in prior studies, along with its implication in stimulus evaluation, made variability in its latency an ideal candidate as the neural generator of the cognitive variability reported for individuals high in trait neuroticism (Woody, 1967).

The primary analysis of intraparticipant variability, measured with P3b latency SDs, did not yield any meaningful results linking high levels of N to greater P3b latency variability, $F(1, 35) = 0.018$, $p = .89$, $\eta^2 = .00037$. A lack of main effect for stimulus type, $F(2, 70) = 0.16$, $p = .85$, $\eta^2 = .001$, further reinforces the notion that earlier stages of perceptual processing or later stages of response selection may be driving the previously reported relationship between N and response variability. An analysis of the summary statistics for P3b latency SDs by N group and stimulus type revealed mean P3b latency SDs that ranged from 229-241 ms in the high group and 233-236 ms in the low group. With the medium effect expected based on the results of prior literature, the present summary statistics do not indicate a meaningful trend towards significance. Due to the low sample size, it is possible that the lack of significant results was from a

nonrepresentative sample. Additionally, longitudinal studies of neuroticism have found peak neuroticism scores at age 20 (Aldinger et al., 2014). It is plausible that the results could change with an older population, as high neuroticism in older age groups may represent a subset of individuals at greater risk for high neuroticism related outcomes. The current results, however, suggest that other neural generators are driving the relationship between N and cognitive variability. This interpretation is also supported by the higher-powered, 42%, follow-up analyses that did not find significant relationships between W, V, rumination, and intraparticipant P3b latency variance, *rs* = -.09 - .12, *ps* > .05. The observed variability previously reported in those with high N could be due to early visual processing as opposed to later stage stimulus evaluation, making the P1 a potential candidate for future research due to its relationship to visual encoding (Dias et al., 2011). Future studies may also wish to assess the N1 in those high in neuroticism, as prior studies of schizophrenia - a psychiatric phenotype linked observationally and biologically to trait neuroticism (Luciano et al., 2018; Van Os & Jones, 2001) - have found significant relationships between N1 parameters and aspects of schizophrenia (Oribe et al., 2013).

High N, N-Related Psychiatric Illness, and Delays in P3b Latency

While prior studies have assessed the relationship between psychiatric phenotypes, domain level N and discrepancies in P3b amplitude and latency, little is known about the relationship between N's lower order aspects derived from factor analysis and prolonged latencies in choice RT tasks. Additionally, few studies include lag-adjusted latency averages when assessing the relationship between personality traits, psychiatric phenotypes, and ERP latencies. Traditional ERP analysis involves averaging the signal over multiple trials, which can lead to an inaccurate interpretation of latencies and amplitudes (Gavin et al., 2019). Computing

lag-adjusted latency averages could result in a measure that is more sensitive to the intertrial variability in P3b latency resulting in a more accurate measure of true latency averages.

Large scale clinical applications of ERPs to psychiatric illness have to a great extent focused on schizophrenia. One of the largest longitudinal studies to date of psychiatric phenotypes, P3b parameters, and RT delays sought to assess differences in cognitive processing that could predict conversion to schizophrenia (Hamilton et al., 2019). This study involved 552 participants who met the criteria for psychosis risk syndrome (PRS) - a disorder that some researchers have argued is an endophenotype between healthy controls and schizophrenia - and 236 controls. Due to the high degree of variance in outcomes, however, using PRS as a sole criterion for early intervention efforts remains controversial. The variance in outcomes for those who reach self-report criteria, further delineates the need for biomarkers that can predicted conversion from intermediate phenotypes to psychiatric illness. The results of Hamilton et al.'s (2019) analysis found significant relationships between diminished P3b amplitude, longer RT, and psychosis conversion in an auditory oddball task, which involves stimulus detection and evaluation. A 1 SD decrease in P3b amplitude was found to increase likelihood of psychosis conversion by 45%, whereas a 1 SD increase in RT was found to increase psychosis conversion by 31%.

The lack of relationship between P3b latency and conversion to psychosis, was likely due to the early stage of illness experienced by participants with recent psychosis conversion. In prior literature, prolongation of P3b latencies has been associated with schizophrenia illness duration (Mathalon et al., 2000), and greater rates of age-relate P3b latency prolongation have been associated with illness severity (O' Donnell et al., 1995). Given the temporal relationship between diminished average P3b amplitude, prolonged P3b latency, and schizophrenia, it is

likely that P3b amplitudes are more sensitive to indexing liability, whereas prolonged latencies may reflect symptomology or lasting effects of pathology. Future research should investigate the relationship between neuroticism aspects and reductions in P3b amplitude, as this parameter may better index liability for conversion to more severe clinical phenotypes than P3b latency.

Studies investigating the relationship between N and P3b latency have yielded mixed results (Vorkapic & Tadninac, 2016). While some studies report negative relationships between domain level N and P3b latency (Stelmack et al., 1993; Pritchard, 1989), others report positive relationships (Vorkapic & Tadninac, 2016), and some no relationship (Fjell et al., 2005). The primary mode of P3b elicitation in studies investigating the N-P3b latency relationship, as well as the psychiatric disorder-P3b relationship, has been auditory oddball tasks (Vorkapic & Tadninac, 2016), which lack the response selection and visual components of the Stoop-style RT task employed in this study. The lack of main effect for stimulus type in the 2x3 mixed designs ANOVA with P3b latency as the outcome, F (2, 70) = 0.14, p = .87, η^2= .00048, as well as prior literature that found a similar lack of effect (Ila & Poich, 1999), supports the notion that the RT discrepancies from congruence conditions reflect variability in response selection rather than stimulus evaluation. Additionally, prior research investigating differences between auditory and visually evoked P3bs has found no significant differences between ERPs elicited from auditory and visual tasks (Comerchero & Polich, 1999), allowing for relatively straightforward comparisons between prior research and the current study.

The current analysis investigating the relationship between P3b latency and N mirrors the discrepancies in prior research investigating these variables. In the first set of analyses with the signal averaged over the parietal montage, the 2x3 mixed designs ANOVA, N group was found to have no significant effect on P3b latencies, F(1, 35) = 1.86, p = .18, η^2 = .044. The value of

η^2, however, makes it likely that an effect would be observed with an increase in power. A violation of parametric assumptions resulted in the application of the more robust Welch's ANOVA, which again found no significant effects of group on average P3b latencies Welch's $F(1, 34.4) = 0.98$, $p = .33$. A follow up analysis at site Pz, however, indicated results with a greater trend towards significance. The results of the one-way Welch's ANOVA at site Pz, $F(1, 33.4) = 2.91$, $p = .097$, $\omega^2 = .049$, would likely be significant with a greater sample size. Additionally, when assessing the relationship between N aspects and P3b latency at site Pz, a significant correlation was observed between the withdrawal subscale and P3b latency $r = .34$, $p = .038$, 95% CI [.02, .60]. A likely explanation for the discrepancy in results between the averaged signal in the parietal montage and site Pz, is the difference in precision between the two signals, with site Pz likely representing a location more sensitive to the P3b (Wei et al., 2020). Future analyses, should further compare the strength and temporal qualities of the P3b across a broader range of parietal electrodes.

The significant relationship between withdrawal and P3b latency at site Pz and lack of relationship between volatility and P3b latency, provides discriminant validity for the lower order N aspects outlined in the BFAS. Additionally, the lack of significant relationship between volatility and P3b latencies, $r = .21$, $p = .21$, provides a potential explanation for the heterogeneity in prior research investigating the relationship between N and P3b latency. It is possible that the heterogeneity in prior results is due to uneven sampling distributions, with some containing a greater proportion of volatile-neuroticism and others contain higher amounts of withdrawn-neuroticism. Additionally, the electrodes chosen in prior studies could be another driving force between the discrepancy in results reported in the literature. Due to prior research that has established prolonged latencies as a biomarker for schizophrenia severity and illness

duration, future research should investigate the predictive validity of BFAS subscales compared to more popular domain level inventories. In particular, it is hypothesized that the BFAS withdrawal subscale would likely have greater predictive validity for disorders that contain social withdrawal as a core feature (i.e., schizophrenia, major depression) compared to inventories that contain broader domain level constructs. Despite the broader constructs contained in more popular domain level N inventories, the predictive validity of trait N remains high, with effect sizes between N and its empirically related outcome measures in the medium-high range (Malouff et al., 2005). Future iterations of this project may also wish to use quartiles or standard deviations of withdrawal as a prescreening measure, as this aspect of neuroticism may be more strongly related to P3b parameters.

Group Level Latency Variance

The relationship between N group and greater interparticipant P3b variability at site Pz, across all task conditions, represents one of the more unique findings of this study. As of date not much literature has investigated the relationship between personality variables and interparticipant variance in single trial ERP averages. The greater interparticipant P3b latency variability observed in individuals high in N mirrors decades of findings that have linked high levels of N to a range of seemingly diverse outcome measures, such as a broad collection of psychiatric illness, worsened disease prognosis, and increased risk of mortality from all causes (Lahey, 2009). The range of outcomes, observed in those high in trait N, could be mediated by numerous variables, that may be driving the observed patterns of association. In turn, these variables could all be associated with unique means of cognitive processing, which could result in greater intergroup variance for those high in neuroticism. Additionally, neuroticism encompasses numerous highly correlated but in part distinct features (i.e., anger, sadness,

anxiety, worry, hostility), which each may result in different styles of cognitive processing. While more research is required to tease apart the mechanisms underlying the greater interparticipant cognitive variance accompanied with high neuroticism, the present analysis represents a unique finding that can contribute to the growing field of personality neuroscience.

Limitations

The analyses conducted in this study are not without limitations, with the most apparent being the limited sample size. A consequence of the sample size was that most analyses were not adequately powered to detect a small or medium effect. Given a larger sample size it is possible that the results could change and thus their interpretations should be considered in this light. Another set of limitations were inherent to the Adaptive Filtering algorithm employed to analyze P3b latencies. One of the major problems in this implementation of Adaptive Filtering is the discrepancy in latency estimates when calculating per stimulus type latency averages compared to total latency averages. When calculating single-trial latencies by stimulus type, the templates were created with a third of the data that was available when creating templates using all trial types. This large discrepancy in available data could bias template creation and result in slight shifts in latency estimates, when comparing single trial latency averages by stimulus type to single trial latency averages across all trials. Other limitations inherent to the Adaptive Filtering algorithm include its tendency to reduce latency variance and sub-optimal performance when the signal-noise ratio is low. For instance, when the signal-noise ratio is low, the algorithm could be maximizing the alignment between noise, as opposed to signal. Despite these limitations, the Adaptive Filtering algorithm has been shown in meta-analyses to perform among the best at isolating single trial ERPs elicited from RT paradigms involving stimulus evaluation (Ouyang, et al., 2018). Finally, another set of limitations arises from the lack of confounds assessed in this

study. Given the relationships between psychiatric disorders and parameters of the P3b ERP, each participant's diagnostic status could mediate the results reported in this study.

CONCLUSION

Neuroticism is a personality dimension related to a broad range of psychiatric disorders and negative outcome measures. Traditional theories conceptualize the construct in terms of greater threat sensitivity, but new hypotheses (i.e., the mental noise hypothesis) are emerging that posit a more chaotic mental control system as the primary mechanism underlying the domains salient features. Support for the mental noise hypothesis comes from a range of experimental paradigms, with cognitive variability being a key feature. The present study sought to assess variability in electrophysiological processes involved in stimulus evaluation as the primary neural generator underlying the cognitive variability reported in prior research. Identifying the underlying neural mechanisms of neuroticism has the potential to refine the current theoretical understanding of the construct and is crucial to avoid the pitfalls of its current method of self-assessment. To that aim, the current project has failed to provide support for variance in cognitive processes underlying stimulus evaluation as the primary mechanism driving the observed patterns of cognitive variability. Areas for future investigations include variability in early stage stimulus processing, as well as late stage response selection, as potential neural generators of the observed variability linked to high levels of neuroticism. While the present analysis failed to provide support for the primary hypothesis, exploratory analyses at site Pz yielded interesting results, which include significant relationships between the neuroticism subscale withdrawal and P3b latency, as well as greater interparticipant P3b latency variance in the high neuroticism group. These results provide discriminant validity for aspects in the BFAS and make unique contributions to the growing field of personality neuroscience. Future research, aiming to establish neurological indexes of liability for conversion to more severe clinical

phenotypes, should investigate links between diminished P3b amplitude and N's lower aspects, as prior research has found this parameter to be more sensitive to liability, as opposed to symptomology or lasting effects of psychopathology.

REFERENCES

Aldinger, M. Stopsack, M., Ulrich, I., Appel, K., Reinelt, E., Wolff, S., Grabe, H. J., Lang, S., Barnow, S. (2014). Neuroticism developmental courses – implications for depression, anxiety and everyday emotional experience; a prospective study from adolescence to young adulthood. *BMC Pyschiatry, 14,* Article 210.

Allen, T. A., & DeYoung, C. G. (2017). Personality neuroscience and the five factor model. In T. Widiger (Ed.), *The Oxford handbook of the five factor model* (Vol. 1, pp. 319–349). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199352487.013.26

Bagby, R., Young, L. T., Schuller, D. R., Bindseil, K. D., Cooke, R. G., Dickens, S. E., Levitt, A. J., Joffe, R. T. (1996). Bipolar disorder, unipolar depression and the Five-Factor Model of personality. *Journal of Affective Disorders*, *41*(1), 25-32. https://doi.org/10.1016/0165-0327(96)0060-2

Cid-Fernández, S., Lindín, M., & Díaz, F. (2019). The importance of age in the search for ERP biomarkers of aMCI. *Biological Psychology*, *142,* 108–115. https://doi.org/10.1016/j.biopsycho.2019.01.015

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24-30. https://doi.org/10.1016/S0168-5597(98)00033-1

Costa, P., & Mccrae, R. R. (1992). Neo PI-R professional manual. *Psychological Assessment Resources*, 223-250.

Cribbie, R., Bewell, C., Wilcox, R. Keselman, H. J. (2007). Tests for treatment group equality when data are nonnormal and heteroscedastic. *Journal of Modern Applied Statistical Methods, 6*(1), 117-132. https://doi.org/10.22237/jmasm/1177992660

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*, 880-896.

Dias, E. C., Butler, P. D., Hoptman, M. J., Javitt, D. C. (2011). Early sensory contributions to contextual encoding deficits in schizophrenia. *Archives of General Psychiatry*, *68*(7), 654-664. DOI: 10.1001/archgenpsychiatry.2011.17

Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, *41*(1), 417–440. https://doi.org/10.1146/annurev.ps.41.020190.002221

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(03), 342-427. https://doi.org/10.1017/S0140525X00058027

Donchin, E., Ritter, W., & Mccallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, & S. H. Koslow (Eds.), *Event-related potentials in man* (pp. 349–411). Academic Press.

Eysenck, H. J. (1958). A short questionnaire for the measurement of two dimensions of personality. *Journal of Applied Psychology*, *42*(1), 14–17. https://doi.org/10.1037/h0041738

Fjell, A. M., Walhovd, K. B., Meling, S., & Johansen, M. B. (2005). Basic information processing of neurotics and stables: and experimental ERP approach to personality and distractibility. *Scandinavian Journal of Psychology*, *46*(6). 493-502, https://doi.org/10.1111/j.1467-9450.2005.00481.x

Flehmig, H. C., Steinborn, M., Langner, R., & Westhoff, K. (2007). Neuroticism and the mental noise hypothesis: Relationships to lapses of attention and slips of action in everyday life. *Psychology Science*, *49*(4), 343–360.

Gavin, W. J., Lin, M., Davies, P. L. (2019). Developmental trends of performance monitoring measures in 7 to 25-year-olds: Unraveling the complex nature of brain measures. *Psychophysiology*, *56*(7), Article 13365. DOI: 10.1111/psyp.13365.

Goldberg, L. R. (1992). The Development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26–42.

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of Clinical Psychiatry*, *76*(2), 155–162. https://doi.org/10.4088/JCP.14m09298

Hamilton, H. K., Boos, A. K., & Mathalon, D. H. (2020). Electroencephalography and event -related potential biomarkers in individuals at clinical high risk for psychosis. *Biological Psychiatry*, *88*(4), 294–303. https://doi.org/10.1016/j.biopsych.2020.04.002

Hamilton, H. K., Roach, B. J., Bachman, P. M., Belger, A., Carrion, R. E., Duncan, E., Johannesen, J., Light, G. A., Niznikiewicz, M. A., Addington, J., Bearden, C. E., Cadenhead, K. S., Cornblatt, B. A., McGlashan, T. H., Perkins, D. O., Seidman, L. J., Suang, M. T., Walker, E. F., Woods, S. W., Cannon, T. D., Mathalon, D. H. (2019). Association between P300 responses to auditory oddball stimuli and clinical outcomes in psychosis risk syndrome. *JAMA Psychiatry, 76(*11), 1187-1197. DOI: 10.1001/jamapsychiatry.2019.2135.

Ila, A., & Polich, J. (1999). P300 and response time from a manual Stroop task. Clinical *Clinical Neurophysiology, 110,* 367-373. DOI:10.1016/S0168-5597(98)00053-7

Jeong, H., Yim, H. W., Lee, S.-Y., Lee, H. K., Potenza, M. N., Kwon, J.-H., Koo, H. J., Kweon, Y.-S., Bhang, S., & Choi, J.-S. (2018). Discordance between self-report and clinical diagnosis of internet gaming disorder in adolescents. *Scientific Reports*, *8*(1), Article 10084. https://doi.org/10.1038/s41598-018-28478-8

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). Personality and major depression: A Swedish longitudinal, population-based twin study. *Archives of General Psychiatry*, *63*(10), 1113–1120. https://doi.org/10.1001/archpsyc.63.10.1113

Klein, R. J., & Robinson, M. D. (2019). Neuroticism as mental noise: Evidence from a continuous tracking task. *Journal of Personality*, *87*(6), 1221–1233. https://doi.org/10.1111/jopy.12469

Kutas, M., McCarthy, G., Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science,197*(4305), 792-795. https://doi.org/10.1126/science.887923

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist, 64*(4), 241-256. DOI:10.1037/a0015309.

Luciano, M., Hagenaars, S. P., Davies, G., Hill, W. D., Clarke, T.-K., Shirali, M., Harris, S. E., Marioni, R. E., Liewald, D. C., Fawns-Ritchie, C., Adams, M. J., Howard, D. M., Lewis, C. M., Gale, C. R., McIntosh, A. M., & Deary, I. J. (2018). Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nature Genetics, 50*(1), 6–11. https://doi.org/10.1038/s41588-017-0013-8

Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five -factor model of personality and symptoms of clinical disorders: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment, 27*(2), 101–114. https://doi.org/10.1007/s10862-005-5384-y

Marchetti, I., Mor, N., Chiorri, C., & Koster, E. H. W. (2018) The brief state rumination inventory (BSRI): Validation and psychometric evaluation. *Cognitive Therapy and Research, 42,* 447-460. https://doi.org/10.1007/s10608-018-9901-1

Mathalon, D. H., Ford, J. M., Rosenbloom, M., & Pfefferbaum, A. (2000). P300 reduction and prolongation with illness duration in schizophrenia. *Biological Psychiatry, 47*(5), 413-427. https://doi.org/10.1016/S0006-3223(99)00151-1

O'Donnell, B. F., Faux, S. F., McCarley, R. W. (1995). O'Donnell, B. F., Faux, S. F., McMarley, R. W. Increased rate of P300 latency prolongation with age in schizophrenia. *Archives of General Psychiatry, 52*(7), 544-549, DOI:10.1001/archpsyc.1995.03950190026004

Oribe, N., Hirano, Y. Kanba, S., Re, E. C. D., Seidman, L. J., Mesholam-Gately, R., Spencer, K. M., McCarley, R. W., Niznikiewicz, M. A. (2013). Early and late stages of visual processing in individuals in prodromal and first episode schizophrenia: an ERP study. *Schizophrenia Research, 146*(1), 95-102. DOI: 10.1016/j.schres.2013.01.015.

Ouyang, G., Hildenbrandt, A., Sommer, W. Zhou, C. (2017). Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods. *Neuroscience and Biobehavioral* Reviews, *75,* 1-21. http://dx.doi.org/doi:10.1016/j.neubiorev.2017.01.023

Pfefferbaum, A., Wenegrat, B. G., Ford, J. M., Roth, W. T., & Kopell, B. S. (1984). Clinical application of the P3 component of event-related potentials. II. Dementia, depression and schizophrenia. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *59*(2), 104–124. https://doi.org/10.1016/0168-5597(84)90027-3

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4), 456–479. https://doi.org/10.1097/00004691-199210000-00002

Rector, N. A., Hood, K., Richter, M. A., & Bagby, R. M. (2002). Obsessive-compulsive disorder and the five-factor model of personality: Distinction and overlap with major depressive disorder. *Behaviour Research and Therapy*, 15.

Robinson, M. D., & Tamir, M. (2005). Neuroticism as mental noise: A relation between neuroticism and reaction time standard deviations. *Journal of Personality and Social Psychology*, *89*(1), 107–114. https://doi.org/10.1037/0022-3514.89.1.107

Rodgers, B. (1996). Reported parental behaviour and adult affective symptoms. 1. Associations and moderating factors. *Psychological Medicine*, *26*(1), 51–61. https://doi.org/10.1017/S0033291700033717

Rodgers, B., & Mann, S. A. (1986). The reliability and validity of PSE assessments by lay interviewers: A national population survey. *Psychological Medicine*, *16*(3), 689–700. https://doi.org/10.1017/S0033291700010436

Schmidtke, J. I., & Heller, W. (2004). Personality, affect and EEG: Predicting patterns of regional brain activity related to extraversion and neuroticism. *Personality and Individual Differences*, *36*(3), 717–732. https://doi.org/10.1016/S0191-8869(03)00129-6

Shipley, B. A., Weiss, A., Der, G., Taylor, M. D., & Deary, I. J. (2007). Neuroticism, extraversion, and mortality in the UK health and lifestyle survey: A 21-year prospective cohort study. *Psychosomatic Medicine*, *69*(9), 923–931. https://doi.org/10.1097/PSY.0b013e31815abf83

Stanzione, P., Fattapposta, F., Giunti, P., D'Alessio, C., Tagliati, M., Affricano, C., & Amabile, G. (1991). P300 variations in parkinsonian patients before and during dopaminergic monotherapy: A suggested dopamine component in P300. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *80*(5), 446–453. https://doi.org/10.1016/0168-5597(91)90093-D

Stelmack, R. M., Houlihan, M., & McGarry-Roberts, P. A. (1993). Personality, reaction time, and event-related potentials. *Journal of Personality and Social Psychology*, *65*(2), 399-409. https://doi.org/10.1037/0022-3514.65.2.399

Trull, T. J., Useda, J. D., Costa, P. T., & McCrae, R. R. (1995). Comparison of the MMPI-2 Personality Psychopathology Five (PSY-5), the NEO-PI, and the NEO-PI—R. *Psychological Assessment, 7*(4), 508–516. https://doi.org/10.1037/1040-3590.7.4.508

Tseng, J., & Poppenk, J. (2020). Brain meta-state transitions demarcate thoughts across task contexts exposing the mental noise of trait neuroticism. *Nature Communications*, *11*(1), Article 3480. https://doi.org/10.1038/s41467-020-17255-9

Van Os, J., & Jones, P. (2001). Neuroticism as a risk factor for schizophrenia. *Psychological Medicine, 31*(6), 1129–1134. https://doi.org/10.1017/S0033291701004044

Vorkapić, S. T., & Tadinac, M. (2016). The relationship of neuroticism, psychoticism, and depression with evoked potentials. *Acta Neuropsychologica, 14*(3), 241 – 254. https://doi.org/10.5604/17307503.1222838

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.

Wilson, R. S., Krueger, K. R., Gu, L., Bienias, J. L., Mendes de Leon, C. F., & Evans, D. A. (2005). Neuroticism, extraversion, and mortality in a defined population of older persons. *Psychosomatic Medicine, 67*(6), 841–845. https://doi.org/10.1097/01.psy.0000190615.20656.83

Wei, X., Ni, X., Liu, J., Lang, H., Zhao, R., Dai, T., Qin, W., Jia, W., Fang, P. (2020). Simulation study on the spatiotemporal difference of complex neurodynamics between P3a and P3b. *Hindwai,* Article 2796809.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering, 5,* 539-554.

APPENDIX A

IRB APPROVAL LETTER

Graduate Studies and Research
Marquette, MI 49855-5301
906-227-2300
www.nmu.edu/graduatestudies/

Memorandum

TO: Jon Barch
Psychological Sciences Department

CC: Jeremey Lawrence
Psychological Sciences Department

DATE: February 16, 2021

FROM: Lisa Schade Eckert
Dean of Graduate Studies and Research

SUBJECT: **IRB Proposal** HS20-1165
IRB Approval Date: 02/16/2021
Proposed Project Dates: 03/1/2021 – 05/1/2021
"The Mental Noise Hypothesis: A relation between Neuroticism and P3 latency variance in choice reaction time tasks"

Your proposal "The Mental Noise Hypothesis: A relation between Neuroticism and P3 latency variance in choice reaction time tasks" has been approved by the NMU Institutional Review Board. Include your proposal number (HS20-1165) on all research materials and on any correspondence regarding this project.

A. If a subject suffers an injury during research, or if there is an incident of non-compliance with IRB policies and procedures, you must take immediate action to assist the subject and notify the IRB chair (dereande@nmu.edu) and NMU's IRB administrator (leckert@nmu.edu) within 48 hours. Additionally, you must complete an Unanticipated Problem or Adverse Event Form for Research Involving Human Subjects.

B. Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding. Informed consent must continue throughout the project via a dialogue between the researcher and research participant.

C. If you find that modifications of investigators, methods, or procedures are necessary, you must submit a Project Modification Form for Research Involving Human Subjects before collecting data. Any changes or revisions to your approved research plan must be approved by the IRB prior to implementation.

APPENDIX B

PYTHON 3.10 ADAPTIVE FILTERING CODE

```
from scipy import signal
import scipy.stats as stats
import numpy as np
import glob
import pandas as pd


index = [str(i) for i in range(0, 65)]
laglist = []
laglist1 = []
latency = []
cl = []
cl1 = []
avg = np.zeros(400,)
template = np.zeros(400, )
template2 = np.zeros(400, )

#Lag Finder w/ Avg
def lag_finder_avg(y1, y2, sr, template):
    n = len(y2)
    corr = signal.correlate(y2, y1, mode='same') /
np.sqrt(signal.correlate(y1, y1, mode='same')[int(n/2)] *
signal.correlate(y2, y2, mode='same')[int(n/2)])
    cl.append(corr)
    delay_arr = np.linspace(-0.5*n/sr, 0.5*n/sr, n)
    delay = delay_arr[np.argmax(corr)]
    if np.amax(corr) > 0:
        laglist.append(delay)
        template += y1.shift(periods = int(delay * sr), fill_value = 0)

#Lag finder w/ Template
def lag_finder_it1(y1, y2, sr):
    global template2
    n = len(y2)
    corr = signal.correlate(y2, y1, mode='same') /
np.sqrt(signal.correlate(y1, y1, mode='same')[int(n/2)] *
signal.correlate(y2, y2, mode='same')[int(n/2)])
    cl1.append(corr)
    delay_arr = np.linspace(-0.5*n/sr, 0.5*n/sr, n)
    delay = delay_arr[np.argmax(corr)]

    if np.amax(corr) > 0:
        laglist1.append(delay)
        template2 += y1.shift(periods=int(delay * sr), fill_value=0)

        x = np.array(y1.shift(periods = int(delay * sr), fill_value = 0))
        z = (np.argmax(x)/sr) * 1000
        latency.append(z)
```

```
#Read Per Subject Per Stim Avg
AvgUnsubC = pd.read_csv('C:\AAvg\LT_02_1hp_ConStim, Ave LT_02.txt', header=
None, sep = '\t')
AvgUnsubI = pd.read_csv('C:\AAvg\LT_02_1hp_IncStim, Ave LT_02.txt', header=
None, sep = '\t')
AvgUnsubX = pd.read_csv('C:\AAvg\LT_02_1hp_XXXStim, Ave LT_02.txt',header=
None, sep = '\t')

#Label Channels
AvgUnsubC.columns = index
AvgUnsubI.columns = index
AvgUnsubX.columns = index

#Montage
AvgC = (AvgUnsubC["34"]+ AvgUnsubC["36"] + AvgUnsubC["37"] + AvgUnsubC["33"]
+ AvgUnsubC["38"] + AvgUnsubC["35"] + AvgUnsubC["39"])/7

AvgI = (AvgUnsubI["34"] + AvgUnsubI["36"] + AvgUnsubI["37"] + AvgUnsubI["33"]
+ AvgUnsubI["38"] + AvgUnsubI["35"] + AvgUnsubI["39"])/7

AvgX = (AvgUnsubX["34"]+ AvgUnsubX["36"] + AvgUnsubX["37"] + AvgUnsubX["33"]
+ AvgUnsubX["38"] + AvgUnsubX["35"] + AvgUnsubX["39"])/7

AvgStim = (AvgC + AvgI + AvgX) / 3
Avg34 = (AvgUnsubC["34"] + AvgUnsubI["34"] + AvgUnsubX["34"])

path = r'C:\ASingle'
all_files = glob.glob(path + "/*.txt")

for filename in all_files:
    df = pd.read_csv(filename, sep = '\t', header = None)
    df.columns = index
    dfs = (df["34"] + df["36"] + df["37"] + df["33"] + df["38"] + df["35"] +
df["39"]) / 7
    lag_finder_avg(dfs, AvgStim, 500, template)
#Clean before next iteration
template = template/len(laglist)
lagsd = stats.tstd(laglist)
print("lag sd w/ avg = ", lagsd)

for filename in all_files:
    df = pd.read_csv(filename, sep = '\t', header = None)
    df.columns = index
    dfs = (df["34"] + df["36"] + df["37"] + df["33"] + df["38"] + df["35"] +
df["39"]) / 7
    lag_finder_it1(dfs, template, 500)

#Clean before presentation
template2 = template2/len(laglist1)
lagsd1 = stats.tstd(laglist1)
print("lag sd w/ template = ", lagsd1)


cla = np.array(cl)
cla1 = np.array(cl1)

print("mean coeff w/ average = ", np.mean(cla), "mean coeff w/ template = ",
```

```
np.mean(cla1))
print("max coeff w/ average = ", np.amax(cla), "max coeff w/ template = ",
np.amax(cla1))
print("trials kept w/ template = ", len(laglist1), "trials kept w/ average =
", len(laglist))

lra = np.array(latency)
lra = lra + 200
lVar = stats.tvar(lra)
lSD = stats.tstd(lra)

print("latency sd = ", lSD)
print("latency mean = ", np.mean(lra))
```

APPENDIX C

R 4.1.3 ANALYSIS CODE

2 x 3 Mixed Designs ANOVA (Latency SD)

```
setwd("C:/RProj")

library(tidyverse)

library(ggpubr)

library(rstatix)

library(car)

data = read_csv(file = "data1.csv")

data$Group = factor(data$Group, levels = c(0, 1), labels = c("High", "Low"))

data = data %>% gather(key = "Stype", value = "lSD", Cong, In, X) %>% convert_as_factor(ID, Stype)

data$Stype = factor(data$Stype, levels = c("X", "Cong", "In"), labels = c("Control", "Congruent", "Incongruent"))

data %>%

  group_by(Stype, Group) %>%

  identify_outliers(lSD)

data %>%

  group_by(Stype, Group) %>%

  shapiro_test(lSD)

data %>%

  group_by(Stype) %>%

  levene_test(lSD ~ Group)

box_m(data[, "lSD", drop = FALSE], data$Group)

res.aov = anova_test(data = data, dv = lSD, wid = ID, between = Group, within = Stype)

res.aov
```

2 x 3 Mixed Designs ANOVA (Latency)

```
setwd("C:/RProj")

library(tidyverse)

library(ggpubr)

library(rstatix)

library(car)

Lat = read_csv(file = "L.csv")

Lat$Group = factor(Lat$Group, levels = c(0, 1), labels = c("High", "Low"))

Lat = Lat %>% gather(key = "Stype", value = "Latency", Cong, In, X) %>%
convert_as_factor(ID, Stype)

Lat$Stype = factor(Lat$Stype, levels = c("X", "Cong", "In"), labels = c("Control", "Congruent",
"Incongruent"))

Lat %>%

  group_by(Stype, Group) %>%

  identify_outliers(Latency)

Lat %>%

  group_by(Stype, Group) %>%

  shapiro_test(Latency)

Lat %>%

  group_by(Stype) %>%

  levene_test(Latency ~ Group)

box_m(Lat[, "Latency", drop = FALSE], Lat$Group)

res.aov2 = anova_test(data = Lat, dv = Latency, wid = ID, between = Group, within = Stype)

get_anova_table(res.aov2)
```

Welch’s ANOVA

```
setwd("C:/RProj")
library(tidyverse)
library(ggpubr)
library(rstatix)
dataE = read_csv(file = "TEa.csv")
dataE$Group = factor(dataE$Group, levels = c(0, 1), labels = c("High", "Low"))
dataE %>% sample_n_by(Group, size = 1)
levels(dataE$Group)
dataE %>%
  group_by(Group) %>%
  identify_outliers(Lat)
dataE %>% levene_test(Lat ~ Group)
res3.aov = dataE %>% welch_anova_test(Lat ~ Group)
res3.aov
get_anova_table(res3.aov)
```