UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

IS IRON STATUS AN ADDITIONAL BARRIER TO SMOKING CESSATION IN FEMALES? AN ERP ANALYSIS.

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

LAILI KHARAZI BOOZARY
Norman, Oklahoma
2022

IS IRON STATUS AN ADDITIONAL BARRIER TO SMOKING CESSATION IN FEMALES? AN ERP ANALYSIS.

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Michael J. Wenger, Chair

Dr. Lauren E. Ethridge

Dr. Dingjing Shi

Dr. Scott D. Gronlund

Dr. Lei Ding

Acknowledgements

My PhD took a lot of hard work, but even more support from my mentors and family. Thank you to my committee members, Dr. Michael Wenger, Dr. Lauren Ethridge, Dr. Dingjing Shi, Dr. Scott D. Gronlund, and Dr. Lei Ding, for their guidance in my work. I wouldn't be who I am today if it wasn't for Dr. Wenger's willingness to accept me into his lab and his mentorship along the way. Additionally, a huge thank you is due to Dr. Darla E. Kendzor for her role in providing me the opportunities to grow as a researcher and scientist, and for giving me vision past the PhD.

Of course, I need to mention Dr. Lisa De Stefano, who has been one of the best mentors and friends I've ever had. Lisa, without you, I never would have made it past the first week. You have been and continue to be a constant support I can turn to for brainstorming solutions to my scientific problems, but more importantly, pep talks and comradery throughout the process. Thank you, Lisa.

Mohammad jan, my husband, you are the one who made this all possible. You are the light that kept me going when I really wanted to quit. Thank you for showing me how good life can be, and for challenging me to dream big. Thank you for reminding me of my strength and resilience, and for keeping me going in the hard times. More importantly, thank you for all of our happy times together, which make the hard times so worth it.

Maman and baba, my parents. You've always worked hard so that I can actualize my full potential. Thank you for being my biggest cheerleaders since day 1. Thank you for always believing in me.

Table of Contents

List of Tables

List of Figures

Abstract

It has been well-established that females achieve lower cessation rates than males in traditional smoking cessation interventions. Research suggests that iron status variations (i.e. deficiency) are common in females. Iron status variations are known to alter dopamine (DA) pathways and learning ability involved in reward processing and addiction. This suggests a mediation model which may partially explain the sex differences in cessation rates wherein iron status predicts abstinence directly and indirectly via changes to reward processing. The current study enrolled 54 females at the onset of smoking cessation intervention and measured cognitive task behavior with concurrent electroencephalography (EEG). In addition, assays were run to extract iron biomarkers from a blood sample. ERP components known to be sensitive to reward learning were extracted from the EEG data. Task behavior was quantified in terms of accuracy and/or reaction time. Lastly, blink rates were extracted from the EEG to act as a proxy for dopaminergic status. After imputing missing data and using a Bayesian variable selection framework to select a final set of mediators, a Generalized Structured Component Analysis was employed to test three mediation models. Results revealed that iron status, task behavior, dopamine status, and neural components accounted for 1/3 of the variance in smoking cessation. These outcomes suggest that iron status may play a role in success or failure in achieving cessation during a quit attempt. These results may have implications for future smoking cessation intervention standards of care.

Chapter 1

1.1 Introduction

Smoking is the leading cause of preventable death in the United States [1], in spite of the $320 million from the NIH [2] and $85 million from the CDC [3] spent annually, and the roughly $6 billion spent to-date from the Tobacco Master Settlement Agreement [4] toward tobacco use prevention and control. Adults who continue to smoke despite these public health initiatives come from vulnerable populations who have difficulty quitting [1]. For example, it has been found that females are significantly less successful at smoking cessation when compared with males [5-8]. Notably, one study comparing a contingency management (CM; i.e. financial incentives for abstinence) intervention against a standard care (SC) intervention found a significant interaction effect such that sex moderated the relationship between treatment group and smoking cessation [9]. Specifically, while in SC males achieved abstinence at higher rates than females, in CM females achieved abstinence at higher rates than males.

While some research has attempted to explain the female disadvantage in smoking cessation by way of differences in factors such as self-efficacy [10], negative affect [11], or others [12; 13], finding sex differences in interventions that include financial incentives highlights differences related to reward processing. Although nicotine exposure alone is enough to impact reward processing via increased dopaminergic activity in structures such as the ventral tegmental area [VTA; 14], the results from CM trials suggest that female reward processing has changed beyond what is observed in males. These results suggest the need for exploring additional reward-related neural mechanisms which disproportionally affect females and which could lead

to differences in cessation. One such mechanism includes the neural changes resulting from variations in iron levels and associated variations in dopaminergic signaling. The WHO estimates that 11.8% of females in the US are anemic [15], while other research has found rates of iron deficiency without anemia (IDNA) to be 10-30% worldwide [16; 17]. In addition to having lower average iron levels compared to men, females also have more variations in iron across the lifespan [18].

Research suggests that disruption to iron homeostasis can lead to cognitive and neural dysregulation [19]. More specifically, studies show that depleted iron levels lead to lower dopamine transporter (DAT) binding [20], increased DA/metabolites [21], and decreased D1-like and D2-like receptors in the striatum and nucleus accumbens [NA; 22]. Erikson et. al. also observed that the changes to the D2-like receptors are markedly higher in comparison with D1-like receptors, with D2 receptor density significantly related to ferritin levels in the striatum [22]. These results correspond with a net decrease in downstream dopaminergic activity. The NA is an important structure role in the mesolimbic dopaminergic pathway and the striatum is noted for its role in receiving inputs from the cortex and substantia nigra pars compacta (SNc) and sending ensuing signals through the direct and indirect pathways of the basal ganglia [23]. The decrease of overall dopaminergic activity, particularly via D2-like receptors going into the indirect pathway of the basal ganglia, caused by reductions in iron may have significant implications for learning, and perhaps more marked effects on negative learning. This mechanism could explain why females are not able to achieve abstinence under normal circumstances but when provided with financial incentives for abstinence (i.e. heightened positive reward) they are significantly more likely to achieve abstinence.

Considering the high rates of iron disruption present in females, these findings demonstrate that the changes to DA availability and function provide one potential mechanism leading to the female disadvantage to smoking cessation.

1.2 Research question and significance

This leads to the primary research question of the current proposal: Do variations in iron pose an additional burden to a woman who is attempting to quit smoking? Answering this question will have notable impacts for the 14.1% of females [1] who continue to smoke in the United States. Tobacco use continues to cause nearly 7 million death and costs more than $300 billion to the US economy annually (including $170 billion in medical costs and $156 billion in lost productivity). The results of this effort could have long-term clinical implications for standard care smoking cessation interventions; recommendations could potentially include blood tests for iron biomarkers, making dietary recommendations, or introducing supplementation for women relatively low in iron even though they may not be iron deficient.

1.3 Definition of Terms

Abstinence: Abstinence will be defined via self-report along with levels of expired carbon monoxide (CO). The self-report is in response to the question, “Have you smoked, even a puff, during the last 7 days?” (yes/no). Expired CO will be acquired at any assessment which the participant attends in-person. Cutoff values for abstinence are defined as an expired CO level of ≤10 parts per million (ppm) for the quit date and ≤6 ppm for any subsequent follow-up assessment [24].

Addiction: The DSM-5 includes eleven criteria for the diagnosis of substance use disorder including: cravings and urges, withdrawal symptoms in absence of use,

continuing use despite adverse effects, and desire yet inability to quit [25]. Addiction goes beyond the chronic use or misuse of a drug to additionally involve unsuccessful quit attempts or an inability to stop use [26].

Basal Ganglia: The current study will use the term “basal ganglia” to refer to a group of sub-cortical nuclei that are implicated in processing of reward, motivation, learning, and movement, including but not limited to: the globus pallidus, striatum (caudate nucleus and putamen), and substantia nigra [SN; 23], nucleus accumbens [27; 28], ventral tegmental area (VTA) [29], and the subthalamic nucleus (STN) [30].

Contingency Management: Providing small, financial incentives for biochemically-verified smoking abstinence, usually alongside standard care [9; 31]. This type of intervention has significant implications for reward-processing due to the financial gain component.

Hemoglobin (Hb): Hb is the oxygen transport pigment of the red blood cells [32], and is the only marker of anemia. Interestingly, Hb is altered in smokers depending on the amount smoked; [3-7 g/L; 32; 33].

Iron: This report distinguishes between iron deficiency with anemia (IDA), iron deficiency without anemia (IDNA). The term “disruption to iron homeostasis” or similar will refer to any level of iron deficiency or toxicity which deviates from normative iron sufficiency. Notable biomarkers include ferritin (sFt), hemoglobin (Hb), and serum transferrin receptor (sTfR). Iron sufficiency is a state in which Hb is ≥12 g/dL and sFt is ≥15 µg/L [32; 34].

Iron Deficiency with anemia (IDA): IDA is a state in which Hb <12 g/dL and sFt <15 µg/L [32; 34].

Iron Deficiency without anemia (IDNA): IDNA is a state in which Hb ≥12 g/dL accompanied by an sFt <15 µg/L [32; 34].

National Health and Nutrition Examination Survey (NHANES) Age-, Sex-, and Race-Adjusted Ferritin Percentile: Ferritin follows a non-normal distribution which creates limitations for using raw ferritin in analyses. More importantly, due to the fact that sFt levels change with age and vary by race, NHANES data on sFt were used to estimate the cumulative density function on levels of sFt which was then used to determine an age- and race-adjusted percentile for each participant's measured level of sFt.

Reward-prediction error: A mechanism of learning is the Reward-Prediction Error, which is equal to the given reward minus the expected reward [35]. At onset of learning, when the received reward is higher than the expected reward, there will be an increase in phasic DA activity as mediated by the D1-like receptors of the direct pathway to code for learning from positive feedback. In the case that the received reward is lower than the expected reward, there will be a suppression of phasic DA activity throughout the basal ganglia as mediated by D2-like receptors to code for learning from negative feedback. As learning continues to take place, there will be less difference between reward attained and reward expected, and thus there will be a decrease in magnitude of dopaminergic activity.

Serum Ferritin (sFt): sFt reflects storage iron contained in the liver [32] and thus is a common measure of iron stores [34]. Iron deficiency first affects sFt before any changes are reflected in Hb. In females sFt cutoff values reflected for ages 20-44, 45-64, and 65-74 to indicate overload are 150, 200, and 300 µg/L [32], while the most recent report from the WHO represents normal values between 15-150 µg/L for healthy menstruating

females of all ages and 15-200 µg/L for non-menstruating females of all ages [34]. Note that median values across all races taken from NHANES data reflect that sFt increases throughout the female lifespan, particularly after menopause, and is lowest at ages 14-18 (i.e. onset of menstruation), pregnancy, and lactation [32].

Serum Transferrin Receptor (sTfR): sTfR is responsible for transferrin uptake into cell bodies and for transport of iron into the brain. The WHO [34] states that the measurement of sTfR reflects the iron supply that is taken into tissues (as opposed to circulating). As ID and IDA develop, sTfR levels increase [32].

Standard Care: This report will use the term “Standard Care” (SC) to refer to standard smoking cessation treatments. The current recommendations for standard care includes the use of counseling of various modes, nicotine replacement therapies (NRTs; i.e. nicotine patch, lozenge, etc.), and other pharmacotherapies such as bupropion or varenicline [36].

Total Body Iron (TBI): TBI is calculated by taking the ratio between sTfR and sFt; the exact equation is as follows: TBI = -[log(sTfR/sFt ratio) - 2.8229]/0.1207 [37].This measure accounts for both short-term decrease in sFt as well as the longer-term upregulation of sTfR.

1.4 Summary

In sum, females’ lack of success in smoking cessation poses a significant public health disparity. As such, it is important to identify other potential barriers to success. One such plausible mechanism exists in dysfunction of reward-processing due to variations in iron status which could lead to decreased rates of cessation and increased response toward financial incentives when provided.

1.5 Assumptions/Limitations

The analysis presented here uses an intent-to-treat outcome (i.e. abstinence). As noted earlier, abstinence is determined based upon self-report and expired CO. In cases for which self-report and CO or CO alone are missing, participants are assumed to be smoking. This is a notable limitation in the field of abstinence outcomes.

A notable limitation of the present analysis includes the indirect measurement of dopamine. The present analysis will utilize blink rates and behavioral data from the Probabilistic Selection Task, both hypothesized correlates of dopamine levels [38; 39].

Additionally, there are generally low rates of smoking cessation outcomes at all time points and this will act as a significant limitation for study results. As noted in previous research, relapse is one of the most common outcomes when attempting to quit tobacco use [40].

Lastly, the current study will utilize data from female subjects, lacks a clear experimental vs. comparison group, and does not screen or correct for iron status. Thus, a necessary limitation in the present analysis is that it employs a non-experimental, observational design.

Chapter 2

2.1 Nicotine addiction: public health problems/cost

Lung cancer, which is primarily caused by cigarette smoking, is the leading cause of cancer death in the U.S. [41]. Despite prevalence reaching an all-time low reflecting a decline in smoking of two-thirds since 1964 [i.e. the first Surgeon General report on Smoking and Health; 42], today 34.2 million (13.7%) of U.S. adults continue to smoke tobacco products. As tobacco use causes more than 7 million deaths per year it is unsurprising that it is still the leading cause of preventable death [1]. Furthermore, tobacco use is responsible for causing 20% of all cancers; 30% of all cancer deaths are caused by smoking [41].

The costs and public health burden associated with tobacco use has led to the development of empirical and evidence-based smoking cessation interventions. Standard smoking cessation interventions have proven effective for many populations [36]. However, these programs have also led to the development of disparities in cessation rates [42]. Racial minorities, homeless adults, low-SES individuals, and females have lower rates of smoking cessation compared to the general population's rate of cessation from interventions [1].

2.2 Sex differences in cessation

It has been well-established that females are significantly less successful at smoking cessation when compared to males [5-8]. Research has attempted to explain this difference in cessation in a variety of ways. For example, some research has focused on behavioral and emotion-regulation issues such as self-efficacy or negative affect [10; 11], finding that females experiencing low self-efficacy or negative affect will

relapse at higher rates than males. Other research has suggested that differences in gonadal hormones may play a role in cessation [43]. Additionally, many researchers focus on sex differences in the hypothalamic-pituitary-adrenal (HPA) axis in response to cessation, specifically related to stress and withdrawal symptoms [44].

Contingency management interventions (CM; i.e. financial incentives for abstinence) provide a unique insight into sex differences due to the unique responses to money as a reward. Kendzor et. al. compared a CM intervention against a standard care (SC) intervention [9]. The financial incentive intervention involved incrementally increasing gift cards for time points at which participants achieved biochemically-verified abstinence; if participants relapsed, financial incentives would be reset to the base value. The primary finding of this study showed that overall, participants given financial incentives had significantly higher quit rates than the standard care group. More interestingly, there was a significant sex by group interaction effect such that in SC, females achieved abstinence at lower rates than males, while in CM, females achieved abstinence at higher rates than males. This interaction effect suggests that there are differences in reward processing between male and female and female smokers in response to money.

2.3 Implications for understanding reward processing

Finding sex differences in interventions that include financial incentives highlights potential sex differences related to reward processing. Although nicotine exposure alone is enough to impact reward processing by way of increased dopaminergic activity in structures such as the ventral tegmental area [VTA; 14], the results from CM trials suggest that female reward processing has changed *beyond* what is observed in males.

These results suggest the need for exploring additional reward-related neural mechanisms which disproportionally affect females and which could lead to differences in cessation. One possible mechanism is the neural changes resulting from variations in iron levels and associated variations in dopaminergic signaling. The WHO estimates that 11.8% of females in the US are anemic [15], while other research has found rates of iron deficiency without anemia (IDNA) to be 10-30% worldwide [16; 17]. These rates of ID in females are higher compared to rates of ID in adult males, which are found to be as low as 2% [45]. In addition to having lower average iron levels compared to men, females also have more variations in iron across the lifespan [18].

2.4 Iron in the brain

Iron plays four critical roles in the brain including oxygen transport, myelination, neurogenesis, and notably, iron's functions in neurotransmitter synthesis and regulation [46]. There are notable implications of early life iron deficiency on neural functions [47]. However, neural changes and/or dysfunction as a result of iron deficiency can occur at any time during a woman's reproductive years [48; 49]. Because of the importance of maintaining iron homeostasis, there are tightly regulated systems for the uptake of iron from the bloodstream into the brain [19].

Iron is primarily contained in subcortical structures. White matter in the brain has some of the highest iron concentrations, and studies show that high concentrations of iron exist in the basal ganglia [50; 51]. More specifically, the globus pallidus, caudate nucleus, putamen, and substantia nigra are those identified as having the highest concentrations of iron [50], and it has been established that the globus pallidus contains a higher concentration of the brain's iron [51] than other structures. The brain has about

1/3 of the ferritin found in the liver, and about 1/3 of the iron present in brain structures is present as ferritin [51].

This is functionally important for many reasons. Because of iron's significant roles noted above, and because of its high distribution in brain structures implicated in reward, learning, and memory, any disruptions to iron homeostasis will have functional implications for reward processing and learning.

2.4.1 ID effect on dopamine in the basal ganglia

The structures of the basal ganglia---those that are differentially dependent on iron---are implicated heavily in reward, learning, and memory, as well as movement. Both the direct and indirect pathways of the basal ganglia (see Figure 1) receive input from the cortex and the substantia nigra pars compacta (SNc) into the striatum. The direct pathway is thought to result in excitation going to the cortex (i.e. beginning or continuing a behavior), while the indirect pathway is thought to result in inhibition going to the cortex (i.e. termination of a behavior). The direct and indirect pathways collaborate to result in the net cortical output [28].

The current view of reward processing in the basal ganglia is the reward prediction error hypothesis [RPE; 29; 35]. When reward received is more than that which is expected, there will be a rapid increase in phasic firing of dopamine, leading to a feeling of pleasure. When reward received is less than that which is expected, dopaminergic activity will be suppressed leading to a feeling of lack of pleasure. Because the direct pathway is in charge of "starting" or "doing" behaviors, it is thought that positive reward (i.e. pleasure) is mediated through the direct pathway of the basal ganglia. On the other hand, because the indirect pathway is in charge of "stopping" or

“not doing”, it is thought that negative reward (i.e. lack of pleasure) is mediated through the indirect pathway.

There are five dopaminergic receptors (D1, D2, D3, D4, and D5) which are sub-classified into D1-like receptors (D1, D5; D1-Rs) and D2-like receptors [D2, D3, D4; D2-Rs; 52]. These receptors input into the basal ganglia and act on cyclic adenosine monophosphate (cAMP) which then activates protein kinase-A (PKA) which then executes a reaction; in this case, excitatory signals through the downstream basal ganglia. D1-like receptors are primarily responsible for inputs into the direct pathway, and they modulate reward/reinforcement by initiating cAMP and thus PKA production when they are stimulated; this will then stimulate more excitatory signals downstream through the pathway [52]. On the other hand, D2-like receptors, primarily responsible for inputs into the indirect pathway and modulating aversion [28], do not result in formation of cAMP/PKA and so result in downstream inhibition. In other words, increased activity of the direct pathway (i.e. D1-like receptors) will correspond to integrating positive feedback (i.e. positive learning), while increased activity of the indirect pathway (i.e. D2-like receptors) will correspond to integrating negative feedback (i.e. negative learning).

2.5 Iron deficiency effects on DA in reward-related structures

There are data showing that ID results in changes to dopaminergic systems that have implications for positive and negative learning. Burhans et. al. [20] fed rats iron deficient or iron sufficient diets from weaning through adulthood. Thin sections of rat brain were isolated, and binding capacity of the dopamine transporter (DAT) was measured. These researchers found that iron deficient rats had significantly lower DAT binding in the striatum and the substantia nigra. Decreased DAT binding indicates that

there is less DA activity because the transporter is more active in the reuptake of DA. Thus, DA spends less time in the synaptic cleft and the overall synaptic transmission of DA is reduced. Notably, these results implicate the striatum and the SN, which are two early structures in the basal ganglia pathways; these results will have an impact on downstream basal ganglia activity which will then affect learning both from positive and negative feedback.

Erickson et. al. [22] similarly fed female and male rats iron deficient or iron sufficient diet from weaning for 6 weeks. Sections and regions of rat brain were isolated, and D1-R and D2-R concentrations in a number of brain regions were compared. These researchers found that ID rats had fewer D1-Rs in the striatum in both males and females, while only males had decreased D1-Rs in the NA. Interestingly, there were small increases in D1-Rs in the PFC and VTA. More notably, this study also found a decrease of D2-Rs in the striatum and changes to in the NA depending on sex (i.e. females had significant increases of D2-Rs, while males had significant decreases of D2Rs) in the iron deficiency condition. Decreases in the D1-Rs and D2-Rs in the striatum would indicate overall less activity in both indirect and direct pathways. However, the researchers additionally found a significantly positive correlation between striatal iron content and D2-R density ($r = 0.91$), indicating that the magnitude of change to the D2 receptors was markedly higher than the change to D1-Rs, as a function of iron content. These results suggest that iron deficiency may lead to decreased learning from negative environmental feedback, while learning from positive environmental feedback may remain unchanged.

2.6 Iron deficiency effect on learning and addiction

ID is associated with deficits to attention, memory, and learning in humans. Falkingham et. al. [53] published a systemic review and meta-analysis on the effects of supplementing adults with iron, and concluded that iron supplementation can improve behavioral outcomes. Though these results are positive, Beard et. al.'s foundational report analyzing iron concentrations in multiple brain regions in rats with iron deficiency or sufficiency showed that despite iron repletion, the group of rats made ID at young age still had 20% lower brain iron concentrations compared to control rats after a period of iron repletion [50]. This suggests that there are neural deficits that cannot be fully restored despite repletion, but that functional deficits may be rescued.

Studies show behavioral and/or cognitive deficits as a function of iron deficiency. Murray-Kolb and Beard [54] compared human females on baseline iron status along a range of cognitive domains. In this analysis, participants with IDA had lower accuracy and slower speed in tasks related to attention, memory, and learning. Furthermore, Blanton et. al. [55] enrolled healthy, non-anemic females, assessed iron levels with Fe and sTfR, and assessed cognition with the Tower of London task. This task requires participants to move an object from an initial position to a goal position within a defined number of moves (2, 3, 4, or 5 moves), which required participants to plan the sequence before beginning execution. Time of task execution was measured using a stopwatch. Results from this study showed that Fe was significantly negatively related to time to complete the task such that lower iron levels corresponded to more time to plan, particularly whenever more moves were defined (i.e. 4 and 5 moves). Research protocols involving rats corroborate these findings. Jenney et. al. [56] compared iron deficient rats against control rats in a learning paradigm. In this protocol, rats were

provided with an active lever supplying cocaine and an inactive lever supplying no reward. Results from this study showed that ID rats did not differ in their responses to rewarded and non-rewarded lever responses, but their non-rewarded lever responses were significantly more frequent compared to the IS rats. These results illustrate a dysfunction in negative learning in the ID rats.

Recent work by Baker et al. [57] provides evidence that learning, particularly changes to positive and negative learning, could have implications for abstinence and relapse in smoking cessation. These researchers enrolled current smokers and asked them to complete the probabilistic selection task (PST), a task noted for sensitivity to reinforcement learning from both positive and negative learning, under three smoking conditions: smoking as normal, 24-hour period of abstinence, and resuming smoking after the abstinent period. The researchers found that when participants were abstinent, positive learning decreased while negative learning increased. In a period of relapse, the opposite happens: positive learning increased and negative learning decreased. This would indicate that whenever participants attempt to quit smoking, the negative symptoms are heeded while positive symptoms are discounted – and the opposite is true whenever participants relapse. Considering these results in the context of the data about ID, this may suggest that ID subjects would result in disruptions to learning during abstinence attempts which would make it even harder to quit smoking.

2.7 Low iron related to cognition

While research analyzing iron's relationship to cognitive and behavioral outcomes is frequently limited to participants with iron deficiency, there is some evidence to suggest that low iron corresponds to poor cognitive/behavioral outcomes.

Scott and Murray-Kolb [58] analyzed the relationship between iron and cognition on a continuous scale in human participants. Participants completed varying tasks including Attentional Network Task (ANT), Go/No Go Task (GNG), and Sternberg memory search (SMS); biomarkers of iron gathered included Fe, Hgb, sTfR, and TBI. Significant associations were found between accuracy and reaction time on multiple tasks and various biomarkers, which indicates that iron is significantly related to cognition at various levels (i.e. including before ferritin reaches deficiency).

2.8 Restatement of the central hypothesis

The findings presented in this chapter provide evidence that ID leads to disruptions in dopaminergic signaling, and may have more significant impacts on D2-Rs. Changes to D2-Rs suggests that negative learning will be differentially affected. Additionally, there is evidence that reliance on negative learning increases during a smoking cessation quit attempt. Compromised negative learning could represent an additional challenge to successful smoking cessation by females who are low in iron.

Chapter 3

3.1 Data Collection, Setting, Timeline

The OUVNL established a collaboration with the Tobacco Settlement Endowment Trust (TSET) Health Promotion Research Center (HPRC; https://healthpromotionresearch.org) of OU Health Sciences Center. Their clinic offers a free cessation intervention program called the Tobacco Treatment Research Program (TTRP; https://healthpromotionresearch.org/Free-Help-to-Quit) from which the current protocol will be recruiting participants. Recruitment began in December 2018 and ended in October 2021.

When participants are enrolled for the TTRP, they are scheduled for a baseline appointment (referred to as “TTRP baseline”). When they come to the TTRP baseline appointment, participants are scheduled for a target quit date and for visits at 1-4, 12, and 26 weeks post-quit date. At TTRP baseline, participants fill out a number of surveys related to their sociodemographic characteristics and tobacco use and history. At this appointment, clinical staff screen and enroll participants to optional additional research studies, including the current protocol. Participants will be scheduled for their first appointment of the current protocol (i.e. “the appointment” or similar) after the TTRP baseline and before their TTRP quit date.

3.2 Research Design, hypotheses

At their appointment, participants will be asked to complete surveys related to their dietary habits, physical activity, socioeconomic status, and medical and menstrual health. Upon completion, they will be fit for an electroencephalogram (EEG) cap system. After EEG cap system configuration, participants will be fit into a chair and

chinrest (both adjustable for height) approximately 62 centimeters away from the computer monitor with which they will be interacting. Upon completion of the task protocol, participants will be asked to complete a quick blood draw and will then be provided with a gift-card.

Participants were initially provided $50 gift card at the conclusion of each visit. With the onset of COVID-19, participants were facing more difficulty with in-person participation and as such, gift cards were increased to $100 at baseline and $150 at the two follow-up appointments.

3.3 Participants

Participants are eligible to participate if they are 19-70 years old, are not pregnant or lactating, have normal or corrected-to-normal vision, have no neurological injuries or deficits, and have a BMI between 18-35. Additionally, anybody who has received treatment for cancer within the past 6 months will be ineligible. Lastly, participants with a diagnosis of chronic obstructive pulmonary disease (COPD) will be ineligible due to the movement and noise related to excessive coughing during EEG sessions.

3.4 Cognitive Tasks

All cognitive tasks will be programmed using Psychtoolbox-3 [59-61] running in Matlab R2017b [62]. The tasks will be presented on a Windows computer with a Dell monitor (model# 1907FPt) which has a 19-inch display size with a resolution of 1280 x 1024 pixels and a 75 Hz refresh rate.

3.4.1 Resting State

EEG and video will be recorded during a 5-minute resting state protocol. Participants will be instructed to be still and silent, keep their chin in the chinrest and eyes open for a period of 5 minutes. Eye video recording (via the laptop's webcam) and EEG recordings will be initiated and terminated simultaneously. The resting state data will be collected for the analysis of spontaneous blink rate. The eye video will be recorded in order to provide ability to assess blink rate manually in cases of EEG file error or otherwise. After the resting state protocol is completed, the following four tasks will be administered in random order per participant.

3.4.2 Probabilistic Selection Task (PST)

The PST assesses positive and negative learning [39; 57]. Previous research has established that the PST is sensitive to dopamine status [30; 39; 63; 64]. The task has two phases: training and testing. Participants are shown two symbols (black Japanese Kanji characters on a gray background) that appear on the screen. During the training phase, participants are shown three stimulus pairs (denoted AB, CD, and EF) which have different probabilities of being correct. The AB pair is 80% correct for A and 20% correct for B. The participant should learn to approach stimulus A and avoid stimulus B. The CD pair is 70% correct for C and 30% correct for D. The EF pair is 60% correct for E and 40% correct for F. At the beginning of each trial, participants are shown a square in the center of the screen and must press the space bar to see the stimulus pair. Upon pressing the space bar, they are shown a fixation cross for a random interval between 500-1500 ms. When the pair is shown on the screen, the participant will select the right stimulus by pressing the letter "M" or the left stimulus by

pressing the letter “C.” The assignment of character to side of the display will be determined randomly on each trial. After a random interval of 300-500 ms, participants are shown feedback about their choice which will read “Correct!” or “Incorrect!” for a period of 2000 ms.

The training phase will be made up of up to three blocks of 60 trials each. In each block there will be 20 trials for each stimulus pair. Ideally, participants will only move to the testing phase after they reach an accuracy criteria defined as 65% correct A selections, 60% correct C selections, and 50% correct E selections [63]. Participants will move to the test phase after the block in which they reach these criteria, or until 3 blocks of 60 trials are completed (i.e. 180 trials).

There is one final block for the testing phase, consisting of 60 trials. In this block, feedback is no longer provided after each response. Additionally, each trial will consist of novel stimulus pairs exclusively (i.e. AC, AD, etc.). Accuracy and latency of responses during this block will indicate the extent of positive learning (i.e. approach A) and negative learning (i.e. avoid B).

3.4.3 Iowa Gambling Task (IGT)

The IGT assesses the extent of learning with uncertainty [65]. Participants will press a button (space bar) to initiate each trial. Participants will then be presented with a fixation cross in the middle of the screen for a random duration of 500-1500ms, after which the stimulus is shown. The stimulus for every trial is four decks of cards, arranged symmetrically around the center of the screen. The participant must select which deck to play on each trial. Feedback about their choice is presented after a random delay of 300-500ms in the form of monetary gain or loss. The four decks represent one of four

conditions: four high gains + 1 higher loss; four low gains + 1 higher loss; four high loss + 1 higher gain; and four low loss + 1 higher gain; conditions were randomized to decks per participant. Participants will not be instructed on the deck conditions at the start of the task, rather they must use the feedback provided to optimize their gains and minimize their losses. Participants start the task with $2,000, and they will be additionally incentivized by being provided with one piece of fun-sized candy per $100 over the initial amount that they end with.

3.4.4 Serial Reaction Time (SRT)

Beyond providing reaction time data, the SRT assesses implicit learning [66]. In this task, there is a fixation cross in the center of the screen with two blue vertical lines to the left and two blue vertical lines to the right of the cross. Each trial consists of one line turning red, at which point participants indicate which line is red by pressing a button on the keyboard corresponding to that line. The next trial will not initiate until the correct response is provided. Participants use the four fingers of their dominant hand to provide responses; right-handed participants use (n m , .) while left-handed participants use (z x c v).

The task has 456 trials in total, consisting of 38 blocks of 12-trial sequences. There is one repeating sequence (i.e. 1-4-2-3-4-4-3-1-2-3-4-2) which is presented interleaved with random sequences (i.e. randomly generated each time). Of the 38-blocks, 30 blocks are of the repeating sequence and 8 blocks are of the random sequence. The order of blocks is as such: two random blocks, then five repeating and one random block repeat until the 38 blocks are completed.

3.4.5 Paired Associates Learning (PAL)

The PAL cues into episodic memory, utilizing visual patterns (i.e. sketched drawings of objects) at various locations around the screen [67]. With varying levels of difficulty corresponding to the number of visual patterns to remember the location of (lesser value of patterns is easier). Participants will be shown either four or six objects at one of six specified locations around the screen. Once all objects are shown, participants must press a button to initiate the recall portion of the task. At this point, participants will be shown each of the objects in the center of the screen; they will then press a button on the keyboard corresponding to the location of the screen in which the object was originally shown. They will not be provided with feedback about their responses. There are 12 blocks in total, six blocks with six objects and six blocks with four objects.

3.5 Blood measures:

At the end of the appointment, participants will undergo a simple blood draw. The blood draw will either occur on-site by IRB-approved personnel or at the OU Children's Hospital phlebotomy lab by staff phlebotomists. In the case of on-site blood draws, a serum separator tube and a purple EDTA tube will be collected in that order. Both tubes will be inverted gently 5-10 times and serum separator tubes will be set aside until a clot has formed. Serum separator tubes will then be centrifuged at 1200 x g for 10 minutes until they are separated into whole blood components, then both tubes will be transported in an insulated bag with an ice pack to OU Children's hospital for assays. Phlebotomists at OU Children's will follow similar instructions.

The assays which will be run on the blood sample include indices of iron status: ferritin, complete blood count (including hemoglobin, hematocrit, etc.), and serum transferrin receptor. Additionally, c-reactive protein (CRP) will be assayed; CRP is an indicator of inflammation, and inflammation corresponds to artificially upregulated ferritin; therefore, this assay may be needed to correct ferritin levels if inflammation is detected.

3.6 EEG data collection and preparation

The electroencephalogram (EEG) will be digitized at 500 Hz, high-pass filtered at 0.5 Hz, low-pass filtered at 100 Hz, and notch filtered between 58-62 Hz. The data will be recorded with BrainVision Recorder version 1.21.0303 [68] with a gel-based 32-channel BrainVision actiCAP [69] compatible with the actiCHamp [70] amplifier [71; 72]. Impedances will be checked at the start of recording and between each task; ideally, impedances will be ≤30 kΩ throughout the recording session but instances of exception were made. EEG data will be processed using EEGLAB 2021.1 [73] for Matlab R2020a [74].

Data will be plotted to visually identify artifacts and noise such as coughs, muscle movements, and occasional ocular events. Time points during which these unwanted artifacts exist will be removed. At this stage unacceptable channels will also be visually identified and interpolated. Next, independent component analysis will be applied to the data [73]. Once this step completed, the components identified will be plotted. Those with more noise than signal or that correspond to identifiable artifacts (e.g., eye movements, muscle, cardiac activity) will be removed while the components with more signal than noise will be retained. Once these steps have been completed and a final

series of checks have been done (epoch the data, plot ERP images, plot the final continuous EEG) to determine that the final data is sufficiently cleaned and retains sufficient enough brain data, the data will be prepared for further analysis.

3.7 Data Analysis Plan

To test the hypothesized pathways through which iron deficiency affects cessation, the final analysis plan will be based on three hypothesized mediation models: mediation by dopaminergic variations, by learning variations, and by brain dynamics. See Figures 2-4 for representations of the proposed mediation models. Because the focal interest of the current study is to determine the predicted direct and indirect effects of iron on abstinence, the causal variables will be the iron status biomarkers and the outcome variable will be smoking abstinence in all three models. The three models differ only in the suggested pathway through which iron's effects on abstinence are mediated. Note that the three models are not intended to be mutually exclusive hypotheses. Rather, they each serve as heuristics for understanding the direct and indirect effects of variations in iron status on abstinence.

3.7.1 Mediation Analyses

Mediation analyses will be conducted to test the effect of either dopaminergic activity, learning, or brain dynamics as mediators of the relationship between iron level and smoking cessation. All mediation analyses will be fit in SAS 9.4 [75] using the maximum likelihood estimator. In all models, two parameters will be estimated: the direct effect (i.e. the variance in smoking cessation explained by iron level not mediated by any of the three) and the indirect effect (i.e. the variance in smoking cessation explained by iron level mediated by the three respective constructs). As standard, the

total effect will be computed as the sum of the direct and indirect effects, and the proportion of the mediation effects will be computed as (indirect effects/total effect)*100.

3.7.2 Model 1

The first model proposes DA as the mediating path (see Figure 2). To test this, the following indirect measures of dopamine will be utilized: spontaneous blink rate (BR), task-related BR, and PST behavioral data (i.e. response frequencies and latencies for choose A/avoid B and high and low conflict choices). It is suggested that BR serves as an indirect measure of dopamine status [39]. Previous research has utilized spontaneous BR [39], however, the present analysis will additionally explore the utility of task-related BR (i.e. BR during cognitive engagement in all four tasks). BR will be extracted in one of two ways. The primary method will be to extract the blinks from the continuous EEG data using the BLINKER plugin [76]. However, there are some rare instances for which the BLINKER plugin does not work (i.e. low-amplitude blink signal, etc.); for these instances, and as a secondary method, the eye video recording will be reviewed by study staff and blinks will be manually recorded. In addition to utilizing BR for measures of DA, it has been shown that sensitivity to positive/negative learning behavior on the PST is sensitive to DA [30; 39; 64]. For this reason, the accuracy and latency from choice behavior during testing phase of the PST will be utilized as mediators.

3.7.3 Model 2

The second model proposes learning outcomes as the mediating path. Learning will be represented by the behavioral data from the four cognitive tasks included in the protocol.

3.7.3.1 PST

Evans and Hampson demonstrated that when the training phase of the PST expires after a set number of trials regardless of reaching appropriate criteria, only 78% of participants reached criteria [77]. Notably, their protocol defined test criteria only with respect to the threshold for AB test pairs (i.e. choose A in ≥70% of AB trials during training), and maximum training phase trials were 480 (compared to 180 in the current study). Reaching criteria is indicative of adequate learning to prefer A over all other stimuli and prefer any other stimuli over B (i.e. positive and negative learning). For this reason, having reached criterion (yes vs. no) will be an important marker of learning to be used in the mediation models. Accuracy will be defined as choosing A and avoiding B in respective testing phase trials [30]. Reaction times of these trials will be equally important in accessing relative brain processing in contrasting these trial types.

Frank et. al. [30] demonstrated that dopamine dysfunction characterized by Parkinsonism conditions (including on/off medication and on/off deep brain stimulation) led to changes in reaction time in a conflict-modulated manner. High-conflict trials are defined as test trials in which the two choice stimuli have similar reinforcement values (i.e. AC = 90% and 80%). Low-conflict trials are defined as test trials in which the two choice stimuli have “easily discriminable values” (i.e. BD = 20% and 30%). Frank et. al. found that when receiving deep brain stimulation, participants RTs for high-conflict trials was lower than RTs low-conflict trials, while the normal pattern of action would be the opposite (i.e. higher RT in high-conflict trials compared with low-conflict trials). The increased reaction times corresponded to higher frequency of error trials. Consequently,

both choice frequencies and latencies in the high and low conflict trials will serve as another indicator of learning in the mediation analyses.

3.7.3.2 IGT

The four deck conditions represent advantageous (four high loss + 1 higher gain; four low loss + 1 higher gain) and disadvantageous (four high gains + 1 higher loss; four low gains + 1 higher loss) conditions. Mapelli et. al. [78] analyzed differences in learning characterized by frequency of selecting an advantageous deck minus frequency of selecting a disadvantageous deck in each 20-trial block (10 blocks for the current protocol); a positive trajectory over the blocks would indicate learning is taking place while a neutral or negative slope would indicate the opposite. These researchers found that most participants with Parkinson's preferred disadvantageous decks, while most control participants preferred advantageous decks; they additionally found that learning (i.e. trajectory over time) occurred slower in Parkinson's patients compared with controls. Thus, it will be important to use both indices (learning overall and trajectory of learning over time) in our mediation models.

3.7.3.3 SRT

Robertson [66] provides a detailed view into analyzing behavioral data from the SRT. If implicit learning has taken place, it is expected that the RT of the last repeating block trials will be faster (i.e. less than) the RT of the random block. To analyze this, we will take the median RT on the last random block minus the median RT on the last repeating block. Positive values will indicate that implicit learning has taken place while negative values will indicate that there is disruption to implicit learning. The magnitude of the learning or disruption will be indicated by the value of RT change. Lastly, another

way to categorize implicit learning is to analyze the trajectory of median RT over each sequential block (i.e. 30 blocks total). The overall trajectory can inform the speed of implicit learning (or lack thereof). Both the magnitude of the RT difference between the random and repeating block, and the rate of change in RT over sequential repeating blocks will be important in characterizing learning.

3.7.3.4 PAL

Two indices of learning will be used for the PAL: accuracy and median RT on the final 6-trial block *minus* accuracy and median RT on the first 6-trial block. Accuracy will be defined as the number of trials answered correctly divided by the number of trials attempted (i.e. maximum of 6 per block). RT will be defined as the latency to respond of trials for which the first response is the correct response. A final accuracy score with a positive value and RT score with a negative value will indicate that learning has taken place.

3.7.4 Model 3

The third model proposes brain dynamics as the mediating path. For this analysis, ERP components of interest were isolated based on those which can be interpreted as indices of learning on feedback-dependent learning paradigms. This is only possible for EEG data from the PST and IGT tasks, and as such EEG data from only these two tasks will be used.

Chapter 4

4.1 Participants

4.1.1 Timeline and COVID-19 Impacts

Recruitment began in December 2018 and was completed in October 2021. Participants initially received a $50 gift card for participation. Due to COVID-19, data collection was paused in March 2020 and began again in June 2020. The last participant seen before COVID-19 occurred on 3-10-2020, and the first participant seen after COVID-19 closures occurred on 8-3-2020. When recruitment began again after COVID-19 closures (June 2020), for the ensuing seven months, only four participants were enrolled, of which two were lost to attrition. Consequently, financial incentives for participation were increased to $100 beginning January 2021 in order to improve enrollment and retention. The remaining 25 participants were consented and scheduled between January and October 2021.

4.1.2 Pre-Screening Limitations

Additionally, recruitment initiatives and procedures, as well as methods of record keeping, changed over the duration of the study due to low recruitment and impacts of COVID-19. One of these changes was implementing pre-screening at the initial phone call to determine whether a participant would be able to attend the baseline appointment in-person in order to have full-screening done at that time. In many cases, participants were pre-screened but then not fully screened due to limitations in the baseline appointment (i.e. if remote, or otherwise). There were many instances of participants who were screened at the baseline appointment, but were never pre-screened during the phone call. Due to the inconsistencies and limitations in recording for both pre-

screen and screening, it will not be possible to report on individuals preceding the full screener administered at the baseline appointment.

There were a number of individuals were marked ineligible at the pre-screen, but were not shown as ever having been asked pre-screen or screener questions. This can only be attributed to staff error. Because it is impossible to know whether these individuals were ever screened and why exactly they were marked ineligible, it is assumed that these instances were marked ineligible due to staff error, and they will not be included in individuals screened.

Additionally, there were some instances of staff error in record keeping in screening and eligibility. For example, there are individuals who answered at least one screening question, but were not marked as eligible or ineligible (n=11). These individuals may or may not have been provided the opportunity to participate in the current study, and thus will not be included in the final count of individuals screened.

4.1.3 Screening

See Figure 5 for the CONSORT diagram for the flow of participants. There were 416 individuals screened for eligibility 429 times; 11 individuals were screened two times and one individual was screened three times. Of the 429, 280 individuals were ineligible. The 280 is further reduced to 267 individuals (due to duplicate screenings). Of the 267 ineligible individuals, 88 were ineligible due to meeting at least one exclusion criterion, 75 due to BMI > 35, 71 individuals due to a COPD diagnosis, 14 were ineligible due to having abnormal vision, 11 due to having neurological injuries or deficits, four individuals due to having a current cancer diagnosis, two individuals were ineligible due to current pregnancy/lactation, and two due to having had cancer

treatment within six months. The CONSORT diagram reflects the number of individuals who endorsed each reason (including those who endorsed more than one reason).

The questions “Do you have a current diagnosis of cancer (are you not cancer free)?” & “Have you received cancer treatment (e.g., chemotherapy, radiation, surgery, medications, or other treatments) in the past six months?” were added shortly after recruitment began, and 14 of those screened were missing responses for these questions. The question “Do you have any neurological injuries or deficits that may affect memory, cognition, or perception?” was added much later in screening, and 130 of those screened were missing responses for this question. No participant had missing responses to all of the screening questions.

4.1.4 Eligible Individuals

Of the remaining 149 individuals who were eligible, 77 declined participation and 72 were enrolled. Reasons for declining were not provided. These 72 individuals signed informed consents and were scheduled for their appointment into the current study. Of these, 54 individuals (75%) were present at their appointment (i.e. 18 or 25% lost to attrition). Note: three out of the 54 individuals declined EEG participation upon arrival to their appointment; therefore, only 51 participants provided EEG data.

Sociodemographic and smoking behavior comparisons (i.e. age, race/ethnicity, income, marital status, education, cigarettes smoked per day, years of smoking, heaviness of smoking index score at baseline, and depression; see section 4.2.1 for details on how variables were operationalized) were conducted between those who declined and those who enrolled. No significant differences were found between the two groups. Additionally, these same comparisons were made between those who were lost

to attrition and those who completed their appointment. No significant differences were found between these two groups.

4.2 Sociodemographic Characteristics

4.2.1 Methodology

The following sociodemographic variables were assessed: age, race/ethnicity, income, marital status, education, cigarettes smoked per day, years of smoking, heaviness of smoking index score at baseline, depression, body mass index (BMI), and menopause status. Age and years of smoking were measured in years and treated as continuous self-report variables.

Race and ethnicity were assessed by asking the two questions, “Are you Hispanic or Latino?” and “How would you best describe your race?” Answers for the two questions were collapsed into one variable such that if participant reported ethnicity as Hispanic, they were categorized as Hispanic; otherwise, participant was characterized by race. Final categories were defined as: White, Black or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian / Alaska Native, More than one race, or Hispanic/Latino. Due to the low numbers, participants were then collapsed into three total categories for race: White, Black/African American, or Other.

Income was assessed by asking “Which of these categories best describes your total combined family income for the past 12 months?” Response options were categorical in nature: $0 to $10,999, $11,000 to $20,999, $21,000 to $30,999, $31,000 to $40,999, $41,000 to $50,999, $51,000 to $60,999, $61,000 to $70,999, $71,000 to $80,999, $81,000 to $90,999, $91,000 to 100,000, $100,000 or greater, or Refuse to

Answer. Of the enrolled participants, no responses were missing, and only two endorsed “Refuse to Answer.”

Participants were additionally asked “What is your present marital status?” with answer choices of: single, married, divorced, widowed, living with significant other, or separated. There were no missing responses.

To assess education, participants were asked “How many years of education have you COMPLETED?” Answer options were categorical and included: zero (no formal schooling), one thru 12 years (selecting 12 years refers to GED or High School diploma attained), 13 years (Some college/technical school), 14 years (Associates Degree), 16 (Bachelor Degree), 17 (Some Post-graduate School), 18 (Master Degree), and 20 (Post-graduate Degree: M.D., Ph.D., etc.). Responses were not transformed or collapsed in any way.

To assess depression, participants were asked to completed a modified version of the Patient Health Questionnaire (PHQ)-9 for Major Depressive Disorder or Other Depressive Syndromes [79], which asked all questions listed in Appendix A. The questions ask about the frequency of depression symptoms over the past two weeks; the question assessing suicidal ideation was omitted because clinical setting was not equipped for responses (i.e. emergency mental health services). Responses were summed for a total score of zero to eight, reflecting the number of symptoms experienced on more than half of days over the past two weeks (i.e. the number of items out of 8 possible for which participants endorsed response option 2 or 3). Participants who endorsed questions 1 and/or 2 and had a total score greater than four

were classified as depressed; otherwise participants were classified as not depressed. Depression classification was not used as exclusionary criterion.

Participants completed the two-item Heaviness of Smoking Index [80] to assess smoking dependence. The questions ask “How soon after you wake do (or did) you smoke your first cigarette or cigarillo?” (with response options: 3 = within 5 minutes | 2 = 5-30 minutes | 1 = 31-60 minutes | 0 = after 60 minutes) and “On days when you smoke(d), how many cigarettes or cigarillos do (or did) you usually smoke per day?” (with response options: 0 = 0, 1 = 1-5, 2 = 6-10, …, 7 = 31 or more). Responses were summed to a total score ranging from 0-6. Scores of 0-1 indicate low dependence, 2-4 indicate moderate dependence, and 5-6 indicate high dependence [81; 82].

Cigarettes smoked per day (CPD) was assessed by asking “On average, how many cigarettes do you smoke a day?” Participants were allowed to enter any value between 0-100. However, there was a high level of missingness for this response (~20-25%). Almost all participants who were missing responses for CPD had answered another question which asks, “On days when you smoke(d), how many cigarettes or cigarillos do (or did) you usually smoke per day?” which required categorical responses. A second version of this variable was created based on CPD (if available) or the second question (if CPD missing). Therefore, this second CPD variable has categorical responses only, with few instances of missing (~ 0-3%).

Anthropomorphic measures were gathered at the baseline TTRP visit. Participants’ height (in inches) and weight (in pounds) was measured, and body mass index (BMI) was calculated with the following equation: $(\text{Weight} / \text{Height}^2) * 703$.

Lastly, menopause status was assessed at their EEG appointment with the question, “How would you describe your current menstrual status?” Participants were able to select from three answer options: 1) Premenopause (before having menopause; having regular periods); 2) Perimenopause/menopause transition (changes in periods, but have not gone 12 months in a row without a period); or 3) Postmenopause (after menopause).

4.2.2 Results

Participants (n=54) were all female, primarily of White (51.9%, n=28) or Black (27.8%, n=15) race, and single (66.7%, n = 36). Additionally, participants were on average 51 years of age and had 13 years of education; most participants had attained a high school diploma or greater (90.7%, n = 49). The majority of participants reported current menopause status as either peri- (11.5%, n=6) or post-menopause (59.6%, n=31). Only 9% of participants (n=5) were classified as depressed based on responses to the PHQ. See Table 1 for full sociodemographic characteristics.

4.3 Iron Biomarkers

4.3.1 Laboratory Analysis

Whole blood samples were collected at the conclusion of every session in lavender-top EDTA Vacutainer® brand tubes for analysis of complete blood count (including Hb). A second gold-top Vacutainer® brand serum separator tube (SST) was collected and centrifuged to separate serum for analysis of sFt, soluble transferrin receptor (sTfR), and C-reactive protein (CRP). Both tubes were collected by a trained phlebotomist, packed into a bio-hazard bag on ice, and transported to the OU Medical Pathology Laboratory for analysis.

4.3.2 Methodology

Of the 54 participants, one participant refused to receive a blood draw and two other samples were lost by OU Medical Pathology. Two additional participants' results for CBC were missing and two results for sTfR were missing. For this reason, all blood results represent 51 participants excepting Hb (n=49) and sTfR (n=49).

Hb is known to be affected by smoking status, and criteria have been defined to adjust for amount of cigarettes smoked per day [32; 83; 84]. As Sullivan et. al. suggest, Hb will be reduced by a range of 0.3 to 0.7 depending on number of cigarettes smoked per day. Both adjusted and unadjusted Hb will be reported.

Additionally, CRP, a biomarker of inflammation, was relatively high in the present sample (M=6.99; SD=13.68; Median = 3). One outlier was identified (CRP = 88.10) and was subsequently removed. The final CRP variable (n=50) still reported high levels of inflammation (M=5.37; SD=7.35; Median = 3). Of the 50 participants with remaining CRP values, 29 had values outside the norm (i.e. met diagnostic criteria for inflammation; 58.0%). For this reason, ferritin values were adjusted to account for inflammation [85; 86]. Both raw serum ferritin (sFt) and CRP-adjusted ferritin (asFt) are reported in Table 2 and subsequently used in Cook's equation to calculate TBI [37].

Lastly, because ferritin is known to be affected by sex, age, and race, ferritin percentile rank (sFtP) was calculated with both sFt and asFt. This was done using the National Health and Nutrition Examination Survey (NHANES) data (i.e. a national sample): a cumulative distribution function was used to transform the data into a percentile value. In this way, each participants' sFtP reflects the percentage of their age and race groups' sFt equal to or below their sFt (females only).

4.3.3 Results

See Table 2 for full report of the iron biomarkers. The average raw Hb was 14.04 g/dl (SD = 1.47). Surprisingly, CPD-adjusted Hb varied minimally from raw Hb at 13.82 g/dl (SD = 1.46). Few participants were classified as anemic using either raw Hb (4%, n=2) or CPD-adjusted Hb (n=3, 6.1%).

The average raw sFt was 103.12 (SD = 97.88; median = 75.70), which is perhaps unsurprising considering the high rates of reported peri- and post-menopause in the sample. Inflammation-adjusted ferritin (asFt) was considerably lower than raw sFt, with an average of 76.14 (SD = 68.66; median = 52.07). Of the total sample, a handful of participants met the criteria for iron toxicity using both sFt (n=11; 21.6%) and asFt (n=6; 11.8%).

Average sFtP was 46.3 (SD = 28.9; median = 43.6). Consistently, asFtP was lower with an average of 39.8 (SD = 26.7; median = 36.4). This means that the sample was generally representative of the average level of sFt in the population.

Lastly, average sTfR was 18.24 (SD = 7.31; median = 16.50). These values were used alongside sFt and asFt to calculate TBI. Using sFt, average TBI was 28.03 mg/kg (SD = 4.57; median = 28.67); using asFt, average TBI was a bit lower at 27.12 (SD = 4.38; median = 27.42).

4.4 Smoking Outcomes

4.4.1 Methodology

See Table 3 (Abstinence Frequency Follow-ups) for full details. Smoking outcomes included biochemically-verified abstinence, self-reported abstinence, and expired carbon monoxide levels and difference score, alternative tobacco product (ATP)

use, and marijuana smoking. Smoking cessation outcome variables were gathered from the parent study called the Tobacco Treatment Research Program (TTRP). The TTRP financially incentivizes in-person participation (and thus, providing carbon monoxide [CO] breath sample) at weeks 4, 12, and 26. However, weeks 1-3 are not incentivized and, thus, many participate in those appointments remotely; no breath sample is provided. Additionally, due to impacts from COVID-19, many appointments that would normally have been in-person became remote; many participants were not able to or otherwise did not provide a breath sample in order to corroborate self-reported abstinence. For this reason, both biochemically-verified abstinence and self-reported abstinence will be reported.

Self-reported abstinence was assessed with the question, “Have you smoked, even a puff, during the last 7 days?” (yes vs. no). Participants then provided a breath sample and a CO level of ≤6 parts per million [ppm; 24] was used as threshold for abstinence. Participants who reported no smoking within past 7 days and had an expired CO ≤ 6 ppm were considered abstinent. In cases of missing both self-report and CO or CO only, participants were considered to be smoking.

Self-reported abstinence only considers responses to the question asked above. Participants were either characterized as smoking (self-report smoking or missing) or abstinent (self-report abstinent).

An additional smoking cessation descriptive is raw CO reading. Higher levels indicate more smoking. Raw CO and change in CO from baseline (i.e. index of magnitude of reduction) are included in Table 3.

ATP and e-cigarette (EC) use are important indices for cigarette-abstinent nicotine use. Because the current study is interested in recovery from nicotine addiction, both ATP and EC use will be described at baseline and follow-up appointments. To assess for ATP, participants were asked the question, “In the last seven days, have you used any other form of tobacco (e.g., cigar, pipe, chewing tobacco, or snuff)?” (yes vs. no). To assess e-cigarette use, participants were asked the question, “In the last seven days, have you used any electronic cigarette or nicotine delivery device (e.g., vaping, JUULING)?” (yes vs. no).

Lastly, because smoking marijuana/cannabis will contribute to elevated CO levels, frequency of marijuana use will be described. Participants were asked the question, “During the past seven days, on how many days did you use marijuana/cannabis for medical or recreational purposes” (0-7). Participants were classified as marijuana users if their response was one or more. These participants were then asked “When you used marijuana in the past seven days, what was the method of delivery?” Response options to this question were: 1) I smoked it, for example, in a joint, bong, pipe, or blunt; 2) I ate it, for example, in brownies, cakes, cookies, or candy; 3) I drank it, for example, in tea, cola, or alcohol; 4) I vaporized (vaped) it, for example, in an e-cigarette-like vaporizer or another vaporizing device; 5) I dabbed it, for example, using waxes or concentrates; 6) I dissolved it in my mouth, for example tablets, oils, or strips; 7) I applied it to my skin, for example lotions, oils, or patches; 8) I used it some other way. Only participants who were classified as marijuana users and endorsed “smoking” as method of delivery were considered marijuana smokers. Marijuana smoking is described in Table 3.

4.4.2 Statistical Analysis

To compare the mean differences in expired CO across time points, ANOVA methodology was employed. ANCOVA, including iron as a covariate, was considered. However, the results did not differ meaningfully, and thus the simpler ANOVA was selected as the final model. Upon significant findings, paired samples t-tests were employed for the pairwise comparisons in order to find which time points were significantly different. Because the pairwise analysis had six comparisons in total, Benjamini & Hochberg's methodology was utilized to control for a false discovery rate (FDR) [87]. Note: all other comparisons made used FDR to control for all comparisons within a task and measure (i.e. PST behavioral comparisons corrected for four analyses, etc.).

4.4.3 Results

Generally, abstinence rates for weeks 1-3 were low (7.4%, 9.3%, and 13.0% respectively). When considering self-reported abstinence, rates improved slightly (24.1%, 20.4%, and 22.2% respectively). Note that participants were not financially incentivized to attend in-person appointments during these weeks, thus, there are high rates of remote participation or no-shows.

ATP and EC use at all follow up appointments tended to be quite low (0-7.4% and 1.9-7.4% of the sample, respectively). While the rate of marijuana smoking was not high, it was higher than use of ATP and ECs (5.6-13.0%). Note that marijuana smoking will impact expired carbon monoxide levels, and could at least partially explain the difference between self-reported abstinence and biochemically-verified abstinence.

At week 4, abstinence frequency was 24.1% (n=13) while self-reported abstinence was 31.5% (n=17). Biochemically-verified abstinence rates at weeks 12 and 26 were much lower than that of week 4 (20.4% and 13.0% respectively). See Figure 6 for abstinence frequencies and Figure 7 for changes in expired CO.

With respect to expired CO, a repeated measures ANOVA found significant differences in CO across time points ($F(3, 39) = 5.68$; $p = 0.003$). Follow-up pairwise t-test (controlling for six analyses) showed two significant differences: the difference from baseline to 4-weeks post quit ($t(33) = 6.66$; $p < 0.0001$; *p-adjust* < 0.0001) and the difference from baseline to 12-weeks post-quit ($t(21) = 3.37$; $p = 0.003$; *p-adjust* = 0.009).

4.5 Indices of Dopamine

4.5.1 PST Accuracy (choose A/avoid B) & Conflict Latency

See section 4.6.1.1 and 4.6.1.2 for PST methodology, data cleaning, and results. Choose A & Avoid B accuracy & conflict latency will be used as indices of DA.

4.5.2 Blink Rates

4.5.2.1 Methods

For all files, the BLINKER plugin Version number [88] was used on the raw EEG data collected. This plugin first filters out EEG occurring <1 and >20 Hz. Then, the program isolates the best candidate signal which is determined by a threshold of number of blinks a signal must have in order to be considered the signal. At this point, potential blinks are identified as intervals during which signal amplitude measures > 1.5 standard deviations above its mean. For each blink, the best linear fit for the inner 80% of the blink curve is computed. Blinks and good blinks are defined according to the R^2

value of the correlation between the computed blink trajectory with the actual blink itself. A ”blink” is defined as any potential blink, and a “good blink” is defined as $R^2 > 0.90$. To clarify, a good blink is a blink which meets additional goodness-of-fit criteria.

For any datafile which provided an error (i.e. low amplitude blinks, no EEG data recorded), data were manually plotted and counted. The resting state data was recorded alongside eye video recordings (video recorded by laptop webcam at data collection). Thus, manual viewing of the eye recording provided the number of blinks for this task. For the remaining tasks, EEG and behavior data alone was recorded. For these tasks, EEG data was filtered to retain 1:100 Hz activity. The EEG data was plotted.

In order to extract blink rate, total number of blinks was divided by total time of the file. Time of the file was defined as the length from first stimulus until the last stimulus. Importantly, the PST task periods were separated into training phase and testing phase. Thus, PST provides two task-related BR values, while the other tasks provide one each (i.e. total of six BR indices).

4.5.2.2 Results

On average, spontaneous BR was 6.4 blinks per minute (BPM) while task-related BR was 5.5, 9.2, 10.4, 10.4, and 11.1 respectively during the SRT, IGT, PST Test, PST Train, and PAL tasks, respectively. It is possible that BR was highest during the PAL task due to the increased complexity of interaction with keyboard for responses and confirmation of responses.

4.6 Indices of Learning

Trials with RTs outside of lower/upper boundaries were discarded. The lower bound was defined as 200 ms, such that trials with RT < 200ms were discarded from all behavioral tasks. The upper bounds were defined as the mean RT + three standard deviations for each task. See Table 5 for all behavioral task outcomes.

4.6.1 PST

4.6.1.1 Methods

Of all 54 participants, one participant did not complete the behavioral PST task, and two participants began the task but did not complete due to computer error. Thus, data from 51 participants were included in the behavioral analysis. Additionally, four individuals had 100% accuracy on A trials, thus, only 47 individuals' RT data was available for incorrect A trials.

Beyond removing trials with RT < 200ms, the upper bound was defined as 10,243ms. From training phase, 12 trials were removed due to RT being below lower threshold, and 714 trials were removed due to RT being above upper threshold. From testing phase, 20 trials were removed due to RT being below lower threshold and 679 trials were removed due to RT being above upper threshold. Thus, 2,334 trials were included in total for the last training phase (mean of 45.8 trials/participant) and 2,361 trials were included in total for the testing phase (mean of 46.3 trials/participant).

4.6.1.2 Results

Reached criteria

Of the 51 participants with data, only 21 (41.2%) reached the criteria to move on to the test phase. However, 33 participants (64.7%) reached criteria respective to just the AB pairs.

Accuracy and latency in choose A/avoid B

On average, participants had a mean accuracy of 64.1% (Median = 65; SD = 21.9) on A trials and 44.8% (Median = 45; SD = 17.3%) on B trials. A paired samples t-test show that the difference in learning to choose A and learning to avoid B was significant ($t(50) = 5.40$; $p < 0.0001$; *p-adjust* = 0.08). This indicates that most participants learned from positive feedback but not negative feedback. Average RT for choose A trials was 1210 ms when correct (Median = 1132; SD = 653) and 1328 ms when incorrect (Median = 1152; SD = 841). Average RT for avoid B trials was 1309 ms when correct (Median = 1110; SD = 697) and 1375 ms when incorrect (Median = 1157; SD = 940). See Figure 8.

Latency in High vs. Low Conflict Trials

On average, RT on high conflict trials (Mean = 1369; Median = 1247; SD = 680) was higher than in low conflict trials (Mean = 1311; Median = 1117; SD = 767). Follow-up paired samples t-test showed that this difference was not statistically significant for choose A trials ($t(46) = -1.96$; $p = 0.06$; *p-adjust* = 0.08) nor avoid B trials ($t(50) = -0.11$; $p = 0.91$; *p-adjust* = 0.91).

4.6.2 IGT

4.6.2.1 Methods

Data from one participant was lost due to software error; this participant did not complete the task and her behavioral data was not included in final analyses. All behavioral analyses for IGT included 53 participants. Beyond removing trials with RT < 200ms, 8724 ms was identified as the upper bound. Of the 10,600 trials in the dataset, 285 trials were removed due to RT being below lower threshold and 71 trials were

removed due to RT being above upper threshold. In total, there were 10,244 trials (mean of 193.3 trials per participant).

Indices of learning during the IGT task include reaction times, learning score, and trajectory of learning. First, each trial's response was classified as playing from advantageous or disadvantageous decks. Median reaction time was calculated per participant per deck type (i.e. median reaction time for advantageous trials vs. disadvantageous trials).

Next, learning score was calculated per participant for all 200 trials. This was operationalized as the frequency participant played from an advantageous deck minus frequency they played from a disadvantageous deck [78]. For example, scores could range from -200 (i.e. if participants played exclusively from disadvantageous decks) to 200 (i.e. if participants played exclusively from advantageous decks).

Additionally, learning score was calculated per trial block (i.e. 10 blocks). Each block has a possible learning score ranging from -20 to 20, as each block includes 20 trials. Trajectory of learning over time was calculated as the slope of a linear regression equation for which learning score was regressed onto block. In this way, each participant has a unique slope calculated. The magnitude of the slope will provide information toward rate of learning; values of greater magnitude indicate greater rate of learning.

4.6.2.2 Results

Advantageous vs. Disadvantageous Frequency & RT

On average, the frequency of playing from advantageous decks was 91.3 (SD = 20.7; range = 46-149) while the frequency of playing from disadvantageous decks was

108.8 (SD = 20.7; range = 51-154). Paired samples t-test showed that there was a significant difference in frequency played from advantageous and disadvantageous decks ($t(52) = -3.10$; $p = 0.003$; *p-adjust* = 0.03). See Figure 9 for frequency of deck play over blocks. Interestingly, the average RTs were similar in both play conditions; average advantageous deck RT was 862 (SD = 392; Median = 794) while average disadvantageous deck RT was 866 (SD = 397; median = 754). The small difference did not reach statistical significance ($t(52) = -0.11$; $p = 0.90$; *p-adjust* = 0.90).

Learning Score and Slope

The average learning score (advantageous frequency minus disadvantageous frequency) was -17.5 (SD = 41.3; range = -108, 98). Slope of learning score over the 10-blocks was 0.21 (SD = 0.91; range = -1.66, 2.82). Following Mapelli et. al., [78] an ANOVA used to compute the effect of block (i.e. time) on learn score found no significance ($F(5.44, 277.22) = 1.31$; $p = 0.255$), meaning that the trajectory of learn score throughout the duration of the task did not change significantly (despite the low magnitude decrease in selecting from disadvantageous decks depicted in Figure 9). This implies a dysfunction to learning during the IGT task.

4.6.3 SRT

4.6.3.1 Methods

Of the total 54 participants, data for SRT was not collected from two due to declining further participation. As such, all behavioral analyses from the SRT is based on data from 52 participants.

Because the SRT task does not proceed to the next trial until the correct response is provided (i.e. unlimited responses per trial), any trial for which the first

response was not the correct response was eliminated from analysis. Beyond removing trials with RT < 200ms, 4153 ms was identified as the upper bound. Of the total 23,712 trials in baseline sessions (456 trials per 52 participants), 1014 were eliminated due to incorrect first response. Of the remaining 22,698 trials, 51 trials were removed due to RT being below lower threshold and 112 trials were removed due to RT being above upper threshold. The final dataset included 22,531 trials (mean of 433.3 trials per participant).

Indices of learning include both learning score and trajectory of RT. The overall learning score was calculated as the median RT on the last random block minus the median reaction time on the last repeated block [66]. Positive values indicate that implicit learning took place overall (i.e. if difference reaches statistical significance). Trajectory was calculated similar to procedure for the IGT task (i.e. slope). Negative slope values of RT indicate implicit learning is occurring over time (i.e. taking less time on trial n+1 than on trial n).

4.6.3.2 Results

The mean reaction time in the last repeated sequence was 568 ms (SD = 165; median = 549). The mean reaction time in the last random sequence was 616 ms (SD = 170; median = 578). Although the difference score (RT on the last random sequence minus the RT on the last repeated sequence) ranged from -306 to 415, the participant average was 48 (SD = 119; median = 50). This difference reached the criteria for statistical significance ($t(51) = -2.90$; $p = < 0.01$). The difference in RT indicates implicit learning did take place overall.

Lastly, the slope of RT over the 30 blocks (see Figure 10) was an average of -1.6 (SD = 7.84), meaning that RT during the 30 blocks trended toward reduction. Despite this downward trajectory, an ANOVA modelling the effect of block (i.e. time) on reaction time was not statistically significant ($F(3.57,181.82) = 2.315$; $p = 0.07$).

4.6.4 PAL

4.6.4.1 Methods

Beyond removing trials with RT < 200ms, 10,304ms was identified as the upper bound. Of the 3,180 trials included, 16 trials were removed due to RT being below lower threshold and 66 trials were removed due to RT being above upper threshold. A total of 3,098 trials remained (mean of 58.5 trials per participant). The first 6-trial block and last 6-trial block were isolated. Accuracy and RT were calculated for each block.

4.6.4.2 Results

The average accuracy (range 0-6) for block 1 was 4.3 (SD = 1.6; median = 5). Average block 2 accuracy was 4.87 (SD = 1.4; median = 5). RT for block 1 was 3087 ms (Median = 2804; SD = 1283), while the RT for block 2 was 2458 ms (Median = 2157; SD = 981). This difference was statistically significant ($t(52) = -3.98$; $p < 0.001$). While this could indicate that participants learned over the task, there was no feedback or repetitions occurring throughout. Thus, reduced RT may be reflective of practice effect (i.e. improving interaction with computer).

4.7 Indices of Neural Activity

Standard in the field, many EEG recordings were unusable due to high levels of noise. Noteworthy, being a smoker has been identified as a predictor of noisy neural

data collection [89]. Thus, it can be expected that a number of data files must be discarded due to noise.

For the PST, EEG data from five participants were not recorded (three refused EEG recording, two did not complete PST at all). Of the remainder, two files were lost due equipment error and eight files were removed due to high levels of noise. Therefore, final analyses of the PST EEG data included data from 39 participants.

For the IGT, EEG data from four participants was not recorded (three refused EEG recording, one did not complete IGT). Of the remainder, three files were lost due to equipment error and seven files were removed due to high levels of noise. Therefore, final analysis of the IGT EEG data included data from 40 participants. See Table 6 for details.

ERP methodology

The signal was filtered at 40 Hz, then epoched around feedback or response (respectively), at the time windows identified for the component, and averaged across the selected electrodes (frontocentral [FC] = Fz, FC1, FC2, C3, and C4; centroparietal [CP] = Pz, CP1, and CP2). Minimum trial count was 25 trials per trial condition per participant; any that contained less than this were eliminated from analyses. With respect to trial-matching, response-locked IGT showed a significant difference in trial count between trial types, while the remaining tasks and trial-types did not. Trial-matching did not alter results for the response-locked IGT, therefore, none of the analyses reported herein employed trial-matching in order to maintain congruency.

Feedback-Related Negativity (FRN)

Following Frank et. al., [63], trials for extracting FRN were epoched from 100 ms before feedback onset through 1000 ms after feedback onset. Baseline correction was applied with respect to the 100 ms pre-event period (-100 – 0 ms). The feedback related negativity was extracted by averaging across frontocentral electrodes, and according to Frank et. al. [63], was identified as the difference between the first negative peak (within the time window of 190-300ms) and the preceding positive peak. The preceding positive peak to the peak FRN itself was defined on the 150-190 ms time window. The FRN was event-locked separately to correct (FRN-cor) and incorrect (FRN-inc) feedback; or win money (FRN-win) vs. lose money (FRN-loss) in the case of IGT task. The same electrodes, time window, and definition were used to assess error-related negativity [63].

Error-Related Negativity (ERN)

Frank et. al. [63] state that error-negativity is informative with respect to a decrease in dopaminergic activity following incorrect responses and error feedback. Thus, to thoroughly examine negativity in the ERPs from the IGT and PST, both response-locked negativity and the feedback-locked negativity were examined. Trials were separated according to positive and negative (or win/loss) feedbacks and (advantageous/disadvantageous) correct and incorrect responses, respectively.

Following Frank et. al., [63], trials for extracting ERN and CRN were epoched from 800 ms before feedback onset through 2000 ms after feedback onset. Baseline correction was applied with respect to the first 100ms period (-800 to -700 ms). The feedback related negativity was extracted by averaging across frontocentral electrodes,

and according to Frank et. al. [63], should be identified as the difference between the first negative peak (within the time window of 50-130ms) and the preceding positive peak. The preceding positive peak to the peak FRN itself was defined on the 30-70ms time window.

P300

Polich [90] states that the current understanding of the P300 component is that it "reflects an information-processing cascade" composed of attentional, memory, and inhibitory processes. He also states that P300 is hypothesized to be guided by "neuroinhibition" [90]. Mapelli et. al. analyze the IGT's P300 between 350-450ms after feedback-onset [78] and West et. al. analyze the PST's P300 in a similar way [91]. Mapelli et. al. consider electrodes Fz, Cz, and Pz (separately) while West et. al. consider electrodes FCz, Fz, FC1, and FC2 for what they call P3a and Pz, CPz, P1, and P2 for what they call P3b (all separately rather than averaged). To integrate the methods here, both the PST and IGT data, feedback-locked trials averaged both frontocentral electrodes and centroparietal electrodes separately, during the time windows 350-450 ms after feedback onset.

P200

The feedback-locked P200 component was identified from the IGT task. Mapelli et. al. identify the P200 component as the average amplitude during the time window 150-250ms after feedback onset [78]. The P300 was also event-locked separately to correct (P300-cor) and incorrect (P300-inc) feedback. The P200 was extracted in both frontocentral and centroparietal electrodes.

4.7.1 PST Results

Correct and Error-related negativity

The average ERN amplitude was -0.52 µV (± 0.30) while the average CRN amplitude was -0.49 µV (± 0.25); see Figure 11. Latency to ERN was 66 ms (± 21) and latency to CRN was 65 ms (± 21). Paired t-tests showed that there were no significant differences between CRN and ERN amplitude ($t(38) = 0.60$; $p = 0.55$; *p-adjust* = 0.63) nor latency ($t(38) = -0.33$; $p = 0.74$; *p-adjust* = 0.74).

Feedback-related negativity

The mean amplitude for FRN on correct trials -2.12 µV (± 1.52), while the mean amplitude for FRN on incorrect trials was -2.90 µV (± 1.76); see Figure 12. The latency to FRN on correct trials was 245 ms (± 33) and the latency to FRN on incorrect trials was 256 ms (± 2). Paired t-tests show that there was a significant difference between correct and incorrect amplitudes ($t(36) = 4.03$; $p < 0.001$; *p-adjust* = 0.03), but not latencies ($t(36) = -1.49$; $p = 0.15$; *p-adjust* = 0.24).

P300

In the fronto-central region, the average correct P300 amplitude was 0.77 µV (±1.21), while the average incorrect P300 amplitude was 1.53 µV (± 1.69); see Figure 13. In the centro-parietal region, the average correct P300 amplitude was 2.59 µV (±1.82), while the average incorrect P300 amplitude was 2.29 µV (±2.29); see Figure 14. Paired t-tests show that there was a significant difference between correct and incorrect P300 amplitudes in the fronto-central region ($t(38) = -5.25$; $p < 0.0001$; *p-adjust* = 0.03), but not centro-parietal ($t(38) = 1.38$; $p = 0.18$; *p-adjust* = 0.24).

4.7.2 IGT Results

Error-related negativity

The average ERN amplitude was -0.48 μV (±0.31), while the average CRN amplitude was -0.50 μV (±0.34); see Figure 15. Latency to ERN was 64 ms (±20), while latency to CRN was 60 ms (±16). Paired t-tests show that there was no significant difference between ERN and CRN amplitudes ($t(37) = -0.05$; $p = 0.96$; *p-adjust* = 0.96) or latencies ($t(37) = -1.17$; $p = 0.25$; *p-adjust* = 0.56).

Feedback-related negativity

One additional dataset was unable to be analyzed for FRN due to erroneous triggers in the file. Thus, the FRN analyses included data from 43 participants.

The average FRN for wins (i.e. win money) amplitude was -3.02 μV (±1.74), while for FRN for losses (i.e. lose money) amplitude was -3.33 μV (±1.76); see Figure 16. Latency to FRN for wins was 287 ms (± 22) and FRN for losses was 287 ms (±23). Paired samples t-tests show that there was a significant difference between FRN for wins and FRN for losses amplitudes ($t(38) = -2.44$; $p = 0.02$; *p-adjust* = 0.06) but not for latencies ($t(38) = -0.24$; $p = 0.81$; *p-adjust* = 0.96).

P300

In the fronto-central region, the average win P300 amplitude was 0.70 μV (±0.97), while the average loss P300 amplitude was 0.95 μV (± 1.24); see Figure 17. In the centro-parietal region, the average win P300 amplitude was 3.54 μV (±2.10), while the average loss P300 amplitude was 3.43 μV (±2.13); see Figure 18. Paired t-tests show that there was a significant difference between win P300 and loss P300 amplitude

in the fronto-central region ($t(39) = -2.72$; $p < 0.01$; *p-adjust* = 0.06), but not the centro-parietal region ($t(39) = 0.74$; $p = 0.47$; *p-adjust* = 0.74).

P200

In the fronto-central region, the average win P200 amplitude was 2.36 µV (±1.26), while the average loss P200 amplitude was 2.50µV (± 1.43); see Figure 19. In the centro-parietal region, the average win P200 amplitude was 0.88 µV (±1.53), while the average loss P200 amplitude was 0.82 µV (±1.47); see Figure 20. Paired t-tests show that there was a significant difference between win and loss P200 amplitude in the fronto-central region ($t(39) = -2.35$; $p = 0.02$; *p-adjust* = 0.06), but not the centro-parietal region ($t(39) = 0.69$; $p = 0.49$; *p-adjust* = 0.74).

4.8 Bayesian Variable Selection

Rationale

The current analysis tested the three mediation models identified in chapter 3. However, it must be acknowledged that the number of variables identified as mediators is high (i.e. 51). Eventually, these mediators will be indicators of a common factor (i.e. learning, dopamine, neural activity). However, due to limitations in sample size of our current data, the current analysis aims to reduce the number of variables/indicators going into the final mediation analysis.

For this reason, the EMPub approach [92] was applied to the data to select the most influential variables in the data. EMPub is a useful tool for Bayesian variable selection; it is particularly meant for guiding selection of causal mediators in a data-driven and exploratory way (i.e. when theory is lacking). This approach quantifies each potential mediator's influence between predictor(s) and outcome variable(s). Each

mediator is defined as providing a chance probability of mediation. The probability of mediation is updated through iterations of a Markov Chain Monte Carlo (MCMC), resulting in posterior probabilities which are bound between 0-1 and provide information which can be used to make inferences about each mediator. Posteriors close to 1 indicate high performance in explaining the relation between the predictor and outcome. Additionally, this method contains one parameter which informs the number of mediators recommended to select using a data-driven approach.

Due to the exploratory nature of the analysis, four predictors (Ferritin, Ferritin Percentile, Hgb, and TBI) and 2 outcomes (Change in CO at 4 weeks and 12 weeks) were selected. Additionally, because the outcomes contained significant missing data, 3 approaches were selected in handling these. First, missing values remained missing (i.e. the EMPub model is able to handle this). Second, because the intent-to-treat paradigm assumes that participants missing their TTRP appointment are smoking, this assumption was extended to CO values: missing assumed no change from baseline (i.e. values replaced with 0). Third, missing values were imputed using predictive mean matching (see below). All three versions of the outcomes were modeled to ensure similar results regardless of how missingness was treated. In total, 24 iterations of the EMPub model were analyzed (4 predictors X 2 outcomes X 3 missing outcome strategies).

Missing Data Imputation

The EMPub model cannot handle missing data in the predictors or mediators. Therefore, missing values were imputed in the dataset containing the final predictors, mediators, and outcomes. Typical imputation methodology includes regression

imputation (i.e. replacing missing values with their estimated values based on other predictor variables). However, there are notable limitations relevant to the current dataset when using regression as the imputation strategy: 1) binary or categorical variables cannot be imputed, 2) estimations are imprecise when sample size is small, and 3) there are limitations in handling datasets wherein variables have different variances. To account for the limitations, instead of regression, the current study employed predictive mean matching imputation using the “MICE” package built for R version 4.0.5 [93-95]. MICE uses its algorithm (implementing Fully Conditional Specification) [94] to model missing data for each variable individually, and it can handle most types of variables for imputation (i.e. continuous, categorical, binary). It is anticipated that imputations resulting from PMM will be similar to multiple imputation

Results

The 24 total versions of the models recommended selecting 40-45 variables in total (Mean = 43; Median = 43).

A conservative and a liberal approach were undertaken, and final variable selection was an average of the two. In the conservative approach, only variables that were identified as useful in all 24 models (n=32) were selected. In the less conservative approach, and particularly because the all models recommended selecting 43 mediators on average, each model’s recommended variable retention was combined into one dataset, and each variable was counted and listed in order from identified in all to identified in the fewest of models. The difference was 11 variables. Of the 11 variables, a few of them were theoretically important while others were identified as least

influential in a number of EMPub models. Therefore, the first half of the 11 (i.e. six) were retained as final mediators. In total, 38 mediators were selected.

Final Mediators

The final mediators indicating dopamine were RT on low conflict PST trials, RT on high conflict PST trials, spontaneous BR and task-related BR. The final mediators indicating learning were having met criteria at end of training phase (yes/no), percent accuracy on A trials during PST test, percent accuracy on B trials during PST test, RT in A-trials when participant chose favorably, RT in A trials when participant chose unfavorably, and RT in B trials when participant chose unfavorably. There was a total of 23 mediators selected to indicate neural activity. In the interest of reducing these even further, only those which were considered theoretically interesting were selected. The final mediators indicating neural activity were: correct and incorrect FRN amplitudes (both IGT and PST), correct and incorrect P200 amplitudes for IGT (FC region), incorrect P300 amplitudes for PST (FC region).

4.9 Generalized Structural Component Analysis

Rationale

The current data is limited in its ability to test mediation models. A traditional SEM model cannot be employed (i.e. small sample size, etc.). For this reason, Generalized Structural Component Analysis (GSCA) framework was selected to test the three mediation models [96] as it is specifically able to handle small sample size SEM. By using GSCA to run three structural models, the current analysis will avoid biased estimates due to inflated Type I error. Additionally, the GSCA methodology does not

make distributional assumptions, and therefore, can handle categorical or dichotomous predictor and/or outcome variables.

Measurement and Structural Models

The measurement model is entirely exploratory in nature. The indicators that were used for each other three latent factors in our model was informed by the results from the EMPub analysis.

Additionally, the current analysis amended the structural models defined in chapter 3 (see Figures 2, 3, and 4). The paths between DA and abstinence as well as the paths between iron and abstinence must be two-way paths. It has been well-established that cigarette smoking will impact DA systems, causing a hyperdopaminergic state in relevant reward/addiction pathways [14]. Additionally, as noted in 4.3.1, evidence has shown that cigarette smoking will impact iron levels, particularly Hb and Ferritin [33]. See Figures 21-26 for final structural and measurement model specification.

Outlier Removal

Note that the predictive mean matching imputation was applied to the data to remove missingness. Outliers were identified both visually and using the well-established inter-quartile range criteria (IQR). In this method, the IQR is defined as 3rd Quartile *minus* the 1st Quartile; any data points identified at greater than or less than 1.5 times the IQR are deemed outliers. Those identified for removal using IQR were accepted with visual inspection. After removal, missing data points were again imputed using the “MICE” package built for R version 4.0.5 [93-95].

Results

Model Fit Indices:

See Table 8 for a summary of model fit indices, and Tables 9-14 for correlation matrices. FIT indicates variance in outcome accounted for by the model structure. AFIT is complexity-adjusted FIT index; it cannot be interpreted with respect to variance in outcome, though it can be used to compare competing models. An acceptable Goodness-of-Fit-Index (GFI) is any ≥ 0.89, while an acceptable Standardized Root Mean Squared Residual (SRMR) is ≤ 0.09.

In the current analyses, models had FIT = 0.34-0.37, indicating that they accounted for about 34-37% of the variance in abstinence at 4- and 12-weeks post-quit. The GFI and SRMR indices indicate a borderline acceptable fit for models in which DA or learning are mediators; these indices for models in which neural activity is the mediator are less acceptable. Because these models are exploratory in nature, one can conclude that the predictor and mediators account for an interesting and not insignificant amount of variance in abstinence outcomes.

Factor Loadings

The factor loadings (see Tables 15, 16, and 17) are consistent between the multiple models. Noteworthy results are detailed here.

Only three factors loaded significantly onto iron: RBC (β = 0.864), HCT (β = 0.878), and Hb (β = 0.967). Interestingly, neither sTfR nor ferritin percentile loaded significantly onto iron. In alternative models in which raw ferritin and/or TBI were included, neither were significant. This seems to suggest that, in these particular models, RBC, HCT, and Hb were the most significant markers of iron.

Only task-related BR loaded significantly onto dopamine. Spontaneous BR did not load significantly, although the confidence interval may be suggestive (95% CI = -0.066, 0.682). RT for both low-conflict and high-conflict trials during the PST testing phase were not significant.

Path Coefficients

See Table 16 for all path coefficients (or Figures 30-35). Even with the low sample size, the path from iron to learning reaches significance when the outcome is 4-week abstinence and 12-week abstinence ($\beta = -0.331$ and $\beta = -0.338$), indicating that iron status has a significant negative impact on learning such that as iron increases, learning decreases. Though this finding is contrary to our hypothesis, it could be guided by the limitations of our sample (i.e. 11-21% of sample characterized by iron toxicity). Not entirely unexpected due to small sample size, the remaining paths do not reach significance. However, the direction of relation for some path coefficients is worth mentioning.

The paths from iron to abstinence at four weeks and abstinence at four weeks to iron were positive, indicating that higher iron status allows for abstinence, while those who are abstinent have a higher iron level. Paths between iron and abstinence at 12 weeks were much closer to 0, indicating that the relationship may be less causal for longer-term smoking cessation outcomes. In contrast, iron and abstinence at four weeks are related at a higher magnitude (though, as mentioned, still not significantly so).

The paths between DA and abstinence at 4 weeks are negative, while those between DA and abstinence at 12 weeks are positive. This interesting tradeoff may hint at a paradigm wherein more DA availability at baseline leads to less likelihood of

quitting when the quit date was more recent (i.e. just four weeks ago) and when participant has quit there was less DA availability at baseline. Conversely, at 12-weeks post-quit, those who were abstinent had higher DA availability at baseline, and higher DA at baseline led to increased likelihood of quitting.

Lastly, the path between learning and abstinence at four weeks follow the hypothesized directionality (i.e. higher learning scores lead to increased likelihood of abstinence).

Chapter 5

5.1 Sample Characteristics

The present analysis aimed to analyze impacts of variations to iron status on smoking cessation, indirectly by way of learning changes, changes to indicators of neural activation, and dopaminergic changes.

The present study was developed as a preliminary, observational study. As such, the sample recruited was limited in several notable ways. Namely, participants in the sample were, on average, middle-aged. As iron levels increase with age [18], the average ferritin was reported as >100 µg (almost median percentile compared to nationally-representative data). It is also worth mentioning that most participants were either peri- or post-menopausal (72%), which also corresponds to higher iron levels [97; 98]. Not only were the participants generally high in iron, but additionally, a significant amount of the sample was deemed iron toxic (22%), meaning that about one-quarter of the sample were considered to have excessive iron. Excessive iron likewise has a poor impact on health, particularly neurodegeneration and similar diseases [99; 100]. Thus, it is possible that the results found may reflect a sample with heightened iron status rather than low iron status. Follow-up analyses split the sample into low iron (CRP-adjusted ferritin $\leq$ 15; n = 9), normal iron (150 > CRP-adjusted ferritin > 15; n = 36), and high iron (CRP-adjusted ferritin > 150; n = 9). In employing a one-way analysis of variance (ANOVA) to analyze differences in group means, no significant differences were found in the following outcome variables: learning outcomes (SRT Total Score, Frequency Played from Advantageous Deck, Frequency Played from Disadvantageous Deck, Median RT on Advantageous Trials, Median RT on Disadvantageous Trials, IGT Learn

Score, Accuracy and RT in A trials, Accuracy and RT in B trials), ERP components (CRN, ERN, Correct and Error FRN, and frontocentral Correct and Error P300 Amplitudes during both IGT and PST, as well as frontocentral Correct and Error P200 Amplitudes during IGT only), nor dopamine status (spontaneous and task-related BR). Lack of significance may indicate that the analyses were not powered to pick up on these differences.

The last notable limitation in our sample is the low prevalence of smoking cessation. The current rates of abstinence were comparable to previous studies of similar interventions (24% abstinent at 4-weeks post-quit in the current study compared to 20% in the parent sample from which participants were drawn) [101]. However, with such a low sample size, only 13 individuals achieved abstinence. Follow-up t-tests comparing abstinent individuals against non-abstinent individuals revealed that there were no significant differences in the same outcomes noted above (i.e. learning outcomes, ERP components, and dopamine status), as well as iron biomarkers (CRP-adjusted ferritin, CRP-adjusted ferritin percentile, sTfR, Hb, TBI, HCT).

5.2 Implication of learning outcomes on dopaminergic state

As previously stated, the basal ganglia and dopaminergic pathways are included in systems which process environmental feedback to integrate into future decision-making (i.e. learning) [102-104]. In review, the nucleus accumbens (NA) and striatum receive cortical inputs and send signals through the direct and indirect pathways of the basal ganglia [23]. The outputs of the basal ganglia, the thalamus and cortex, guide future behavior based on the inputs received and integrated previously. Notably, DA is the crucial neurotransmitter which aids in projecting cortical inputs to the NA and

striatum. Thus, dopaminergic status and learning ability are dependent on one another. Crucially, dysfunctional dopaminergic status can have significant implications on the ability to learn. Likewise, abnormalities in learning outcomes can be used to make inferences about dopaminergic dysfunction.

In the current sample, a minority of participants reached criteria to move from the learning to the test phase of the PST. This creates difficulty in interpretation of results from the PST. For example, Frank et. al., studied a sample with dopaminergic dysfunction (i.e. Parkinson's disease, which is known to have a hyperdopaminergic state in basal ganglia structures; PD); they found that in the PD group, RT during high-conflict testing trials were lower than low-conflict testing trials, while the control sample had higher RT in high-conflict trials vs. low-conflict trials. The current sample showed no significant difference in RT for high- vs. low-conflict trials, thus these low rates of meeting training criteria could be indicative of lack of learning during training phase.

Most participants learned from positive feedback but not negative feedback as exhibited by accuracy in PST and IGT outcomes. Evens et. al. conducted a meta-analysis (25 articles in total) of behavioral performance on the IGT in PD (established as hypodopaminergic state) both on and off medications; those off medications discontinued medication for period of a few days. These researchers concluded that on average, PD groups both on medication ($\beta = -8.27$, CI = [-11.27, -5.27]) and off medication ($\beta = -16.27$, CI = [-29.75, -2.79]) have lower accuracy and chose disadvantageous decks more frequently when compared with healthy controls [105]. Thus, participants in a hypodopaminergic state, with or without treatment, had poorer outcomes reflecting poor learning from negative feedback. These findings support our

hypothesis that this sample had disrupted learning; the particular pattern of disruption seems to prefer positive feedback integration, while negative feedback does not seem to be integrated. Individuals with disrupted integration of negative feedback while they are smoking may find that the positive effects of smoking (e.g. [106] demonstrates that smoking leads to phasic dopamine bursts which feels pleasurable) could outweigh the negative effects (i.e. financial cost, health impacts, social pressure, etc.), leading to lower quit rates.

Interestingly, the markers from the SRT (i.e. RT during last repeated vs. random sequence) show that implicit learning remained intact. Uddén et. al. conducted a review about the neural mechanisms guiding implicit learning [107]. They concluded that implicit learning is a dopamine-dependent function such that even “moderate” depletions of dopamine would lead to impaired implicit learning. While reward learning implicates dopaminergic function in the four major dopaminergic pathways and the direct and indirect pathways of the basal ganglia, Uddén et. al. concluded that neural mechanisms implicated in implicit learning are those which overlap between dopaminergic and serotonergic systems (i.e. serotonin-dependent dopamine release, dopamine/serotonin mutual impact on acetylcholine release, and DA and acetylcholine co-function in the striatum).

Evidence from the current sample shows that implicit learning remains intact (i.e. RT on repeated sequence decreased compared to random sequence), but reward learning is compromised (i.e. via IGT/PST). This could imply that the unique dysfunction of this sample is limited to neural mechanisms guiding reinforcement learning alone (i.e. does not spread to additional systems and functions such as serotonergic or

acetylcholine interactions). Although these results are in line with what could be expected from low iron states, which can mimic a PD-like hypodopaminergic state, previous research does not evaluate cognitive outcomes (i.e. reward and/or implicit learning) in high iron samples. However, and quite interestingly, it has been found that the gene responsible for high iron states (i.e. hemochromatosis) is a significant risk factor for the development of PD [108]. This data may suggest that both low and high iron disruptions have similar effects on learning.

5.3 Relation between iron and learning

Despite the low power, the path from iron status significantly predicted learning. The hypothesized relationship was positive (i.e. higher iron predicts higher learning scores), however the opposite was found here; the current study found that higher iron status reflected less learning. Chen et. al. detail the pathways through which an excess of iron can impact neuronal function [109]. An excess of iron can lead to oxidative stress and/or mitochondrial dysfunction (i.e. causing less energy production). Additionally, excessive iron states are associated with Parkinson's and Alzheimer's Diseases, among others (note: both excessive and depleted iron can be hallmarks of PD due to their cause of oxidative stress and hypodopaminergic states, respectively [110]).

The learning outcomes found in the current study are similar to those found in studies of Parkinson's samples [63; 78; 105]. Additionally, it is important to note that previous literature analyzing learning outcomes in relation to iron status show significant positive relations [56; 111]. On the other hand, Adreeva et. al. analyzed the relation between serum iron and markers of cognition during midlife (i.e. considering that higher age corresponds to higher iron status) found significant negative relations [112]. Other

research shows evidence that higher brain iron is associated with poorer cognition in an older age group [113-115]. Results from the current study paired alongside previous research provides evidence in support of an inverse-U (i.e. curvilinear) relationship between iron and cognition (including reward learning), such that individuals with iron status outside normative values may have compromised learning.

5.4 Implications of ERP patterns

Interestingly, the ERN and CRN amplitudes and latencies were similar in both PST and IGT tasks. The ERN component (locked to either error response or correct response), though debated in literature due to varying methodologies, can be said to be reflective of error awareness [116]. Because there were no significant differences in the signals, the implication is that participants (on average) were not aware of their error responses. For this reason, this result could be reflective of a lack of learning.

In the following frontocentral ERP components, the incorrect-locked or loss-locked trial activity was of significantly higher magnitude compared with correct-locked or win-locked trials: FRN, P300, and P200.

The present study found that the FRN had a higher magnitude amplitude (i.e. more negativity) when participants saw incorrect feedback compared to correct feedback. Initially, it was difficult to explain this neural pattern, as participants integrated positive feedback more accurately into future selections. Pfabigan et. al. studied amplitude differences in healthy controls locked to expected vs. unexpected feedback of differing valence (positive or negative) [117]. They found that the FRN amplitude was significantly higher in magnitude for ‘unexpected’ feedback conditions – irrespective of valence (i.e. whether positive or negative). Thus, it was concluded that the FRN

indicates the reward prediction error (i.e. higher when feedback is different). Other research, including Huang and Yu's study [118], Walsh and Anderson's review paper [119], and others [120; 121], confirm Pfabigan et. al.'s conclusion that the FRN reflects the difference between the actual and expected value of reward. Thus, the results from viewing incorrect feedback that elicited higher amplitude FRN amplitude suggest that this feedback elicited a greater difference between the expected and received feedback (compared with correct feedback). Thus, it can be concluded that negative feedback was not integrated at a level of expectation (i.e. learning did not take place). The results from behavioral analyses (i.e. low accuracy in negative-feedback trials) corroborate this conclusion.

With respect to the P300 during the IGT and PST tasks, there was significantly higher average amplitude when viewing negative feedback compared with positive feedback. A study conducted by Liu and Huo studied various ERPs during PST performance in a sample of healthy adults [122]. These researchers studied the relation between external feedback vs. internal monitoring (error-signaling) changes during the task, and found that there was a tradeoff in reliance of these two systems. Additionally, they found that the P300 amplitude increased over the duration of the task, a reliance on internal monitoring increased, and reliance on external feedback processing decreased. The conclusion of this study was that the P300 acted as an indicator of error-signaling. The results from the current study indicate that participants' internal monitoring was substantially higher when viewing negative feedback. If reliance on error-signaling increases, but learning has not taken place, one can conclude that

accuracy from negative feedback will not improve throughout the duration of the task. Such was the case in the current sample.

With respect to the P200 during the IGT, there was significantly higher average amplitude when viewing negative feedback compared with positive feedback. Mapelli et. al. compared feedback-evoked response in frontocentral regions during IGT in a Parkinson's disease group (PD; i.e. a hypodopaminergic state) against that of a control group. They found that healthy controls exhibited significantly higher neural amplitude in P200 components when participants won money compared to when they lost money; interestingly, the PD participants had no significant differences in P200 amplitude compared in win vs. loss conditions. Interestingly, and contrarily, Martínez-Selva et. al. studied a sample of 25 healthy females aged 20-21 years old performing the IGT task; they found that the P200 exhibited higher mean amplitude in loss trials compared to win trials [123].

Due to the inconsistency of previous literature, is not clear exactly how to interpret the results from the current study. For example, as this sample exhibits the "opposite" pattern as seen in PD, perhaps this sample mimics the "opposite" dopaminergic problem (i.e. a hyperdopaminergic state). However, literature analyzing participants with schizophrenia has found delayed and depleted P200 compared to healthy controls [124-126]. Thus, it cannot be suggested that the current sample's ERP activation necessarily mimics hyperdopaminergic state.

Notably, follow-up analyses again used ANOVA design to analyze iron group (i.e. low, normal, and high) differences in amplitudes between correct vs. incorrect trials; it was found that there were no significant differences in amplitude between groups.

However, when the samples were split into separate t-tests, the high and normal iron groups had no significant differences in P300 amplitude during IGT task, while the low iron group had a significantly higher P300 amplitude for incorrect trials compared to correct trials. This pattern was not found for P200 nor FRN components of IGT nor PST.

As an alternate explanation, this pattern of P200 activation was observed in another female-only sample [123]. Previous research has shown that there are high rates of iron deficiency among college-age females [16; 17]. Plausibly, because the current sample exhibits some iron disruption, the unique ERP pattern observed here could be driven by the changes to dopamine associated with iron disruption rather than changes to dopamine associated with nicotine use. However, this explanation cannot be concluded with certainty. To date, no studies compare differences in FRN, P300, or P200 activation during reward-feedback paradigms such as the PST or IGT in a sample of smokers vs. non-smokers or in a sample of iron deficient vs. sufficient vs. toxic.

In sum, evidence arising from the ERP components extracted in the current study support the conclusion that negative learning was impaired in the sample of middle-aged, female, adult smokers. This disruption may be, in part, caused by disruptions to iron.

5.5 Relation between DA and abstinence

A noteworthy finding of the present study is the indicators of DA. In both models with DA as a mediator, eight variables were included as potential indicators: spontaneous BR, five task-related BRs, and RT during high- vs. low-conflict PST trials. Interestingly, only the five task-related BRs loaded onto DA significantly, and all factor loadings were positive. Thus, it can be concluded that task-related blink rate can be

used as meaningful indirect measures of dopaminergic status in this sample (i.e. middle-aged, female smokers, with moderate frequency of iron disruption). Notably, previous research has established that BR is not significantly related to dopaminergic status in healthy adults without dopaminergic dysfunction [127], but has been shown to correspond to dopamine levels in samples with dopaminergic dysfunction [38; 39]. Because the five task-related BR indicators loaded significantly onto DA (see below for discussion on DA relation to abstinence), this acts as further support for dopaminergic dysfunction in the present sample.

Although the path between DA and 4-week abstinence was not statistically significant, the confidence intervals were suggestive (i.e. [-0.42, 0.09] and [-0.40, 0.10] respectively). It has been well-established that smoking will impact dopaminergic neurons [14] and dopaminergic release [i.e. phasic bursts of dopaminergic activity; 128; 129], thus this may be one explanation for the near-significant paths. Numerous research studies detail the significant impact of dopaminergic dysfunction, resulting from drug use, will reinforce continued use of the drug [130-133]. Drug use is one well-established pathway that can explain the onset of dopaminergic disruption.

However, previous research shows that iron disruption leads to disrupted dopaminergic state as well [20; 22; 50; 134]. In fact, one early study from Ben-Shachar et. al. goes as far as to show that iron chelators selectively inhibited binding of D2 receptors in some neural structures [135]. Another study from Murray-Kolb et. al. shows that iron supplementation improves behavioral function in a manner suggesting correction of dopaminergic disfunction [54]. These studies together show that iron disruptions have a significant impact on dopaminergic function, and it can be concluded

that it may lead to dopamine dysfunction. As it is established that dopaminergic dysfunction can impact recovery from addiction, this evidence suggests that iron disruption could play an additional role in continued drug use and increased difficulty in cessation.

5.6 Conclusion: Is there evidence that disrupted iron status negatively impact smoking cessation attempt? Principles to guide future research.

In this sample of female adults who are currently smoking, both neural and behavioral evidence shows a dysfunctional learning mechanism, and that this dysfunction is particularly compromising negative learning. Despite a limited sample, analyses conducted in the current study show preliminary evidence suggesting that iron can partially explain the differences in smoking cessation up to 4-weeks post-quit. More specifically, any disruption to normal iron status can have a negative impact on attempted smoking cessation.

If future studies can support this very preliminary evidence, current evidence-based standards of smoking cessation interventions [36] ignore a highly significant pathway to recovery (i.e. the impact of micronutrient [iron] status on ability to recover from addiction from nicotine). Continuing to ignore this significant biological pathway to recovery means that the females attempting smoking cessation will continue to face a significant neurobiological barrier to success. Thus, efforts should be made to continue this path of research.

Future studies should aim to extend the current methodology to a sample of iron deficient, iron sufficient, and iron toxic groups to compare outcomes (i.e. screen based on iron status). Eventually, randomized control trials (RCTs) can be employed to test an evidence-based smoking cessation intervention alongside an iron treatment protocol;

those who are iron deficient should be given supplements [136], while those who are toxic should be given iron chelators [137].

With respect to cognitive tasks and procedures, because the current sample failed to learn in the allotted number of trials in the PST, task procedures should be extended to include more trials and/or blocks. Some versions of the PST require that participants continue in the training phase until criteria is met, then move to testing phase [63]; similar pattern can and should be used in similar samples. Further, although the current study employed a version of the IGT which had double the trial number compared to similar studies [78; 105]; future studies should aim to test longer-trial tasks to determine whether similar samples do not learn (i.e. learn poorly) such as what was concluded herein, or whether they do learn eventually, just more slowly as compared with healthy controls. Lastly, the evidence of implicit learning during SRT indicates that the task procedures were robust for the current sample; no changes are suggested for this task.

References

1. Creamer, M. R., Wang, T. W., Babb, S., Cullen, K. A., Day, H., Willis, G., Jamal, A., & Neff, L. (2019, Jan 6). Tobacco Product Use and Cessation Indicators Among Adults - United States, 2018. *MMWR Morb Mortal Wkly Rep, 65*(52), 1457-1464. https://doi.org/10.15585/mmwr.mm6552a1
2. Reporter, N. (2021). *Estimates of Funding for Various Research, Condition, and Disease Categories*. https://report.nih.gov/funding/categorical-spending#/
3. CDC. (2021). *National Tobacco Control Program Funding*. https://www.cdc.gov/tobacco/stateandcommunity/tobacco-control/program-funding/index.htm
4. General, N. A. o. A. (2021). *MSA Payment Information*. https://www.naag.org/our-work/naag-center-for-tobacco-and-public-health/the-master-settlement-agreement/msa-payment-information/
5. Perkins, K. A., Donny, E., & Caggiula, A. R. (1999). Sex differences in nicotine effects and self-administration: review of human and animal evidence. *Nicotine & Tobacco Research, 1*(4), 301-315.
6. Wetter, D. W., Kenford, S. L., Smith, S. S., Fiore, M. C., Jorenby, D. E., & Baker, T. B. . (1999). Gender differences in smoking cessation. *Journal of consulting and clinical psychology, 67*(4).
7. Garey, L., Peraza, N., Smit, T., Mayorga, N. A., Neighbors, C., Raines, A. M., Schmidt, N. B., & Zvolensky, M. J. (2018, Sep). Sex differences in smoking constructs and abstinence: The explanatory role of smoking outcome expectancies. *Psychol Addict Behav, 32*(6), 660-669. https://doi.org/10.1037/adb0000391
8. Smith, P. H., Bessette, A. J., Weinberger, A. H., Sheffer, C. E., & McKee, S. A. (2016, Nov). Sex/gender differences in smoking cessation: A review. *Prev Med, 92*, 135-140. https://doi.org/10.1016/j.ypmed.2016.07.013
9. Kendzor, D. E., Businelle, M. S., Poonawalla, I. B., Cuate, E. L., Kesh, A., Rios, D. M., Ma, P., & Balis, D. S. (2015, 2015/06/01). Financial Incentives for Abstinence Among Socioeconomically Disadvantaged Individuals in Smoking Cessation Treatment. *American Journal of Public Health, 105*(6), 1198-1205. https://doi.org/10.2105/AJPH.2014.302102
10. Waters, A. F., Businelle, M. S., Frank, S. G., Hebert, E. T., & Kendzor, D. E. (2018, Sep). Understanding the link between contingency management and smoking cessation: The roles of sex and self-efficacy. *Addict Behav, 84*, 99-105. https://doi.org/10.1016/j.addbeh.2018.03.018
11. Hogle, J. M., & Curtin, J. J. (2006, Jul). Sex differences in negative affective response during nicotine withdrawal. *Psychophysiology, 43*(4), 344-356. https://doi.org/10.1111/j.1469-8986.2006.00406.x
12. Wayne, G. F., & Carpenter, C. M. (2009). Tobacco industry manipulation of nicotine dosing. *Handb Exp Pharmacol*(192), 457-485. https://doi.org/10.1007/978-3-540-69248-5_16
13. Jandikova, H., Duskova, M., & Starka, L. (2017, Sep 26). The influence of smoking and cessation on the human reproductive hormonal balance. *Physiol Res, 66*(Suppl 3), S323-S331. https://doi.org/10.33549/physiolres.933724

14. Volodymyr I. Pidoplichko, M. D., John T. Williams, John A. Dani. (1997). Nicotine activates and desensitizes midbrain dopamine neurons. *Nature, 390*, 401-404.
15. WHO. (2021). *Prevalence of anaemia in women of reproductive age (aged 15-49)*. https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-anaemia-in-women-of-reproductive-age-(-)
16. Al-Naseem, A., Sallam, A., Choudhury, S., & Thachil, J. (2021, Mar). Iron deficiency without anaemia: a diagnosis that matters. *Clin Med (Lond), 21*(2), 107-113. https://doi.org/10.7861/clinmed.2020-0582
17. Soppi, E. T. (2018, Jun). Iron deficiency without anemia - a clinical challenge. *Clin Case Rep, 6*(6), 1082-1086. https://doi.org/10.1002/ccr3.1529
18. Zacharski, L. R., Ornstein, D. L., Woloshin, S., & Schwartz, L. M. (2000, Jul). Association of age, sex, and race with body iron stores in adults: analysis of NHANES III data. *Am Heart J, 140*(1), 98-104. https://doi.org/10.1067/mhj.2000.106646
19. Zheng, W., & Monnot, A. D. (2012, Feb). Regulation of brain iron and copper homeostasis by brain barrier systems: implication in neurodegenerative diseases. *Pharmacol Ther, 133*(2), 177-188. https://doi.org/10.1016/j.pharmthera.2011.10.006
20. Burhans, M. S., Dailey, C., Beard, Z., Wiesinger, J., Murray-Kolb, L., Jones, B. C., & Beard, J. L. (2005, Feb). Iron deficiency: differential effects on monoamine transporters. *Nutr Neurosci, 8*(1), 31-38. https://doi.org/10.1080/10284150500047070
21. Felt, B. T., Beard, J. L., Schallert, T., Shao, J., Aldridge, J. W., Connor, J. R., Georgieff, M. K., & Lozoff, B. . (2006). Persistent neurochemical and behavioral abnormalities in adulthood despite early iron supplementation for perinatal iron deficiency anemia in rats. *Behavioural brain research, 171*(2), 261-270.
22. Erikson, K. M., Jones, B. C., Hess, E. J., Zhang, Q., & Beard, J. L. (2001). Iron deficiency decreases dopamine D1 and D2 receptors in rat brain. *Pharmacology biochemistry and behavior, 69*(3-4), 409-418.
23. Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J., & Mack, S. (2000). *Principles of neural science* (Vol. 4). McGraw Hill.
24. Benowitz, N. L., Bernert, J. T., Foulds, J., Hecht, S. S., Jacob, P., Jarvis, M. J., Joseph, A., Oncken, C., & Piper, M. E. (2020, Jun 12). Biochemical Verification of Tobacco Use and Abstinence: 2019 Update. *Nicotine Tob Res, 22*(7), 1086-1097. https://doi.org/10.1093/ntr/ntz132
25. American Psychiatric Association, D. S., & American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5). American psychiatric association.
26. Solinas, M., Belujon, P., Fernagut, P. O., Jaber, M., & Thiriet, N. (2019, Apr). Dopamine and addiction: what have we learned from 40 years of research. *J Neural Transm (Vienna), 126*(4), 481-516. https://doi.org/10.1007/s00702-018-1957-2
27. Ikemoto, S., Glazier, B. S., Murphy, J. M., & McBride, W. J. (1997). Role of dopamine D1 and D2 receptors in the nucleus accumbens in mediating reward. *Journal of neuroscience, 17*(21), 8580-8587.

28. Soares-Cunha, C., Coimbra, B., Sousa, N., & Rodrigues, A. J. (2016, Sep). Reappraising striatal D1- and D2-neurons in reward and aversion. *Neurosci Biobehav Rev, 68*, 370-386. https://doi.org/10.1016/j.neubiorev.2016.05.021
29. Lammel, S., Lim, B. K., & Malenka, R. C. (2014, Jan). Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology, 76 Pt B*, 351-359. https://doi.org/10.1016/j.neuropharm.2013.03.019
30. Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007, Nov 23). Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science, 318*(5854), 1309-1312. https://doi.org/10.1126/science.1146157
31. Businelle, M. S. K., Darla E.; Kesh, Anshula; Cuate, Erica L.; Poonawalla, Insiya B.; Reitzel, Lorraine R.; Okuyemi, Kolawole S.; Wetter, David W. (2014, 3//). Small financial incentives increase smoking cessation in homeless smokers: A pilot study. *Addictive Behaviors, 39*(3), 717-720. https://doi.org/http://dx.doi.org/10.1016/j.addbeh.2013.11.017
32. Gibson, R. S. (2005). Assessment of iron status. In *Principles of nutritional assessment* (2nd Edition. ed., pp. 443-453). Oxford University Press.
33. Milman, N., & Pedersen, A. N. (2009, Jul). Blood haemoglobin concentrations are higher in smokers and heavy alcohol consumers than in non-smokers and abstainers: should we adjust the reference range? *Ann Hematol, 88*(7), 687-694. https://doi.org/10.1007/s00277-008-0647-9
34. WHO. (2020). *WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations* (Licence: CC BY-NC-SA 3.0 IGO).
35. Hu, H. (2016, Jul 8). Reward and Aversion. *Annu Rev Neurosci, 39*, 297-324. https://doi.org/10.1146/annurev-neuro-070815-014106
36. Fiore, M. C., Jaen, C. R., Baker, T. B., Bailey, W. C., Benowitz, N. L., & Curry, S. J. (2008). *Treating Tobacco Use and Dependence: 2008 Update*. USDHHS. Retrieved August 14 from https://www.ncbi.nlm.nih.gov/books/NBK63952/
37. Cook, J. D., Flowers, C. H., & Skikne, B. S. (2003, May 1). The quantitative assessment of body iron. *Blood, 101*(9), 3359-3364. https://doi.org/10.1182/blood-2002-10-3071
38. Jongkees, B. J., & Colzato, L. S. (2016, Dec). Spontaneous eye blink rate as predictor of dopamine-related cognitive function-A review. *Neurosci Biobehav Rev, 71*, 58-82. https://doi.org/10.1016/j.neubiorev.2016.08.020
39. Slagter, H. A., Georgopoulou, K., & Frank, M. J. (2015, May). Spontaneous eye blink rate predicts learning from negative, but not positive, outcomes. *Neuropsychologia, 71*, 126-132. https://doi.org/10.1016/j.neuropsychologia.2015.03.028
40. Kypriotakis, G., Cinciripini, P. M., & Versace, F. (2020, Jul 15). Modeling neuroaffective biomarkers of drug addiction: A Bayesian nonparametric approach using dirichlet process mixtures. *J Neurosci Methods, 341*, 108753. https://doi.org/10.1016/j.jneumeth.2020.108753
41. Siegel, R. L., Miller, K. D., & Jemal, A. (2019, Jan). Cancer statistics, 2019. *CA Cancer J Clin, 69*(1), 7-34. https://doi.org/10.3322/caac.21551
42. CDC. (2019). *Cigarette Smoking Among U.S. Adults Hits All-Time Low* https://www.cdc.gov/media/releases/2019/p1114-smoking-

low.html#:~:text=Cigarette%20smoking%20among%20U.S.%20adults%20has%20reached%20an%20all%2Dtime,the%20health%20consequences%20of%20smoking.
43. al'Absi, M., Nakajima, M., Allen, S., Lemieux, A., & Hatsukami, D. (2015, Apr). Sex differences in hormonal responses to stress and smoking relapse: a prospective examination. *Nicotine Tob Res, 17*(4), 382-389. https://doi.org/10.1093/ntr/ntu340
44. Kudielka, B. M., & Kirschbaum, C. (2005, Apr). Sex differences in HPA axis responses to stress: a review. *Biol Psychol, 69*(1), 113-132. https://doi.org/10.1016/j.biopsycho.2004.11.009
45. Killip, S., Bennett, J. M., & Chambers, M. D. (2007). Iron Deficiency Anemia. *American Academy of Family Physicians 75*(5), 671-678.
46. Beard, J. L. (2008). Why Iron Deficiency Is Important in Infant Development. *The Journal of Nutrition, 138*, 2534-2536.
47. Georgieff, M. K., Brunette, K. E., & Tran, P. V. . (2015). Early life nutrition and neural plasticity. *Development and Psychopathology, 27*, 411-423.
48. Zheng, W., & Monnot, A. D. . (2012). Regulation of brain iron and copper homeostasis by brain barrier systems: Implication in neurodegenerative diseases. *Pharmacology & Therapeutics, 133*, 177-188.
49. Wenger, M. J., Murray-Kolb, L. E., Nevins, J. E. H., Venkatramanan, S., Reinhart, G. A.,, & Wesley, A., & Haas, J. D. (2017). Consumption of a double-fortified salt affects perceptual, attentional, and mnemonic functioning in women in a randomized controlled trial in India. *The Journal of Nutrition, 147*(12), 2297-2308.
50. Beard, J. L., Connor, J. R., & Jones, B. C. (1993). Iron in the Brain. *Nutrition Reviews, 51*, 157-170.
51. Connor, J. R., Menzies, S. L., St. Martin, S. M., Mufson, E. J. . (1990). Cellular distribution of transferrin, ferritin, and iron in normal and aged human brains. *Journal of Neuroscience Research, 27*, 595-611.
52. BENINGER, R. J., & MILLER, R. (1998). Dopamine D1-like Receptors and Reward-related Incentive Learning. *Neuroscience & Biobehavioral Reviews, 22*, 335-345.
53. Falkingham, M., Abdelhamid, A., Curtis, P., Fairweather-Tait, S., Dye, L., & Hooper, L. (2010). The effects of oral iron supplementation on cognition in older children and adults: a systematic review and meta-analysis. *Journal of Nutrition, 9*.
54. Murray-Kolb, L. E., & Beard, J. L. (2007). Iron treatment normalizes cognitive functioning in young women. *American Journal of Clinical Nutrition, 85*(3), 778-787.
55. Blanton, C. A., Green, M. W., & Kretsch, M. J. (2013, Mar 14). Body iron is associated with cognitive executive planning function in college women. *Br J Nutr, 109*(5), 906-913. https://doi.org/10.1017/S0007114512002620
56. Jenney, C. B., Alexander, D. N., Jones, B. C., Unger, E. L., & Grigson, P. S. (2016, Dec 1). Preweaning iron deficiency increases non-contingent responding during cocaine self-administration in rats. *Physiol Behav, 167*, 282-288. https://doi.org/10.1016/j.physbeh.2016.09.007

57. Baker, T. E., Zeighami, Y., Dagher, A., & Holroyd, C. B. (2020). Smoking Decisions: Altered Reinforcement Learning Signals Induced by Nicotine State. *Nicotine and Tobacco Research, 22*(2), 164-171.
58. Scott, S. P., & Murray-Kolb, L. E. (2016, Jan). Iron Status Is Associated with Performance on Executive Functioning Tasks in Nonanemic Young Women. *J Nutr, 146*(1), 30-37. https://doi.org/10.3945/jn.115.223586
59. Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.
60. Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.
61. M, K., D, B., & D, P. (2007). What's new in Psychtoolbox-3? http://psychtoolbox.org/[http://www.perceptionweb.com/abstract.cgi?id=v070821
62. The MathWorks, I. (1984-2017). *MATLAB/Simulink.* In (Version R2017b)
63. Frank, M. J., Woroch, B. S., & Curran, T. (2005, Aug 18). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501. https://doi.org/10.1016/j.neuron.2005.06.020
64. Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsoniam. *Science, 306*(5703), 1940-1943.
65. Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making. *The Journal of Neuroscience, 19*(13), 5473-5481.
66. Robertson, E. M. (2007, Sep 19). The serial reaction time task: implicit motor skill learning? *J Neurosci, 27*(38), 10073-10075. https://doi.org/10.1523/JNEUROSCI.2747-07.2007
67. Barnett, J. H., Blackwell, A. D., Sahakian, B. J., & Robbins, T. W. (2016). The Paired Associates Learning (PAL) Test: 30 Years of CANTAB Translational Neuroscience from Laboratory to Bedside in Dementia Research. *Curr Top Behav Neurosci, 28*, 449-474. https://doi.org/10.1007/7854_2015_5001
68. BrainProducts. (2000-2017). *BrainVision Recorder.* In (Version 1.21.0303)
69. BrainProducts. (2018). *actiCAP.* In [32-channel].
70. BrainProducts. (2018). *actiCHamp Plus.* In
71. Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). Chapter 11: EEG oscillations and waveleet analysis. In *Event-related potentials: A methods handbook* (pp. 229).
72. Urigüen, J. A., & Garcia-Zapirain, B. (2015). EEG artifact removal—state-of-the-art and guidelines. *Journal of neural engineering, 12*(3).
73. Delorme, A. M., S. . (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*, 9-21.
74. The MathWorks, I. (1984-2020). *MATLAB/Simulink.* In (Version R2020a)
75. Institute, T. S. (2016). *SAS.* In (Version 9.4)
76. Kleifges, K., Bigdely-Shamlo, N., Kerick, S. E., & Robbins, K. A. (2017). BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis. *Frontiers in neuroscience, 11*, 12.

77. Evans, K. L., & Hampson, E. (2015). Sex-dependent effects on tasks assessing reinforcement learning and interference inhibition. *Front Psychol, 6*, 1044. https://doi.org/10.3389/fpsyg.2015.01044
78. Mapelli, D., Di Rosa, E., Cavalletti, M., Schiff, S., & Tamburin, S. (2014). Decision and dopaminergic system: an ERPs study of Iowa gambling task in Parkinson's disease. *Front Psychol, 5*, 684. https://doi.org/10.3389/fpsyg.2014.00684
79. Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Group, P. P. C. S. (1999). Validation and Utility of a Self-report Version of PRIME-MD: The PHW Primary Care Study. *Journal of American Medical Association, 282*(18).
80. Kozlowski, L. T., Porter, C. Q., Orleans, C. T., Pope, M. A., & Heathertone, T. (1994). Predicting smoking cessation with self-reported measures of nicotine dependence: FTQ, FTND, and HSI. *Drug and Alcohol Dependence, 34*, 211-216.
81. Heatherton, Kozlowski, Frecker, Rickert, & Robinson. (1989). Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *British Journal of Addiction, 84*(7), 791-799. https://doi.org/10.1111/j.1360-0443.1989.tb03059.x
82. NIDA. (2016). *Heaviness of Smoking Index*. https://datashare.nida.nih.gov/instrument/heaviness-of-smoking-index
83. Amodeo, L. R., Jacobs-Brichford, E., McMurray, M. S., & Roitman, J. D. (2017, May 1). Acute and long-term effects of adolescent methylphenidate on decision-making and dopamine receptor mRNA expression in the orbitofrontal cortex. *Behav Brain Res, 324*, 100-108. https://doi.org/10.1016/j.bbr.2017.02.019
84. Sullivan, K. M., Mei, Z., Grummer-Strawn, L., & Parvanta, I. (2008, Oct). Haemoglobin adjustments to define anaemia. *Trop Med Int Health, 13*(10), 1267-1271. https://doi.org/10.1111/j.1365-3156.2008.02143.x
85. Skikne, B. S., Flowers, C. H., & Cook, J. D. (1990). Serum transferrin receptor: a quantitative measure of tissue iron deficiency. *Blood, 75*, 1870-1876.
86. Gartner, A., Berger, J., Bour, A., El Ati, J., Traissac, P., Landais, E., El Kabbaj, S., & Delpeuch, F. (2013, Sep). Assessment of iron deficiency in the context of the obesity epidemic: importance of correcting serum ferritin concentrations for inflammation. *Am J Clin Nutr, 98*(3), 821-826. https://doi.org/10.3945/ajcn.112.054551
87. Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*(1), 289-300.
88. Kleifges, K., Bigdely-Shamlo, N., Kerick, S. E., & Robbins, K. A. (2017). BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis. *Frontiers in neuroscience, 11*.
89. Ekhtiari, H., Kuplicki, R., Yeh, H. W., & Paulus, M. P. (2019, Jan 23). Physical characteristics not psychological state or trait characteristics predict motion during resting state fMRI. *Sci Rep, 9*(1), 419. https://doi.org/10.1038/s41598-018-36699-0
90. Polich, J. (2007, Oct). Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol, 118*(10), 2128-2148. https://doi.org/10.1016/j.clinph.2007.04.019

91. West, R., Bailey, K., & Anderson, S. (2018, Apr). Transient and sustained ERP activity related to feedback processing in the probabilistic selection task. *Int J Psychophysiol, 126*, 1-12. https://doi.org/10.1016/j.ijpsycho.2018.02.011
92. Shi, D., Shi, D., & Fairchild, A. F. (2022). Variable selection for mediators under a Bayesian mediation model. *Unknown*. Retrieved from osf.io/w9cen
93. Yuan, Y. C. (2005). *Multiple Imputation for Missing Data: Concepts and New Development (Version 9.0)*. SAS Institute Inc.
94. Buuren, S. v., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3).
95. Vink, G., Frank, L. E., Pannekoek, J., & van Buuren, S. (2014). Predictive mean matching imputation of semicontinuous variables. *Statistica Neerlandica, 68*(1), 61-90. https://doi.org/10.1111/stan.12023
96. Hwang, H., & Takane, Y. (2004). Generalized Structured Component Analysis. *Psychometrika, 69*(1).
97. Kim, C., Nan, B., Kong, S., & Harlow, S. (2012, Aug). Changes in iron measures over menopause and associations with insulin resistance. *J Womens Health (Larchmt), 21*(8), 872-877. https://doi.org/10.1089/jwh.2012.3549
98. Milman, N., Kirchhoff, M., & Jorgensen, T. (1992). Iron status markers, serum ferritin and hemoglobin in 1359 Danish women in relation to menstruation, hormonal contraception, parity, and postmenopausal hormone treatment. *Annals of Hematology, 65*, 96-102.
99. Spence, H., McNeil, C. J., & Waiter, G. D. (2020). The impact of brain iron accumulation on cognition: A systematic review. *PLoS One, 15*(10), e0240697. https://doi.org/10.1371/journal.pone.0240697
100. Hershko, C. (2007). Mechanism of Iron Toxicity. *Food and Nutrition Bulletin, 28*(4 (Supplement)).
101. Boozary, L. K., Frank-Pearce, S. G., Alexander, A. C., Waring, J. J. C., Ehlke, S. J., Businelle, M. S., Cohn, A. M., & Kendzor, D. E. (2021, Jul 1). Correlates of e-cigarette use among adults initiating smoking cessation treatment. *Drug Alcohol Depend, 224*, 108724. https://doi.org/10.1016/j.drugalcdep.2021.108724
102. Dayan, P. (2009, May). Dopamine, reinforcement learning, and addiction. *Pharmacopsychiatry, 42 Suppl 1*, S56-65. https://doi.org/10.1055/s-0028-1124107
103. Castillo Diaz, F., Caffino, L., & Fumagalli, F. (2021, Dec). Bidirectional role of dopamine in learning and memory-active forgetting. *Neurosci Biobehav Rev, 131*, 953-963. https://doi.org/10.1016/j.neubiorev.2021.10.011
104. Hyman, S. E., Malenka, R. C., & Nestler, E. J. (2006). Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu Rev Neurosci, 29*, 565-598. https://doi.org/10.1146/annurev.neuro.29.051605.113009
105. Evens, R., Hoefler, M., Biber, K., & Lueken, U. (2016, Oct). The Iowa Gambling Task in Parkinson's disease: A meta-analysis on effects of disease and medication. *Neuropsychologia, 91*, 163-172. https://doi.org/10.1016/j.neuropsychologia.2016.07.032
106. Schultz, W. (2011, Feb 24). Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron, 69*(4), 603-617. https://doi.org/10.1016/j.neuron.2011.02.014

107. Uddén, J., Folia, V., & Petersson, K. M. (2010). The Neuropharmocology of Implicit Learning. *Current Neuropharmacology, 8*, 367-381.
108. Guerreiro, R. J., Bras, J. M., Santana, I., Januario, C., Santiago, B., Morgadinho, A. S., Ribeiro, M. H., Hardy, J., Singleton, A., & Oliveira, C. (2006, Jul 6). Association of HFE common mutations with Parkinson's disease, Alzheimer's disease and mild cognitive impairment in a Portuguese cohort. *BMC Neurol, 6*, 24. https://doi.org/10.1186/1471-2377-6-24
109. Chen, P., Totten, M., Zhang, Z., Bucinca, H., Erikson, K., Santamaria, A., Bowman, A. B., & Aschner, M. (2019, Mar). Iron and manganese-related CNS toxicity: mechanisms, diagnosis and treatment. *Expert Rev Neurother, 19*(3), 243-260. https://doi.org/10.1080/14737175.2019.1581608
110. Moreau, C., Duce, J. A., Rascol, O., Devedjian, J. C., Berg, D., Dexter, D., Cabantchik, Z. I., Bush, A. I., Devos, D., & group, F.-I. s. (2018, Apr). Iron as a therapeutic target for Parkinson's disease. *Mov Disord, 33*(4), 568-574. https://doi.org/10.1002/mds.27275
111. Schmidt, A. T., Waldow, K. J., Grove, W. M., Salinas, J. A., & Georgieff, M. K. (2007, Jun). Dissociating the long-term effects of fetal/neonatal iron deficiency on three types of learning in the rat. *Behav Neurosci, 121*(3), 475-482. https://doi.org/10.1037/0735-7044.121.3.475
112. Andreeva, V. A., Galan, P., Arnaud, J., Julia, C., Hercberg, S., & Kesse-Guyot, E. (2013, Dec). Midlife iron status is inversely associated with subsequent cognitive performance, particularly in perimenopausal women. *J Nutr, 143*(12), 1974-1981. https://doi.org/10.3945/jn.113.177089
113. Daugherty, A. M., Haacke, E. M., & Raz, N. (2015, Apr 29). Striatal iron content predicts its shrinkage and changes in verbal working memory after two years in healthy adults. *J Neurosci, 35*(17), 6731-6743. https://doi.org/10.1523/JNEUROSCI.4717-14.2015
114. Daugherty, A. M., & Raz, N. (2016, Mar). Accumulation of iron in the putamen predicts its shrinkage in healthy older adults: A multi-occasion longitudinal study. *Neuroimage, 128*, 11-20. https://doi.org/10.1016/j.neuroimage.2015.12.045
115. Rodrigue, K. M., Daugherty, A. M., Foster, C. M., & Kennedy, K. M. (2020, Apr 15). Striatal iron content is linked to reduced fronto-striatal brain function under working memory load. *Neuroimage, 210*, 116544. https://doi.org/10.1016/j.neuroimage.2020.116544
116. Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Front Hum Neurosci, 6*, 88. https://doi.org/10.3389/fnhum.2012.00088
117. Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011, May). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology, 48*(5), 656-664. https://doi.org/10.1111/j.1469-8986.2010.01136.x
118. Huang, Y., & Yu, R. (2014). The feedback-related negativity reflects "more or less" prediction error in appetitive and aversive conditions. *Front Neurosci, 8*, 108. https://doi.org/10.3389/fnins.2014.00108

119. Walsh, M. M., & Anderson, J. R. (2012, Sep). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci Biobehav Rev, 36*(8), 1870-1884. https://doi.org/10.1016/j.neubiorev.2012.05.008
120. Bellebaum, C., Polezzi, D., & Daum, I. (2010, Sep). It is less than you expected: the feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia, 48*(11), 3343-3350. https://doi.org/10.1016/j.neuropsychologia.2010.07.023
121. Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., MNishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Cognitive Neuroscience and Neuropsychology, 16*(4).
122. Liu, C., & Huo, Z. (2020, Apr). A tradeoff relationship between internal monitoring and external feedback during the dynamic process of reinforcement learning. *Int J Psychophysiol, 150*, 11-19. https://doi.org/10.1016/j.ijpsycho.2020.01.004
123. Martinez-Selva, J. M., Munoz, M. A., Sanchez-Navarro, J. P., Walteros, C., & Montoya, P. (2019). Time Course of the Neural Activity Related to Behavioral Decision-Making as Revealed by Event-Related Potentials. *Front Behav Neurosci, 13*, 191. https://doi.org/10.3389/fnbeh.2019.00191
124. McCarley, R. W., Faux, S. F., Shenton, M. E., Nestor, P. G., & Adams, J. (1991). Event-related potentials in schizophrenia: their biological and clinical correlates and a new model of schizophrenic pathophysiology. *Schizophrenia Research, 4*, 209-231.
125. Ogura, C., Nageishi, Y., NMatsubayashi, M., Omura, F., Kishimoto, A., & Shimokochi, M. (1991). Abnormalities in Event-Related Potentials, N100, P200, P300, and Slow Wave in Schizophrenia. *The Japanese Journal of Psychiatry and Neurology, 45*(1), 57-65.
126. Kayser, J., Tenke, C. E., Gil, R. B., & Bruder, G. E. (2009, Sep). Stimulus- and response-locked neuronal generator patterns of auditory and visual word recognition memory in schizophrenia. *Int J Psychophysiol, 73*(3), 186-206. https://doi.org/10.1016/j.ijpsycho.2009.02.003
127. Dang, L. C., Samanez-Larkin, G. R., Castrellon, J. J., Perkins, S. F., Cowan, R. L., Newhouse, P. A., & Zald, D. H. (2017, Sep-Oct). Spontaneous Eye Blink Rate (EBR) Is Uncorrelated with Dopamine D2 Receptor Availability and Unmodulated by Dopamine Agonism in Healthy Adults. *eNeuro, 4*(5). https://doi.org/10.1523/ENEURO.0211-17.2017
128. Brody, A. L., Olmstead, R. E., London, E. D., Farahi, J., Meyer, J. H., Grossman, P., Lee, G. S., Huang, J., Hahn, E. L., & Mandelkern, M. A. (2004). Smoking-induced Ventral Striatum Dopamine Release. *American Journal of Psychiatry, 161*, 1211-1218.
129. Volkow, N. D., Wang, G. J., Fowler, J. S., Tomasi, D., & Telang, F. (2011, Sep 13). Addiction: beyond dopamine reward circuitry. *Proc Natl Acad Sci U S A, 108*(37), 15037-15042. https://doi.org/10.1073/pnas.1010654108
130. D.Volkow, N., Fowler, J. S., & Gene-JackWang. (1999). Imaging studies on the role of dopamine in cocaine reinforcement and addiction in humans. *Journal of Psychopharmacology, 13*(4), 337-345.

131. Volkow, N. D., Fowler, J. S., Wang, G. J., Baler, R., & Telang, F. (2009). Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology, 56 Suppl 1*, 3-8. https://doi.org/10.1016/j.neuropharm.2008.05.022
132. Johnson, P. M., & Kenny, P. J. (2010, May). Dopamine D2 receptors in addiction-like reward dysfunction and compulsive eating in obese rats. *Nat Neurosci, 13*(5), 635-641. https://doi.org/10.1038/nn.2519
133. Ashok, A. H., Mizuno, Y., & Howes, O. D. (2019, Apr). Tobacco smoking and dopaminergic function in humans: a meta-analysis of molecular imaging studies. *Psychopharmacology (Berl), 236*(4), 1119-1129. https://doi.org/10.1007/s00213-019-05196-1
134. Lozoff, B. (2011, Apr 1). Early iron deficiency has brain and behavior effects consistent with dopaminergic dysfunction. *J Nutr, 141*(4), 740S-746S. https://doi.org/10.3945/jn.110.131169
135. Ben-Shachar, D., Finberg, J. P. M., & Youdim, M. B. H. (1985). Effect of Iron Chelators on Dopamine D2 Receptors. *Journal of Neurochemistry, 45*, 999-1005.
136. Murray-Kolb, L. E., Wenger, M. J., Scott, S. P., Rhoten, S. E., Lung'aho, M. G., & Haas, J. D. (2017, Nov). Consumption of Iron-Biofortified Beans Positively Affects Cognitive Performance in 18- to 27-Year-Old Rwandan Female College Students in an 18-Week Randomized Controlled Efficacy Trial. *J Nutr, 147*(11), 2109-2117. https://doi.org/10.3945/jn.117.255356
137. Poggiali, E., Cassinerio, E., Zanaboni, L., & Cappellini, M. D. (2012, Oct). An update on iron chelation therapy. *Blood Transfus, 10*(4), 411-422. https://doi.org/10.2450/2012.0008-12
138. Blumenfeld, H. (2010). *Neuroanatomy through clinical cases* (2nd Edition ed.). Sinauer Associates.

Table 1: Baseline Sociodemographic Characteristics

CHARACTERISTIC	MEAN	MEDIAN	SD	RANGE	N
Age (years)	51	54	12	20-70	54
Education (years)	13	13	2	8-20	54
Body Mass Index (BMI)	27.84	28	4.43	18.4-34.8	54
Years Smoking	27	26	14	4-50	52
Cigarettes per day (CPD)	16.34	15	11.32	2-60	41
CHARACTERISTIC	**PERCENT**	**FREQUENCY**	**N**		
Race			54		
White	51.85	28			
Black	27.78	15			
Hispanic	9.26	5			
More than one race	5.56	3			
American Indian/Alaska Native	3.70	2			
Native Hawaiian or Other Pacific Islander	1.85	1			
Asian	0.00	0			
Participated in Clinical Trial (% yes)	22.22	12	54		
Education (≥HS)	90.74	49	54		
Income			52		
0$ TO 10,999$	25.00	13			
11,000 $ TO 20,9999 $	28.85	15			
21,000 $ TO 30,999 $	9.62	5			
31,000 $ TO 40,999 $	7.69	4			
41,000 $ TO 5,999 $	9.62	5			
61,000 $ TO 70,999 $	7.69	4			
71,000 $ TO 80,999 $	3.85	2			
91,000 $ TO 100,000 $	1.92	1			
≥ 100,000 $	5.77	3			
Marital Status			54		
Single	35.19	19			
Married	29.63	16			
Divorced	22.22	12			
Separated	5.56	3			
Living with Significant Other	3.70	2			
Widowed	3.70	2			
Depression, % yes	9.26	5	54		

Table 1: Continued

CHARACTERISTIC	PERCENT	FREQUENCY	N		
Menopause Status			52		
Pre-menopause	29.00	15			
Peri-menopause	12.00	6			
Post-menopause	60.00	31			
CPD (Categorical)			54		
1-5	20.37	11			
6-10	18.52	10			
11-15	16.67	9			
16-20	27.78	15			
21-25	7.41	4			
26-30	3.70	2			
≥31	5.56	3			
HSI			52		
0	5.77	3			
1	9.62	5			
2	19.23	10			
3	26.92	14			
4	23.08	12			
5	9.62	5			
6	5.77	3			

Table 2: Baseline Biomarker Characteristics

Biomarker	**Mean**	**SD**	**Median**	**RANGE**	**N**
Ferritin	103.12	97.88	75.70	5.0-524.0	51
CRP	6.99	13.68	3.00	0.1-88.1	51
CRP- outlier removed	5.37	7.35	3.00	0.1-30.6	50
CRP-Adjusted Ferritin	76.14	68.66	52.07	3.3-340.6	51
Ferritin Percentile	46.33	28.92	43.60	0.5-93.4	51
CRP-Adjusted Ferritin Percentile	39.84	26.72	36.40	0.4-93.8	51
sTfR	18.24	7.31	16.50	9.3-51.8	49
Hb	14.04	1.47	14.00	8.4-17.0	49
Cigarettes per Day Adjusted Hb 1*	13.82	1.46	13.80	8.4-16.5	49
Cigarettes per Day Adjusted Hb 2**	13.76	1.47	13.90	8.4-16.0	37
Total Body Iron‡	28.03	4.57	28.67	15.0- 35.6	49
Total Body Iron‡‡	27.12	4.38	27.42	13.4-34.1	49
Thresholds	**FREQUENCY**	**PERCENT**			
Inflammation	29	58.00			
Iron Toxicity‡	11	21.57			
Iron Toxicity‡‡	6	11.76			
Iron Deficiency Anemia	2	4.08			
Iron Deficiency Anemia*	3	6.12			
Iron Deficiency Anemia**	3	8.11			

sTfR = Serum Transferrin Receptor; Hb = Hemoglobin

*Use categorical CPD (which necessitated estimation for correction).

**Used raw CPD (which had high levels of missingness).

‡ Used raw Ferritin in calculation

‡‡ Used CRP-adjusted Ferritin in calculation

Table 3: Follow-Up Smoking Behavior

	Baseline		Quit Day		Week1		Week2	
	M ± SD	MED	M ± SD	MED	M ± SD	MED	M ± SD	MED
Expired CO	22.15 ± 14.46	20	11.71 ± 8.18	9.5	7.7 ± 9.88	4	8.47 ± 12.16	6
	FREQ	%	FREQ	%	FREQ	%	FREQ	%
Abstinence	--	--	22	40.7	4	7.41	5	9.26
SR Abstinence	--	--	17	31.5	13	24.07	11	20.37
Alternative Tobacco or Smoking Product Use (SR Only)								
SR ATP Use	17	31.48	4	8.89	2	6.06	0	0
SR EC Use	9	16.6	4	8.89	3	9.09	1	3.33
SR Cannabis	0	0	7	13	6	11.11	4	7.41
No-Show	--	--	12	22.22	18	33.33	19	35.19

SR = self-reported

Table 3: Continued

	Week3		Week4		Week12		Week26	
	M ± SD	MED	M ± SD	MED	M ± SD	MED	M ± SD	MED
Expired CO	8.16 ± 8.91	5	9.33 ± 10.08	6	10.17 ± 11.22	6	13.33 ± 11.46	10
	FREQ	%	FREQ	%	FREQ	%	FREQ	%
Abstinence	7	12.96	13	24.07	11	20.37	7	12.96
SR Abstinence	12	22.22	17	31.48	16	29.63	11	20.37
Alternative Tobacco or Smoking Product Use (SR Only)								
SR ATP Use	1	3.23	3	6.98	2	6.25	1	3.45
SR EC Use	2	6.45	2	4.65	3	13.04	1	4.35
SR Cannabis	4	7.41	5	9.26	4	7.41	3	5.56
No-Show	17	31.48	10	18.52	22	40.74	25	46.3

SR = self-reported

Table 4: Description of Blink Rates (BR; blinks per minute) during Cognitive Tasks

Task	Mean	Median	SD	Range		N
Spontaneous BR (Rest)	6.41	4.8	5.86	0.4	30.8	49
Task-related BR (PST Test)	10.39	8.03	7.96	0.45	29.37	43
Task-related BR (PST Train)	10.37	8.55	8.17	0.66	33.14	45
Task-related BR (IGT)	9.23	8.05	6.98	0.29	32.51	46
Task-related BR (SRT)	5.55	3.42	5.06	0.25	20.73	45
Task-related BR (PAL)	11.12	8.02	10.04	0.75	42.71	48

Table 5: Behavior Outcomes from all Cognitive Tasks

Outcome	Mean	Median	SD	Range	
PAL Behavior Outcomes					
Accuracy Block1	4.30	5	1.56	1	6
Accuracy Block6	4.87	5	1.41	0	6
RT - Block1	3087	2804	1283	1057	6726
RT - Block6	2458	2157	981	929	5605
Difference Score	-0.57	0	1.83	-4	2
SRT Behavior Outcomes					
Last Repeat Sequence Median RT	568	549	165	319	1078
Last Random Sequence Median RT	616	578	170	374	1129
SRT Total Score	48.28	49.75	118.98	-306	415
SRT Slope	-1.39	-1.30	6.15	-33.94	15.53
IGT Behavior Outcomes					
Played from Advantageous Deck	91.25	91	20.67	46	149
Played from Disadvantageous Deck	108.75	109	20.67	51	154
Median RT - Advantageous Trials	862	794	392	370	2168
Median RT - Disadvantageous Trials	866	754	397	397	2114
IGT Learn Score	-17.51	-18	41.34	-108.00	98.00
IGT Learning Slope (10 Blocks)	0.21	0.11	0.91	-1.66	2.82
PST Behavior Outcomes					
A pairs					
Accuracy	64.14	65	21.86	15	100
RT - Incorrect	1328	1152	841	380	5266
RT - Correct	1210	1132	653	395	3889
B pairs					
Accuracy	44.78	45	17.31	20	90
RT - Incorrect	1375	1157	940	314	5887
RT - Correct	1309	1110	697	289	3290
Conflict					
High Conflict RT	1369	1247	680	300	4162
Low Conflict RT	1311	1117	767	321	4718
PST Criteria Thresholds	**Frequency**		**Percent**		
Reached criteria					
AB only	33		64.71		
CD only	30		58.82		
EF only	40		78.43		
All pair	21		41.18		

Table 6: ERP Components Descriptive Statistics for EEG Data

ERP Component	Mean	Median	SD	Range	
PST					
Correct					
CRN Amplitude	-0.49	-0.45	0.25	-1.09	-0.09
Latency to CRN	65	52	22	50	128
FRN Amplitude	-2.12	-2.01	1.52	-6.32	0.00
FRN Amplitude (Outliers Removed)	-2.23	-2.14	1.47	-6.32	-0.05
Latency to FRN	245	240	33	190	300
Latency to FRN (Outliers Removed)	248	242	31	190	300
FC P300 Amplitude	0.77	0.88	1.21	-2.16	3.60
CP P300 Amplitude	2.59	1.93	1.82	0.09	6.19
Incorrect					
ERN Amplitude	-0.52	-0.46	0.30	-1.36	-0.07
Latency to ERN	66	58	21	50	126
FRN Amplitude	-2.90	-2.75	1.76	-8.33	-0.42
Latency to FRN	256	252	30	192	300
FC P300 Amplitude	1.53	1.35	1.69	-2.59	5.40
CP P300 Amplitude	2.29	2.43	2.29	-1.77	7.69
IGT					
Advantageous					
CRN Amplitude	-0.5	-0.38	0.34	-1.61	-0.08
Latency to CRN	60	50	16	50	116
FRN Amplitude	-3.02	-2.85	1.74	-8.16	0.00
Latency to FRN	287	298	22	190	300
SPN	-0.15	-0.14	0.13	-0.43	0.16
Fronto-central					
P300 Amplitude	0.70	0.62	0.97	-0.97	3.11
P200 Amplitude	2.36	2.34	1.26	0.11	6.09
Centro-parietal					
P300 Amplitude	3.54	3.61	2.10	0.58	8.87
P200 Amplitude	0.88	1.02	1.53	-3.70	3.42
Disadvantageous					
ERN Amplitude	-0.48	-0.42	0.31	-1.38	-0.06
Latency to ERN	64	5	20	50	126
FRN Amplitude	-3.33	-3.18	1.76	-8.20	-0.62
Latency to FRN	287	296	23	194	300

Table 7: Smoking, Behavioral, and Neural Comparisons

Comparison	t-value	DF	95% CI		*p*	*p*-adjust
Post-hoc Smoking Outcomes (CO)						
Week 0 - Week 4	6.66	33	--		**< .0001**	**< .0001**
Week 0 - Week 12	3.37	21			0.003	**0.009**
Week 0 - Week 26	1.11	19			0.28	0.42
Week 4 - Week 12	-0.65	23			0.52	0.53
Week 4 - Week 26	-2.13	20			0.05	0.09
Week 12 - Week 26	-0.64	15			0.53	0.53
PST						
High vs. Low conflict RTs	-2.14	50	-95.41	-3.08	**0.04**	0.08
A vs. B Accuracy	5.40	50	13.02	28.57	**< .0001**	0.04
Cor vs. Inc RT - A	-1.96	46	-148.61	2.07	0.06	0.08
Cor vs. Inc RT - B	-0.11	50	-81.01	72.64	0.91	0.91
SRT						
Last Repeated vs. Random RT	-2.90	51	-81.40	-15.15	**0.01**	**-**
IGT						
Adv vs. Disadv Frequency	-3.10	52	-28.90	-6.12	**0.003**	**0.03**
Adv vs. Disadv RT	-0.11		-58.22	51.93	0.90	0.90
PST Neural Data						
CRN vs. ERN Amplitude	0.60	38	-0.08	0.14	0.55	0.63
CRN vs. ERN Latency	-0.33	38	-7.60	5.44	0.74	0.74
Cor vs. Inc FRN Amplitude	4.32	38	0.41	1.14	**<0.001**	**0.03**
Cor vs. Inc FRN Amplitude - Outliers Removed	4.03	36	0.38	1.14	**<0.001**	**0.03**
Cor vs. Inc FRN Latency	-1.89	38	-23.71	0.84	0.07	0.14
Cor vs. Inc FRN Latency - Outliers Removed	-1.49	36	-21.58	3.31	0.15	0.24
Cor vs. Inc FC P300 Amplitude	-5.25	38	-1.05	-0.46	**<0.0001**	**0.03**
Cor vs. Inc CP P300 Amplitude	1.38	38	-0.14	0.75	0.18	0.24
IGT Neural Data						
CRN vs. ERN Amplitude	-0.05	37	-0.12	0.12	0.96	0.96
CRN vs. ERN Latency	-1.17	37	-11.35	3.03	0.25	0.56
Adv vs. Disadv FRN Amplitude	-2.44	38	-0.55	-0.05	**0.02**	0.06
Adv vs. Disadv FRN Latency	-0.24	38	-11.12	8.76	0.81	0.96
Adv vs. Disadv FC P300 Amplitude	-2.72	39	-0.43	-0.06	**<0.01**	0.06
Adv vs. Disadv CP P300 Amplitude	0.74	39	-0.20	0.42	0.47	0.74
Adv vs. Disadv FC P200 Amplitude	-2.35	39	-0.27	-0.02	**0.02**	0.06
Adv vs. Disadv CP P200 Amplitude	0.69	39	-0.12	0.25	0.49	0.74
Adv vs. Disadv SPN Amplitude	-0.16	39	-0.04	0.03	0.87	0.96

Cor = Correct; Inc = Incorrect; Adv = Advantageous; Disadv = Disadvantageous

Table 8: Indices of GSCA Model Fit

	MODEL	FIT	AFIT	GFI	SRMR
	4-week				
Model1	*Dopamine*	0.342	0.313	0.913	0.098
Model2	*Learn*	0.369	0.342	0.932	0.091
Model3	*Neural Activity*	0.344	0.316	0.863	0.129
	12-week				
Model4	*Dopamine*	0.346	0.317	0.915	0.097
Model5	*Learn*	0.367	0.34	0.928	0.094
Model6	*Neural Activity*	0.344	0.316	0.864	0.128

FIT = total variance of all variables accounted for by model
AFIT = Adjusted FIT (controls for model complexity)
GFI = Goodness-of-Fit-Index
SRMR = Standardized Root Mean Squared Residual

Table 9: Correlation Matrix for Model in which Dopamine mediates the Relation between Iron and Abstinence at 4-weeks post-quit

	sTfR	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb
sTfR	1							
RBC	-0.07	1						
HCT	-0.13	0.83	1					
MCV	-0.02	-0.40	0.09	1				
MCHC	-0.25	-0.06	-0.11	-0.23	1			
RDW	0.31	0.02	-0.05	-0.06	-0.49	1		
Ferritin Percentile	-0.17	-0.17	-0.15	0.05	0.38	-0.30	1	
Hb	-0.32	0.77	0.86	-0.08	0.36	-0.27	0.05	1
RT Low-Conflict (PST)	0.11	-0.29	-0.04	0.31	-0.12	-0.06	0.07	-0.19
RT High-Conflict (PST)	0.10	-0.22	-0.03	0.30	-0.17	0.00	0.04	-0.18
Resting State BR	0.08	-0.19	-0.27	-0.13	0.12	-0.21	0.27	-0.21
PAL BR	-0.03	-0.24	-0.12	0.15	0.14	-0.18	0.32	-0.04
IGT BR	0.05	0.04	-0.03	-0.17	0.04	-0.02	0.13	0.00
SRT BR	0.32	-0.04	0.05	0.08	-0.13	-0.04	0.03	-0.01
PST BR (Training)	0.21	-0.07	-0.08	-0.07	0.07	0.02	0.12	0.02
PST BR (Testing)	0.10	0.03	-0.02	-0.19	0.00	0.09	0.21	-0.02
4-wk Abs	-0.18	-0.04	0.09	0.13	0.02	-0.09	0.05	0.08

Abbreviations in Tables 9-14:
sTfr = serum transferrin receptor; RBC = red blood cells
HCT = hematocrit; MCV = mean corpuscular volume
MCHC = mean corpuscular hemoglobin concentration
RDW = Red cell distribution width
Ferritin Percentile = CRP-adjusted Ferritin Percentile Rank
Hb = hemoglobin; 4-wk Abs = Abstinence at 4-week post-quit
Cor = Correct; Inc = Incorrect

Table 9: Continued

	RT Low-Conflict (PST)	RT High-Conflict (PST)	Resting State BR	PAL BR	IGT BR	SRT BR	PST BR (Train)	PST BR (Test)	4-wk Abs
sTfR									
RBC									
HCT									
MCV									
MCHC									
RDW									
Ferritin Percentile									
Hb									
RT Low-Conflict (PST)	1								
RT High-Conflict (PST)	0.83	1							
Resting State BR	0.12	0.07	1						
PAL BR	0.05	-0.06	0.38	1					
IGT BR	0.02	0.04	0.43	0.59	1				
SRT BR	-0.06	-0.06	0.19	0.41	0.48	1			
PST BR (Training)	-0.05	-0.06	0.13	0.39	0.55	0.66	1		
PST BR (Testing)	-0.15	-0.16	0.07	0.07	0.30	0.56	0.53	1	
4-wk Abs	0.06	0.08	-0.03	-0.08	0.04	-0.18	-0.13	-0.17	1

Table 10: Correlation Matrix for Model in which Learning mediates the Relation between Iron and Abstinence at 4 weeks

	sTfR	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb
sTfR	1							
RBC	-0.07	1						
HCT	-0.13	0.83	1					
MCV	-0.02	-0.40	0.09	1				
MCHC	-0.25	-0.06	-0.11	-0.23	1			
RDW	0.31	0.02	-0.05	-0.06	-0.49	1		
Ferritin Percentile	-0.17	-0.17	-0.15	0.05	0.38	-0.30	1	
Hb	-0.32	0.77	0.86	-0.08	0.36	-0.27	0.05	1
RT for A trials (Inc)	0.13	-0.26	-0.08	0.22	-0.24	0.05	-0.10	-0.23
RT for A trials (Cor)	0.07	-0.35	-0.06	0.40	-0.10	-0.03	0.07	-0.22
RT for B trials (Inc)	0.09	-0.33	-0.13	0.22	-0.11	0.03	0.03	-0.25
Percent Accuracy (A trials)	-0.01	-0.24	-0.20	-0.06	-0.06	0.14	-0.08	-0.24
Percent Accuracy (B trials)	0.21	0.03	-0.05	-0.16	-0.27	0.46	-0.11	-0.17
Met Criteria (yes/no)	-0.16	-0.17	-0.06	0.19	0.04	0.16	0.13	-0.06
4-wk Abs	-0.18	-0.04	0.09	0.13	0.02	-0.09	0.05	0.08

Table 10: Continued

	RT for A trials (Inc)	RT for A trials (Cor)	RT for B trials (Inc)	Percent Accuracy (A trials)	Percent Accuracy (B trials)	Met Criteria (yes/no)	4-wk Abs
sTfR							
RBC							
HCT							
MCV							
MCHC							
RDW							
Ferritin Percentile							
Hb							
RT for A trials (Inc)	1						
RT for A trials (Cor)	0.75	1					
RT for B trials (Inc)	0.78	0.79	1				
Percent Accuracy (A trials)	0.29	0.18	0.36	1			
Percent Accuracy (B trials)	-0.11	-0.15	-0.14	0.05	1		
Met Criteria (yes/no)	0.05	0.12	0.01	0.28	0.02	1	
4-wk Abs	0.11	0.10	0.14	0.17	-0.01	0.06	1

Table 11: Correlation Matrix for Model in which Neural Activation mediates the Relation between Iron and Abstinence at 4 weeks

	Inc P200 (IGT; fc)	Cor FRN (IGT)	Inc FRN (IGT)	Cor P200 (IGT; fc)	Cor FRN (PST)	Inc FRN (PST)	Inc P300 (PST; fc)
Inc P200 (IGT; fc)	1						
Cor FRN (IGT)	-0.43	1					
Inc FRN (IGT)	-0.47	0.73	1				
Cor P200 (IGT; fc)	0.79	-0.24	-0.23	1			
Cor FRN (PST)	-0.28	0.33	0.35	-0.20	1		
Inc FRN (PST)	0.06	0.04	-0.06	-0.01	0.57	1	
Inc P300 (PST; fc)	-0.09	0.19	0.09	0.18	0.21	-0.06	1
sTfR	-0.34	0.43	0.34	-0.24	-0.10	-0.25	-0.17
RBC	0.14	-0.13	0.18	0.17	-0.06	-0.16	-0.02
HCT	-0.05	-0.04	0.26	0.08	0.02	-0.19	-0.06
MCV	-0.39	0.25	0.18	-0.19	0.22	-0.12	0.06
MCHC	0.31	-0.47	-0.60	0.14	-0.10	0.12	0.03
RDW	-0.11	0.25	0.24	-0.07	-0.16	-0.18	0.07
Ferritin Percentile	-0.08	0.13	-0.23	0.02	0.21	0.18	0.35
Hb	0.12	-0.27	-0.09	0.12	-0.06	-0.14	-0.04
4-wk Abs	0.06	-0.07	0.03	0.09	0.07	0.22	-0.15

Table 11: Continued

	sTfR	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb	4-wk Abs
Inc P200 (IGT; fc)									
Cor FRN (IGT)									
Inc FRN (IGT)									
Cor P200 (IGT; fc)									
Cor FRN (PST)									
Inc FRN (PST)									
Inc P300 (PST; fc)									
sTfR	1								
RBC	-0.07	1							
HCT	-0.13	0.83	1						
MCV	-0.02	-0.40	0.09	1					
MCHC	-0.25	-0.06	-0.11	-0.23	1				
RDW	0.31	0.02	-0.05	-0.06	-0.49	1			
Ferritin Percentile	-0.17	-0.17	-0.15	0.05	0.38	-0.30	1		
Hb	-0.32	0.77	0.86	-0.08	0.36	-0.27	0.05	1	
4-wk Abs	-0.18	-0.04	0.09	0.13	0.02	-0.09	0.05	0.08	1

Table 12: Correlation Matrix for Model in which Dopamine mediates the Relation between Iron and Abstinence- at 12 weeks

	sTfR	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb
sTfR	1							
RBC	-0.07	1						
HCT	-0.13	0.83	1					
MCV	-0.02	-0.40	0.09	1				
MCHC	-0.25	-0.06	-0.11	-0.23	1			
RDW	0.31	0.02	-0.05	-0.06	-0.49	1		
Ferritin Percentile	-0.17	-0.17	-0.15	0.05	0.38	-0.30	1	
Hb	-0.32	0.77	0.86	-0.08	0.36	-0.27	0.05	1
RT Low-Conflict (PST)	0.11	-0.29	-0.04	0.31	-0.12	-0.06	0.07	-0.19
RT High-Conflict (PST)	0.10	-0.22	-0.03	0.30	-0.17	0.00	0.04	-0.18
Resting State BR	0.08	-0.19	-0.27	-0.13	0.12	-0.21	0.27	-0.21
PAL BR	-0.03	-0.24	-0.12	0.15	0.14	-0.18	0.32	-0.04
IGT BR	0.05	0.04	-0.03	-0.17	0.04	-0.02	0.13	0.00
SRT BR	0.32	-0.04	0.05	0.08	-0.13	-0.04	0.03	-0.01
PST BR (Training)	0.21	-0.07	-0.08	-0.07	0.07	0.02	0.12	0.02
PST BR (Testing)	0.10	0.03	-0.02	-0.19	0.00	0.09	0.21	-0.02
12-wk Abs	-0.19	-0.12	-0.04	0.02	0.10	0.06	0.02	0.05

Table 12: Continued

	RT Low-Conflict (PST)	RT High-Conflict (PST)	Resting State BR	PAL BR	IGT BR	SRT BR	PST BR (Train)	PST BR (Test)	12-wk Abs
sTfR									
RBC									
HCT									
MCV									
MCHC									
RDW									
Ferritin Percentile									
Hb									
RT Low-Conflict (PST)	1								
RT High-Conflict (PST)	0.83	1							
Resting State BR	0.12	0.07	1						
PAL BR	0.05	-0.06	0.38	1					
IGT BR	0.02	0.04	0.43	0.59	1				
SRT BR	-0.06	-0.06	0.19	0.41	0.48	1			
PST BR (Training)	-0.05	-0.06	0.13	0.39	0.55	0.66	1		
PST BR (Testing)	-0.15	-0.16	0.07	0.07	0.30	0.56	0.53	1	
12-wk Abs	-0.07	0.03	0.01	0.16	0.15	0.19	0.26	0.21	1

Table 13: Correlation Matrix for Model in which Learning mediates the Relation between Iron and Abstinence at 12 weeks

	sTfR	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb
sTfR	1							
RBC	-0.07	1						
HCT	-0.13	0.83	1					
MCV	-0.02	-0.40	0.09	1				
MCHC	-0.25	-0.06	-0.11	-0.23	1			
RDW	0.31	0.02	-0.05	-0.06	-0.49	1		
Ferritin Percentile	-0.17	-0.17	-0.15	0.05	0.38	-0.30	1	
Hb	-0.32	0.77	0.86	-0.08	0.36	-0.27	0.05	1
RT for A trials (Inc)	0.13	-0.26	-0.08	0.22	-0.24	0.05	-0.10	-0.23
RT for A trials (Cor)	0.07	-0.35	-0.06	0.40	-0.10	-0.03	0.07	-0.22
RT for B trials (Inc)	0.09	-0.33	-0.13	0.22	-0.11	0.03	0.03	-0.25
Percent Accuracy (A trials)	-0.01	-0.24	-0.20	-0.06	-0.06	0.14	-0.08	-0.24
Percent Accuracy (B trials)	0.21	0.03	-0.05	-0.16	-0.27	0.46	-0.11	-0.17
Met Criteria (yes/no)	-0.16	-0.17	-0.06	0.19	0.04	0.16	0.13	-0.06
12-wk Abs	-0.19	-0.12	-0.04	0.02	0.10	0.06	0.02	0.05

Table 13: Continued

	RT for A trials (Inc)	RT for A trials (Cor)	RT for B trials (Inc)	Percent Accuracy (A trials)	Percent Accuracy (B trials)	Met Criteria (yes/no)	12-wk Abs
sTfR							
RBC							
HCT							
MCV							
MCHC							
RDW							
Ferritin Percentile							
Hb							
RT for A trials (Inc)	1						
RT for A trials (Cor)	0.75	1					
RT for B trials (Inc)	0.78	0.79	1				
Percent Accuracy (A trials)	0.29	0.18	0.36	1			
Percent Accuracy (B trials)	-0.11	-0.15	-0.14	0.05	1		
Met Criteria (yes/no)	0.05	0.12	0.01	0.28	0.02	1	
12-wk Abs	0.01	0.06	0.00	0.04	0.14	0.24	1

Table 14: Correlation Matrix for Model in which Neural Activation mediates the Relation between Iron and Abstinence at 12 weeks

	Inc P200 (IGT; fc)	Cor FRN (IGT)	Inc FRN (IGT)	Cor P200 (IGT; fc)	Cor FRN (PST)	Inc FRN (PST)	Inc P300 (PST; fc)	sTfR
Inc P200 (IGT; fc)	1							
Cor FRN (IGT)	-0.43	1						
Inc FRN (IGT)	-0.47	0.73	1					
Cor P200 (IGT; fc)	0.79	-0.24	-0.23	1				
Cor FRN (PST)	-0.28	0.33	0.35	-0.20	1			
Inc FRN (PST)	0.06	0.04	-0.06	-0.01	0.57	1		
Inc P300 (PST; fc)	-0.09	0.19	0.09	0.18	0.21	-0.06	1	
sTfR	-0.34	0.43	0.34	-0.24	-0.10	-0.25	-0.17	1
RBC	0.14	-0.13	0.18	0.17	-0.06	-0.16	-0.02	-0.07
HCT	-0.05	-0.04	0.26	0.08	0.02	-0.19	-0.06	-0.13
MCV	-0.39	0.25	0.18	-0.19	0.22	-0.12	0.06	-0.02
MCHC	0.31	-0.47	-0.60	0.14	-0.10	0.12	0.03	-0.25
RDW	-0.11	0.25	0.24	-0.07	-0.16	-0.18	0.07	0.31
Ferritin Percentile	-0.08	0.13	-0.23	0.02	0.21	0.18	0.35	-0.17
Hb	0.12	-0.27	-0.09	0.12	-0.06	-0.14	-0.04	-0.32
12-wk Abs	0.05	-0.13	-0.20	-0.06	0.05	-0.08	0.15	-0.19

Table 14: Continued

	RBC	HCT	MCV	MCHC	RDW	Ferritin Percentile	Hb	12-wk Abs
Inc P200 (IGT; fc)								
Cor FRN (IGT)								
Inc FRN (IGT)								
Cor P200 (IGT; fc)								
Cor FRN (PST)								
Inc FRN (PST)								
Inc P300 (PST; fc)								
sTfR								
RBC	1							
HCT	0.83	1						
MCV	-0.40	0.09	1					
MCHC	-0.06	-0.11	-0.23	1				
RDW	0.02	-0.05	-0.06	-0.49	1			
Ferritin Percentile	-0.17	-0.15	0.05	0.38	-0.30	1		
Hb	0.77	0.86	-0.08	0.36	-0.27	0.05	1	
12-wk Abs	-0.12	-0.04	0.02	0.10	0.06	0.02	0.05	1

Table 15: Factor Loadings in GSCA Models (DA)

4-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.366	0.21	-0.64	0.21
RBC	**0.864**	**0.11**	**0.489**	**0.97**
HCT	**0.878**	**0.09**	**0.612**	**0.96**
MCV	-0.198	0.23	-0.6	0.31
MCHC	0.267	0.29	-0.37	0.69
RDW	-0.294	0.27	-0.66	0.36
Ferritin Percentile	0.005	0.31	-0.55	0.51
Hb	**0.967**	**0.06**	**0.732**	**0.98**
DA				
RT Low-Conflict (PST)	-0.093	0.33	-0.64	0.57
RT High-Conflict (PST)	-0.125	0.34	-0.6	0.63
Resting State BR	0.437	0.2	-0.07	0.68
PAL BR	**0.664**	**0.16**	**0.218**	**0.81**
IGT BR	**0.78**	**0.12**	**0.449**	**0.88**
SRT BR	**0.823**	**0.05**	**0.708**	**0.89**
PST BR (Training)	**0.817**	**0.06**	**0.67**	**0.89**
PST BR (Testing)	**0.631**	**0.12**	**0.32**	**0.8**
ABSTINENCE				
4-Week Abstinence	1	0	1	1
12-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.353	0.26	-0.67	0.24
RBC	**0.874**	**0.11**	**0.484**	**0.96**
HCT	**0.88**	**0.12**	**0.4**	**0.97**
MCV	-0.209	0.26	-0.59	0.37
MCHC	0.259	0.33	-0.56	0.75
RDW	-0.28	0.27	-0.66	0.42
Ferritin Percentile	-0.007	0.31	-0.62	0.57
Hb	**0.966**	**0.13**	**0.597**	**0.98**
DA				
RT Low-Conflict (PST)	-0.088	0.33	-0.63	0.55
RT High-Conflict (PST)	-0.099	0.34	-0.75	0.53
Resting State BR	0.423	0.18	-0.13	0.7
PAL BR	**0.665**	**0.14**	**0.146**	**0.81**
IGT BR	**0.788**	**0.11**	**0.47**	**0.9**
SRT BR	**0.815**	**0.07**	**0.633**	**0.88**
PST BR (Training)	**0.826**	**0.08**	**0.605**	**0.9**
PST BR (Testing)	**0.631**	**0.13**	**0.344**	**0.83**
ABSTINENCE				
12-Week Abstinence	1	0	1	1

Note: Bold cells indicate significant factor loadings.

Table 16: Factor Loadings in GSCA Models (Learning)

4-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.359	0.25	-0.64	0.3
RBC	**0.876**	**0.1**	**0.64**	**0.98**
HCT	**0.852**	**0.09**	**0.64**	**0.97**
MCV	-0.244	0.22	-0.62	0.17
MCHC	0.286	0.32	-0.56	0.65
RDW	-0.291	0.29	-0.7	0.43
Ferritin Percentile	0.015	0.29	-0.6	0.46
Hb	**0.963**	**0.08**	**0.61**	**0.98**
Learning				
RT for A trials (Incorrect; PST)	**0.892**	**0.03**	**0.82**	**0.93**
RT for A trials (Correct; PST)	**0.884**	**0.04**	**0.79**	**0.93**
RT for B trials (Incorrect; PST)	**0.924**	**0.03**	**0.86**	**0.96**
Percent Accuracy (A trials; PST)	**0.487**	**0.11**	**0.26**	**0.67**
Percent Accuracy (B trials; PST)	-0.145	0.19	-0.49	0.25
Met Criteria (yes/no; PST)	0.184	0.18	-0.23	0.48
ABSTINENCE				
4-Week Abstinence	1	0	1	1
12-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.337	0.23	-0.63	0.26
RBC	**0.889**	**0.09**	**0.63**	**0.98**
HCT	**0.854**	**0.09**	**0.6**	**0.96**
MCV	-0.259	0.22	-0.61	0.24
MCHC	0.276	0.29	-0.42	0.64
RDW	-0.276	0.26	-0.65	0.32
Ferritin Percentile	0.002	0.29	-0.57	0.55
Hb	**0.959**	**0.07**	**0.78**	**0.98**
Learning				
RT for A trials (Incorrect; PST)	**0.894**	**0.03**	**0.83**	**0.93**
RT for A trials (Correct; PST)	**0.888**	**0.04**	**0.79**	**0.94**
RT for B trials (Incorrect; PST)	**0.923**	**0.02**	**0.87**	**0.96**
Percent Accuracy (A trials; PST)	**0.476**	**0.15**	**0.07**	**0.69**
Percent Accuracy (B trials; PST)	-0.147	0.17	-0.46	0.25
Met Criteria (yes/no; PST)	0.185	0.22	-0.3	0.51
ABSTINENCE				
12-Week Abstinence	1	0	1	1

Note: Bold cells indicate significant factor loadings.

Table 17: Factor Loadings in GSCA Models (Neural Activity)

4-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.42	0.27	-0.68	0.39
RBC	**0.826**	**0.13**	**0.45**	0.96
HCT	**0.818**	**0.14**	**0.34**	0.96
MCV	-0.246	0.27	-0.65	0.4
MCHC	0.363	0.34	-0.61	0.79
RDW	-0.346	0.29	-0.69	0.43
Ferritin Percentile	0.039	0.29	-0.49	0.53
Hb	**0.961**	**0.1**	**0.54**	0.98
Neural Activity				
Incorrect P200 (IGT; fc)	**-0.819**	**0.08**	**-0.92**	-0.6
Correct FRN (IGT)	**0.786**	**0.06**	**0.64**	0.88
Incorrect FRN (IGT)	**0.762**	**0.09**	**0.55**	0.9
Correct P200 (IGT; fc)	**-0.655**	**0.18**	**-0.83**	-0.1
Correct FRN (PST)	**0.579**	**0.19**	**0.12**	0.86
Incorrect FRN (PST)	0.159	0.3	-0.37	0.67
Incorrect P300 (PST; fc)	0.163	0.17	-0.21	0.44
ABSTINENCE				
4-Week Abstinence	1	0	1	1
12-Week Abstinence	**Estimate**	**SE**	**95%CI**	
IRON				
sTfR	-0.403	0.25	-0.78	0.37
RBC	**0.839**	**0.16**	**0.23**	0.96
HCT	**0.82**	**0.19**	**0.27**	0.96
MCV	-0.261	0.22	-0.64	0.16
MCHC	0.358	0.36	-0.55	0.79
RDW	-0.335	0.29	-0.73	0.36
Ferritin Percentile	0.03	0.32	-0.59	0.56
Hb	**0.96**	**0.11**	**0.59**	0.98
Neural Activity				
Incorrect P200 (IGT; fc)	**-0.817**	**0.08**	**-0.9**	-0.6
Correct FRN (IGT)	**0.788**	**0.07**	**0.6**	0.91
Incorrect FRN (IGT)	**0.768**	**0.09**	**0.55**	0.91
Correct P200 (IGT; fc)	**-0.649**	**0.18**	**-0.88**	-0.2
Correct FRN (PST)	0.577	0.2	-0.03	0.82
Incorrect FRN (PST)	0.165	0.31	-0.49	0.6
Incorrect P300 (PST; fc)	0.154	0.18	-0.27	0.44
ABSTINENCE				
12-Week Abstinence	1	0	1	1

Table 18: Path Coefficients from GSCA Models

Model	Estimate	SE	95%CI	
4-week abstinence outcomes				
Dopamine				
4-Week Abstinence → Iron	0.077	0.151	-0.171	0.363
Iron → DA	-0.062	0.201	-0.433	0.377
4-Week Abstinence → DA	-0.138	0.137	-0.404	0.100
Iron → 4-Week Abstinence	0.067	0.173	-0.283	0.357
DA → 4-Week Abstinence	-0.138	0.14	-0.421	0.094
Learning				
4-Week Abstinence → Iron	0.06	0.166	-0.249	0.435
Iron → Learning	-0.331	0.147	**-0.583**	**-0.017**
Iron → 4-Week Abstinence	0.124	0.175	-0.191	0.486
Learning → 4-Week Abstinence	0.194	0.152	-0.115	0.497
Neural Activity				
4-Week Abstinence → Iron	0.079	0.154	-0.178	0.455
Iron → Neural Activity	-0.283	0.292	-0.648	0.494
Iron → 4-Week Abstinence	0.074	0.18	-0.235	0.505
Neural Activity → 4-Week Abstinence	-0.016	0.192	-0.402	0.319
12-week abstinence outcomes				
Dopamine				
12-Week Abstinence → Iron	-0.009	0.172	-0.349	0.341
Iron → DA	-0.072	0.196	-0.407	0.346
12-Week Abstinence → DA	0.255	0.178	-0.155	0.568
Iron → 12-Week Abstinence	0.01	0.184	-0.405	0.335
DA→ 12-Week Abstinence	0.257	0.177	-0.156	0.559
Learning				
12-Week Abstinence → Iron	-0.018	0.171	-0.346	0.298
Iron → Learning	-0.338	0.143	**-0.603**	**-0.050**
Iron → 12-Week Abstinence	-0.003	0.187	-0.397	0.345
Learning → 12-Week Abstinence	0.044	0.15	-0.252	0.316
Neural Activity				
12-Week Abstinence → Iron	0.005	0.201	-0.329	0.447
Iron → Neural Activity	-0.283	0.282	-0.721	0.420
Iron → 12-Week Abstinence	-0.021	0.242	-0.387	0.504
Neural Activity → 12-Week Abstinence	-0.094	0.189	-0.415	0.310

Figure 1: Direct & Indirect Pathways of the Basal Ganglia; adapted from Blumenfeld (2010 [138]).

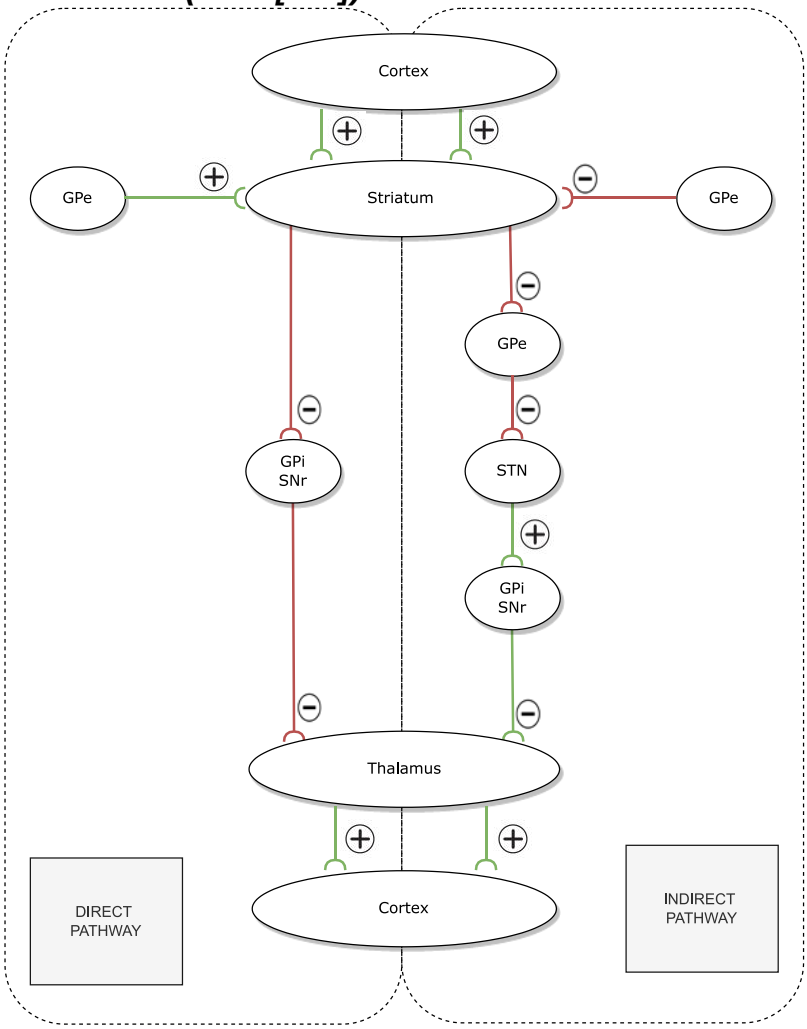


Figure 2: Proposed mediation models between iron and abstinence by way of DA.

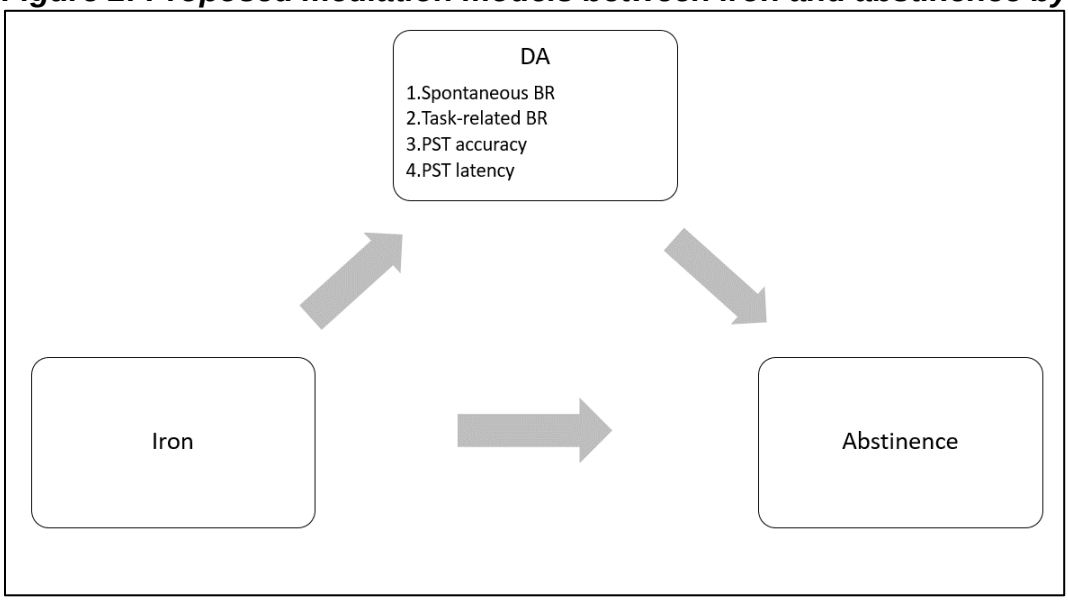


Figure 3: Proposed mediation models between iron and abstinence by learning.

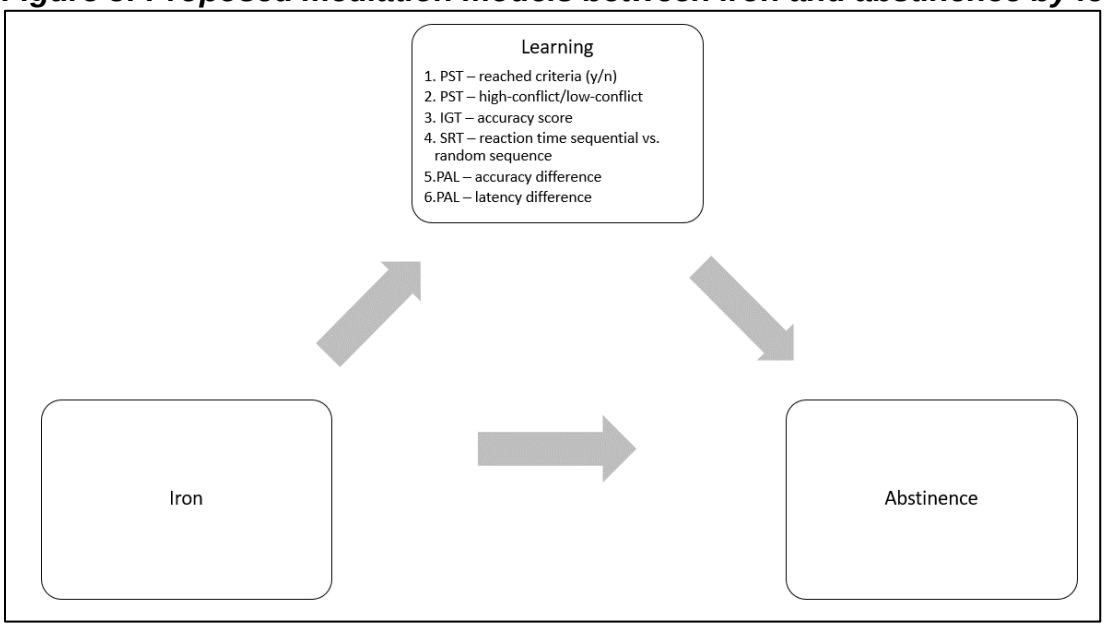


Figure 4: Proposed mediation models between iron and abstinence via brain dynamics.

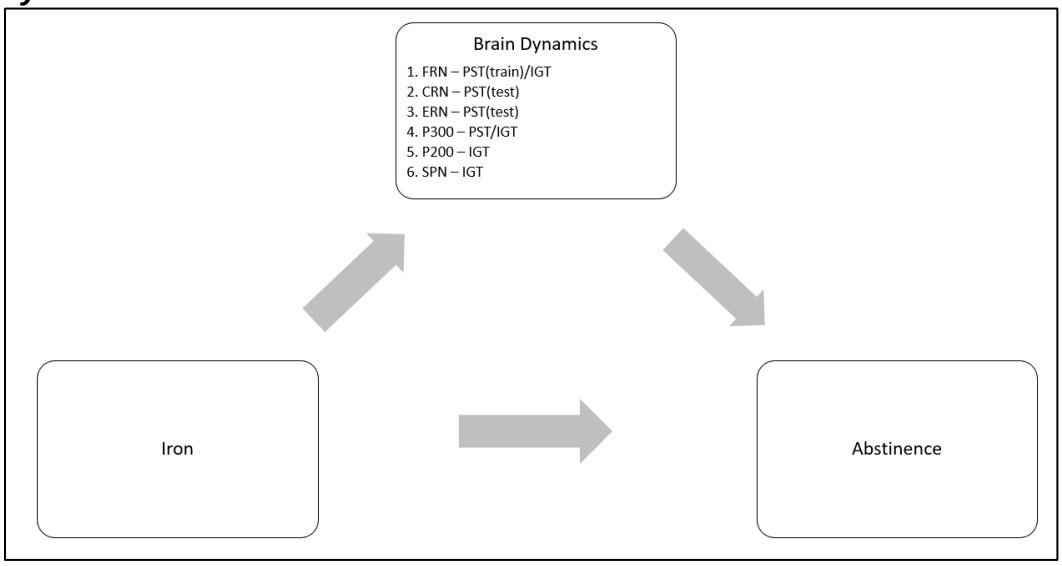
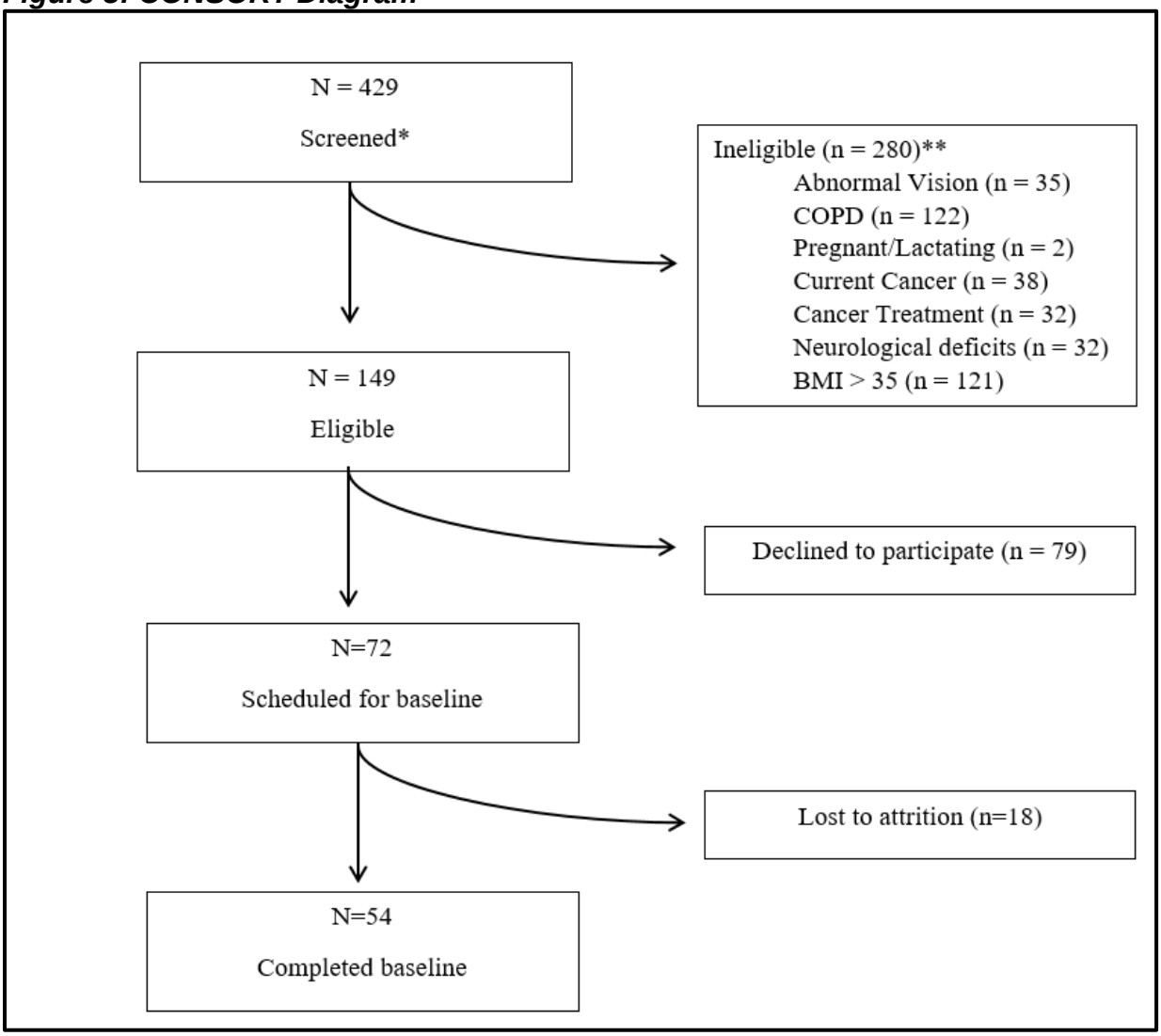


Figure 5: CONSORT Diagram



* 429 instances of screening occurred of 416 individuals; 11 individuals were screened twice; 1 individual was screened 3 times.

** Many participants who were ineligible endorsed more than one reason for ineligibility (n = 88).

Figure 6: Frequency of Abstinence at Follow-Up

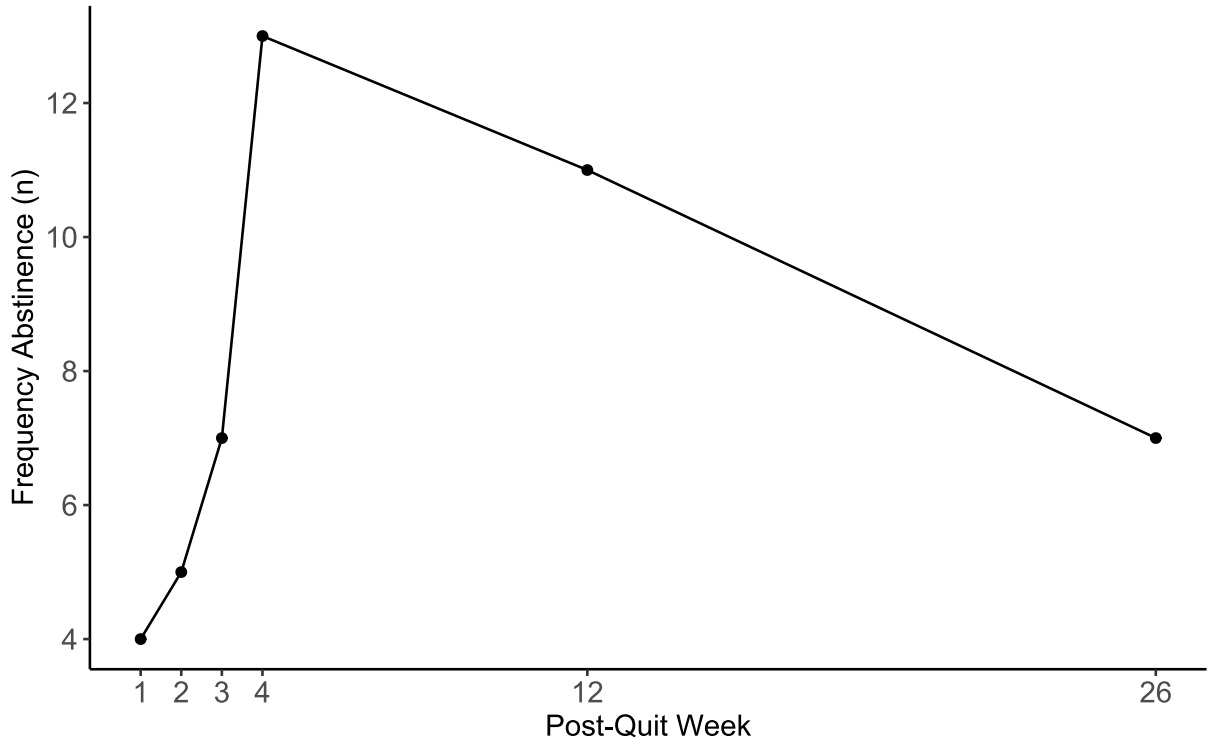


Figure 7: Mean Expired CO Comparisons from Baseline to 26-Weeks Post-Quit

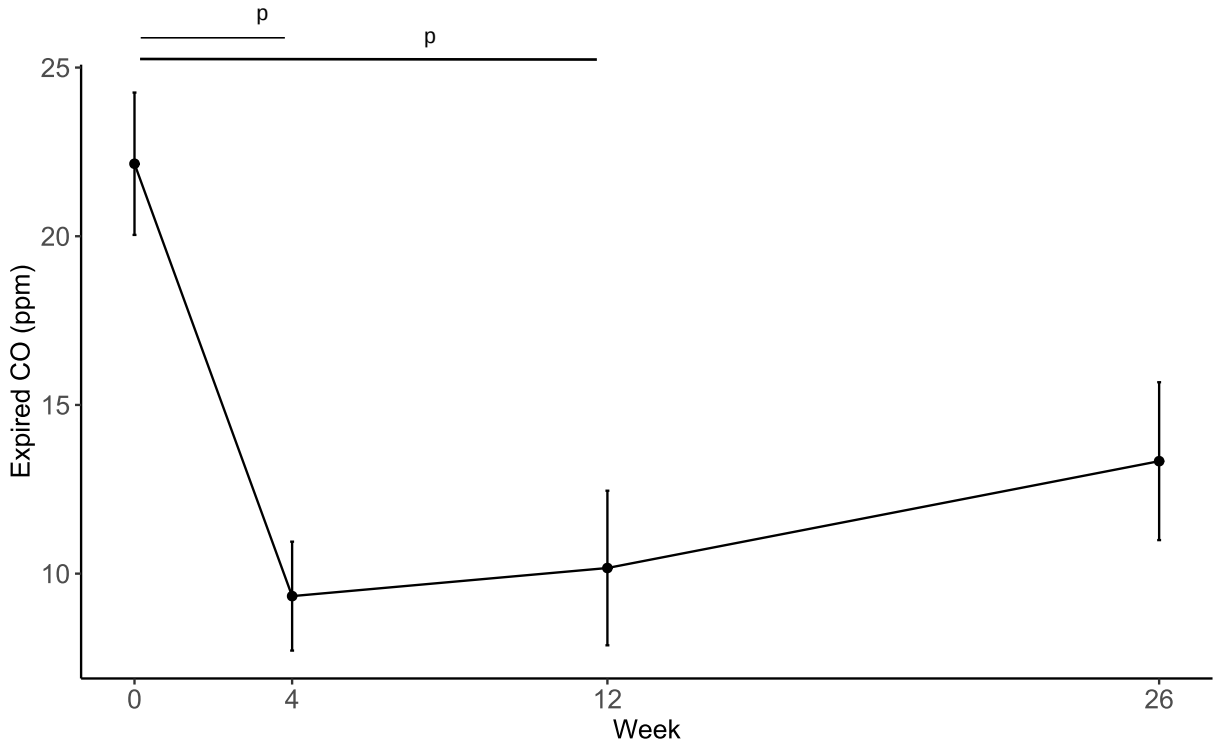


Figure 8: Average Reaction Time for Trial Type by Accuracy Interaction during PST

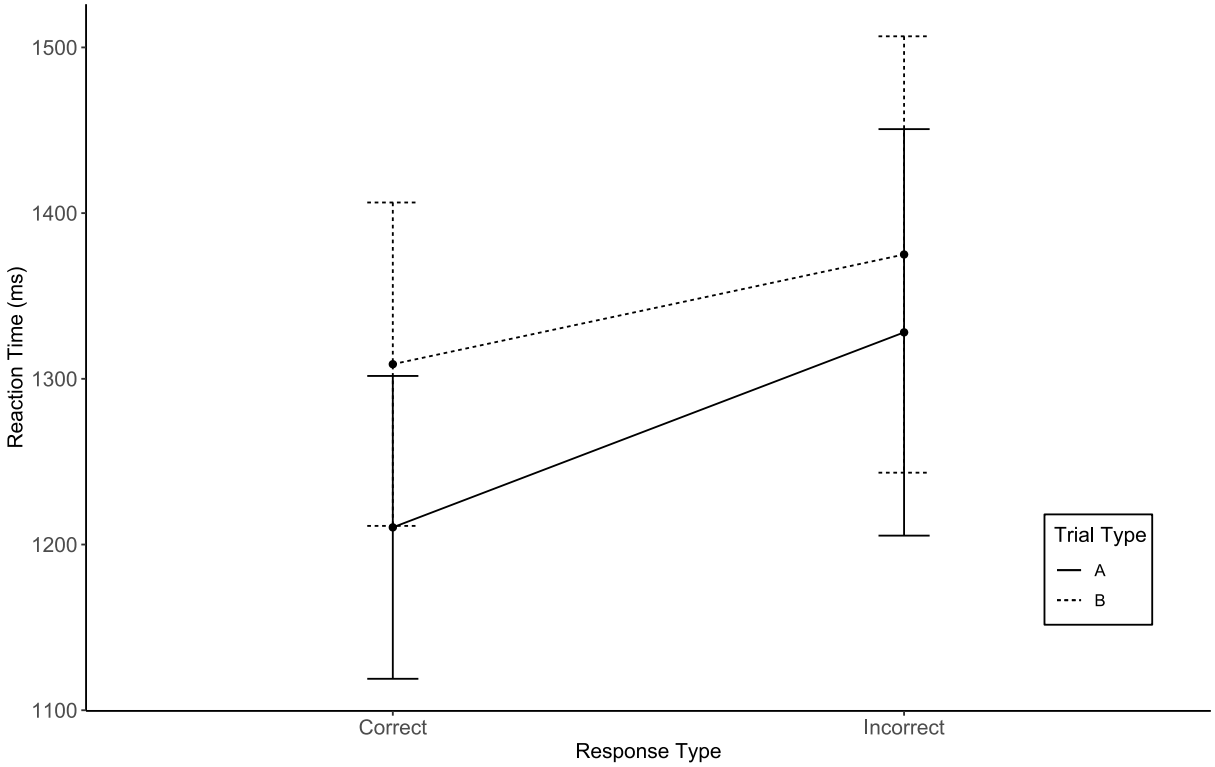


Figure 9: Frequency of Playing from Each Deck Type over 10 Blocks during IGT

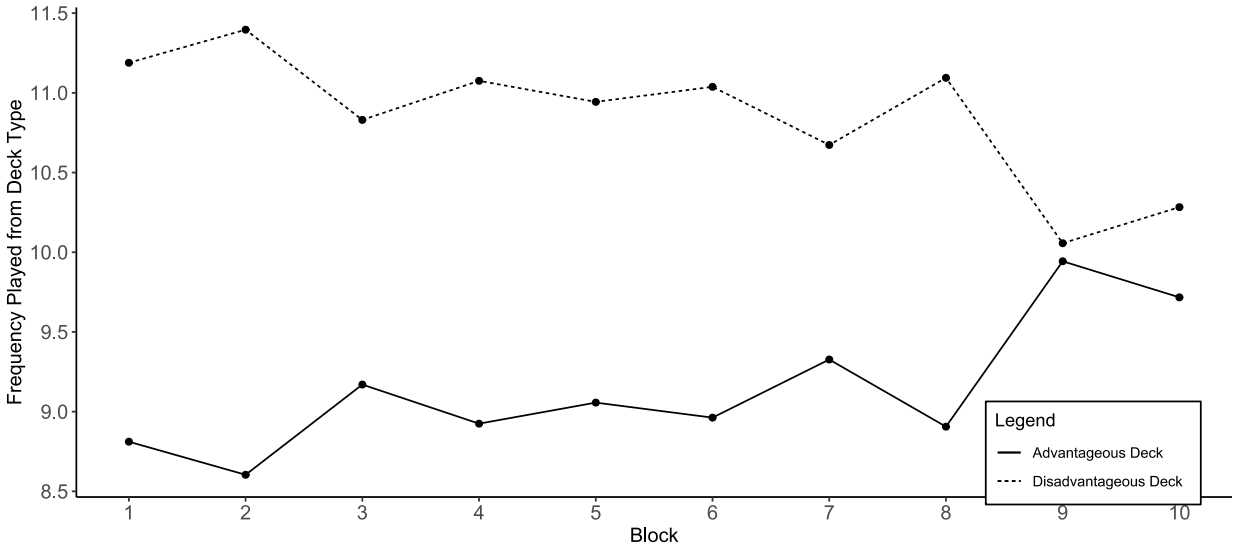


Figure 10: RT during 30 Blocks of SRT

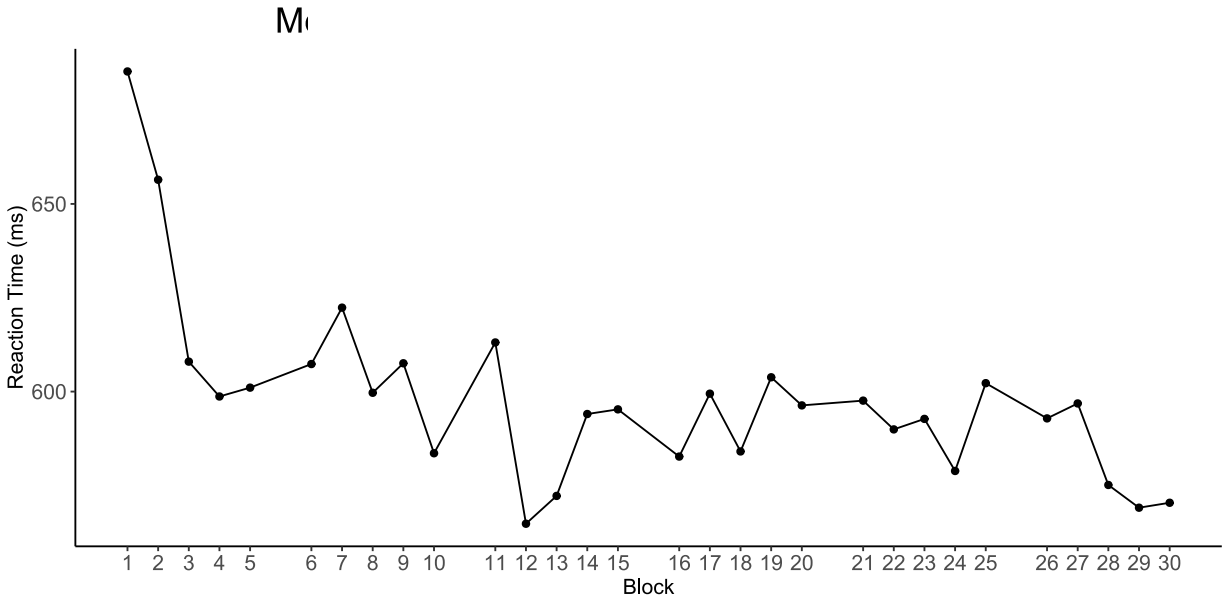
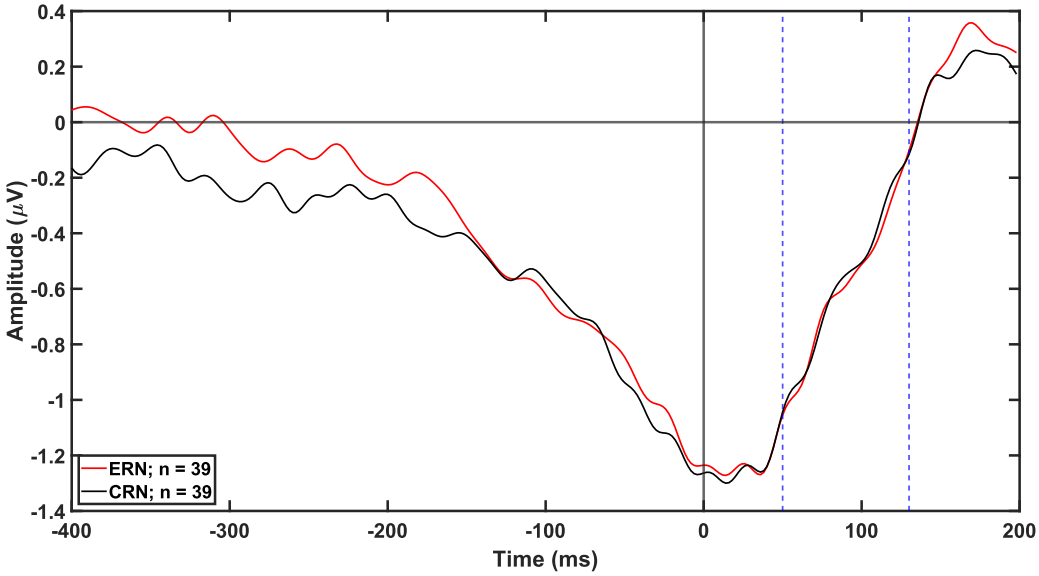
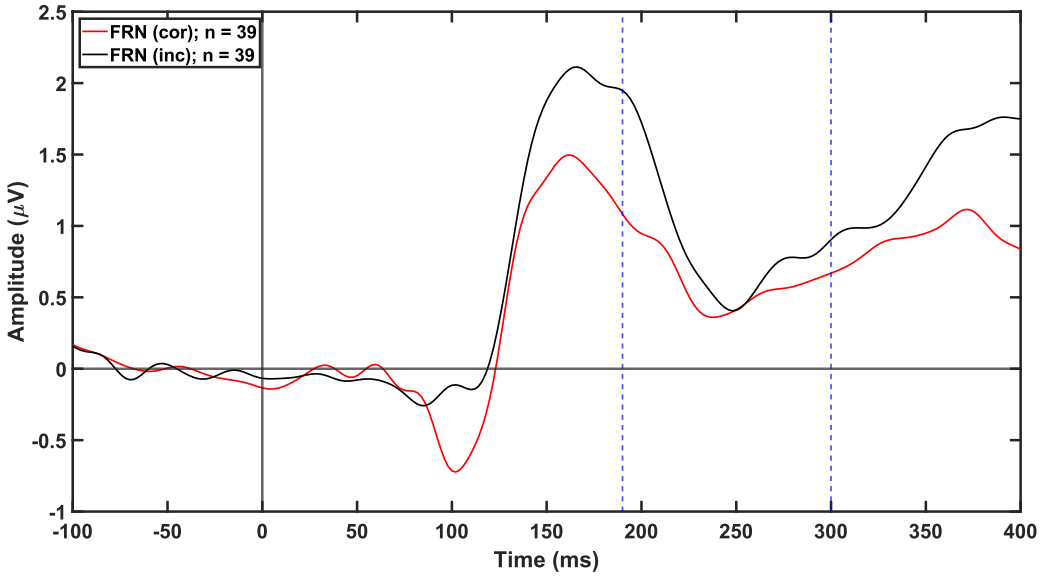


Figure 11: Response-locked ERN and CRN during PST test phase



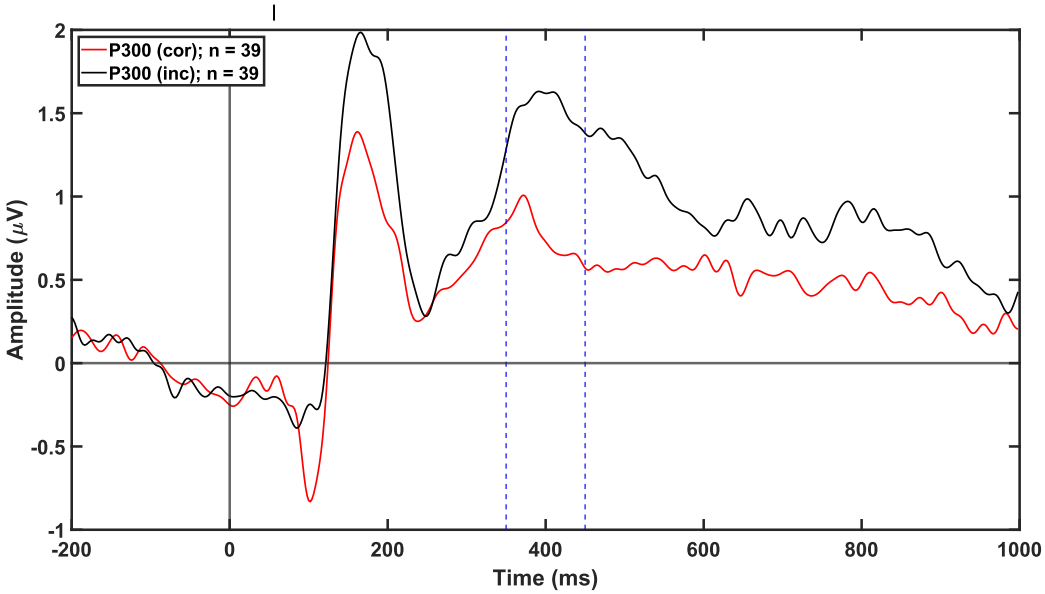
Reference lines illustrate time periods during which component was isolated (ERN/CRN: 50-130ms).

Figure 12: Feedback-locked FRN by feedback type (correct vs. incorrect) during PST training phase



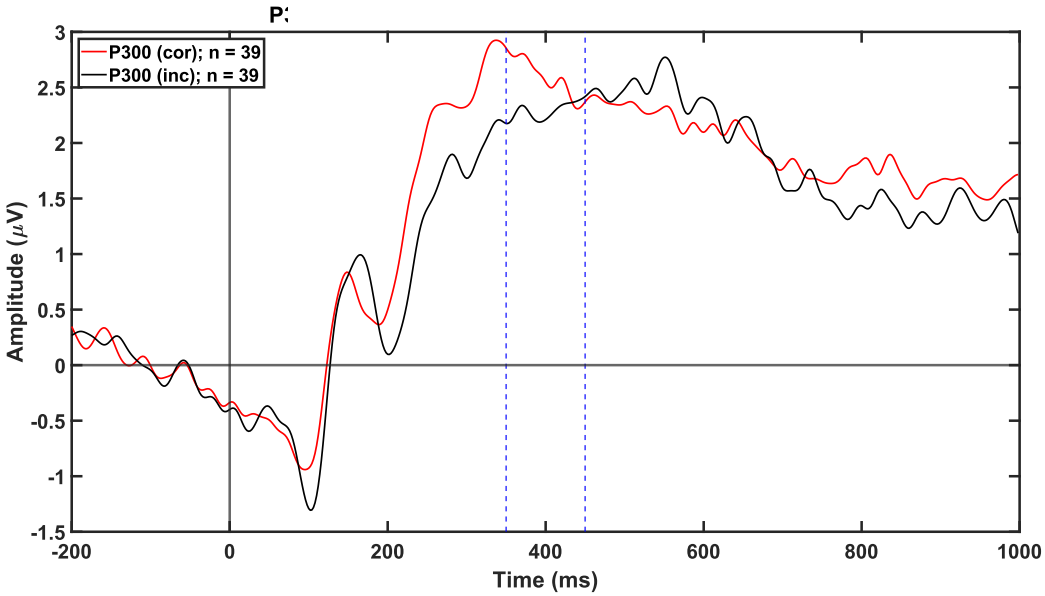
Reference lines illustrate time periods during which component was isolated (FRN: 190-300ms).

Figure 13: Feedback-locked P300 by feedback type (correct vs. incorrect) during PST training phase (fronto-central region)



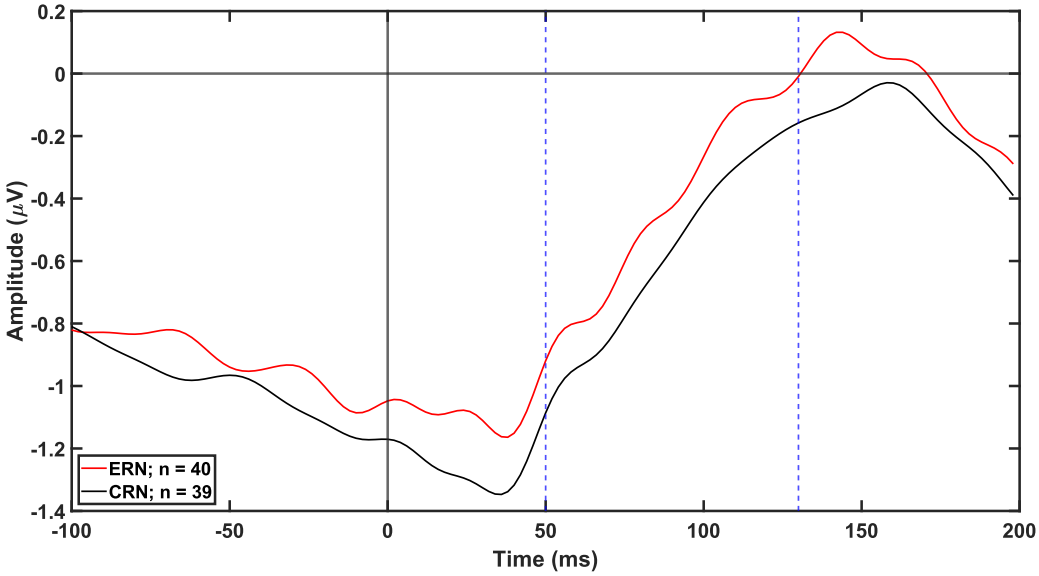
Reference lines illustrate time periods during which component was isolated (P300: 350-450).

Figure 14: Feedback-locked P300 by feedback type (correct vs. incorrect) during PST training phase (centro-parietal region)



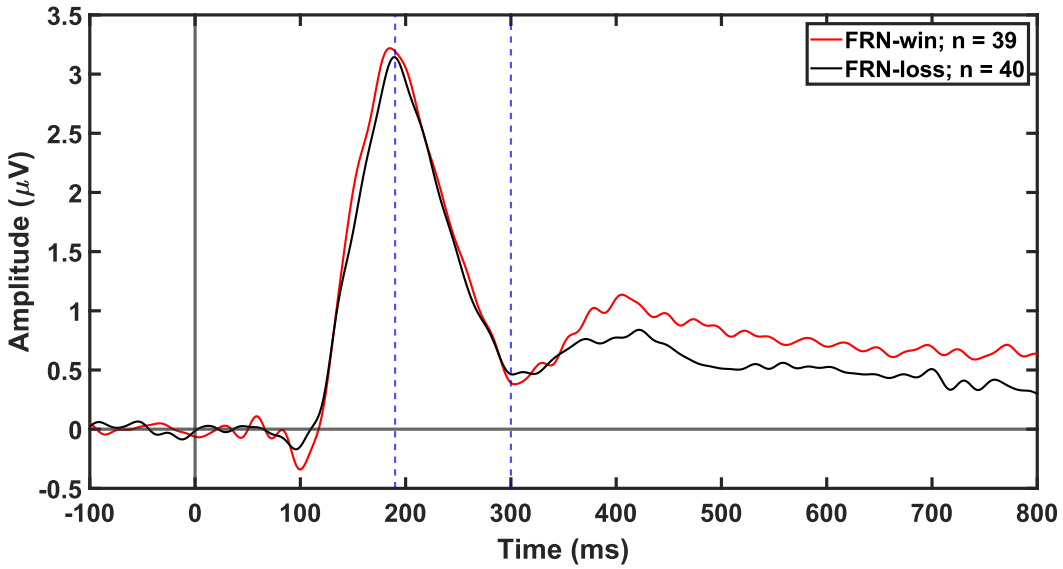
Reference lines illustrate time periods during which component was isolated (P300: 350-450).

Figure 15: Response-locked ERN and CRN during IGT



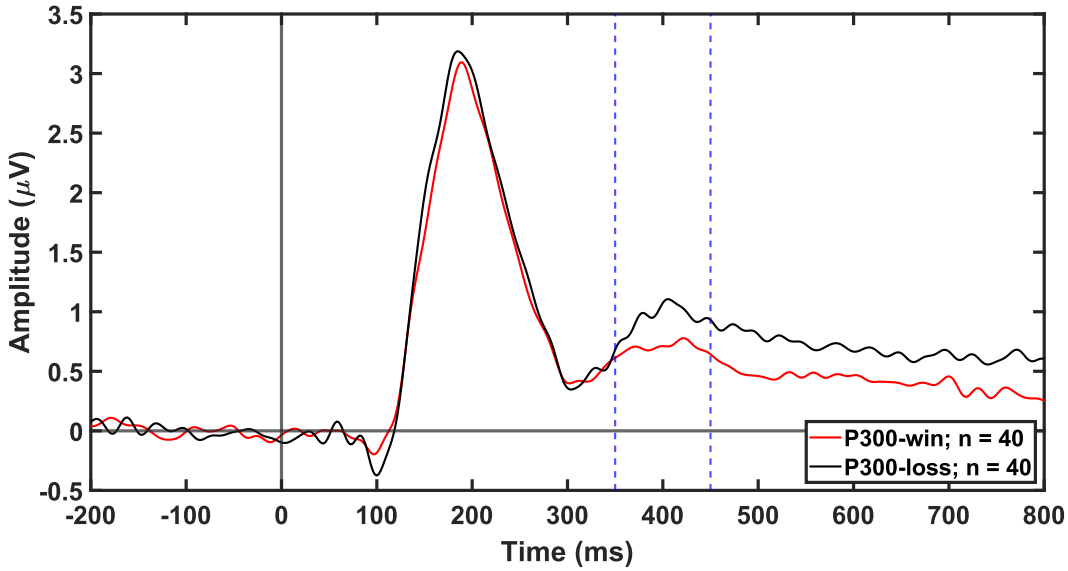
Reference lines illustrate time periods during which component was isolated (ERN/CRN: 50-130ms).

Figure 16: Feedback-locked FRN by feedback type (win vs. loss) during IGT



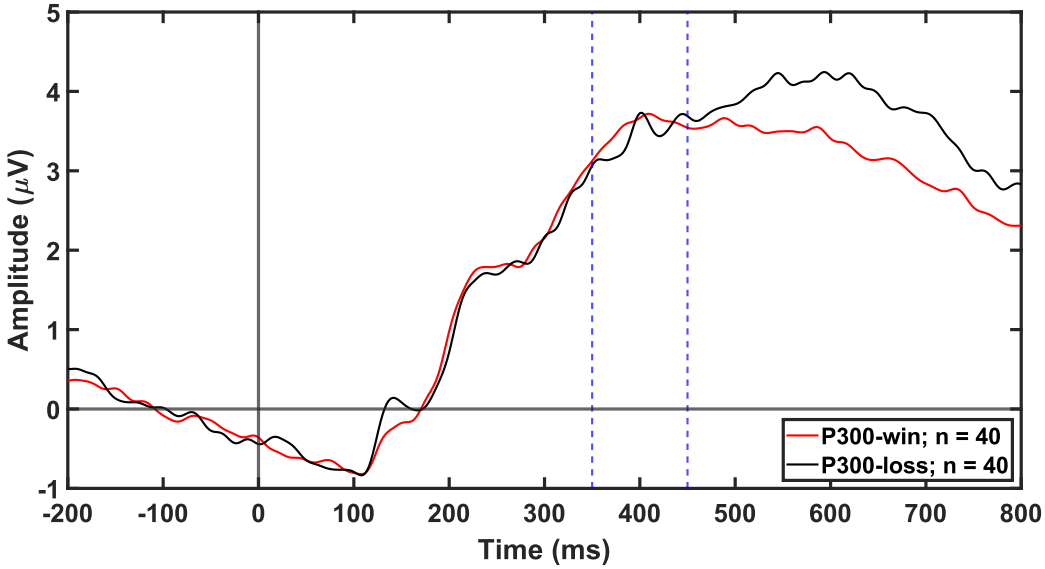
Reference lines illustrate time periods during which component was isolated (FRN: 190-300ms).

Figure 17: Feedback-locked P300 by feedback type (win vs. loss) during IGT (fronto-central region)



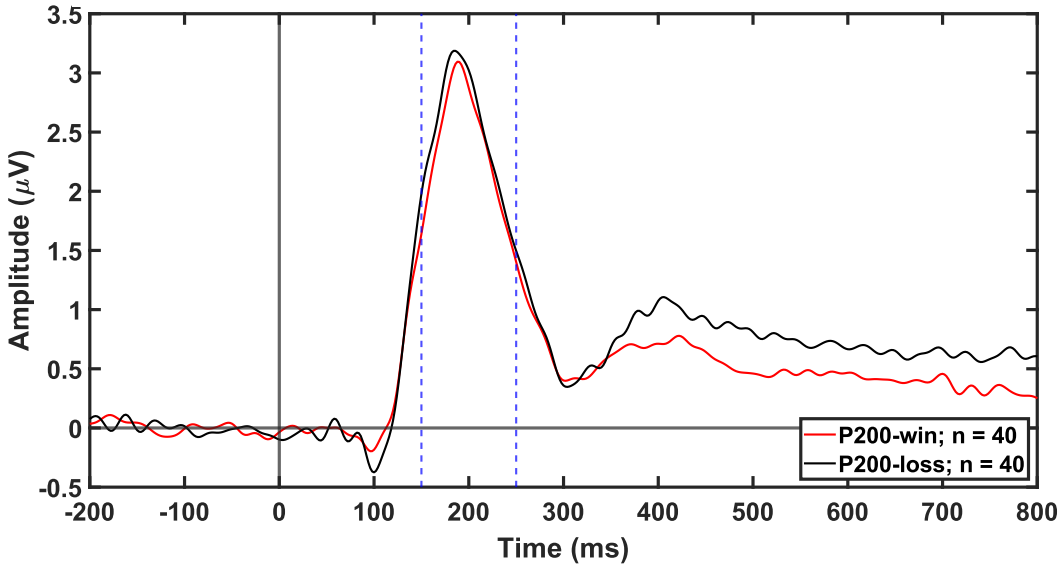
Reference lines illustrate time periods during which component was isolated (P300: 350-450).

Figure 18: Feedback-locked P300 by feedback type (win vs. loss) during IGT (centro-parietal region)



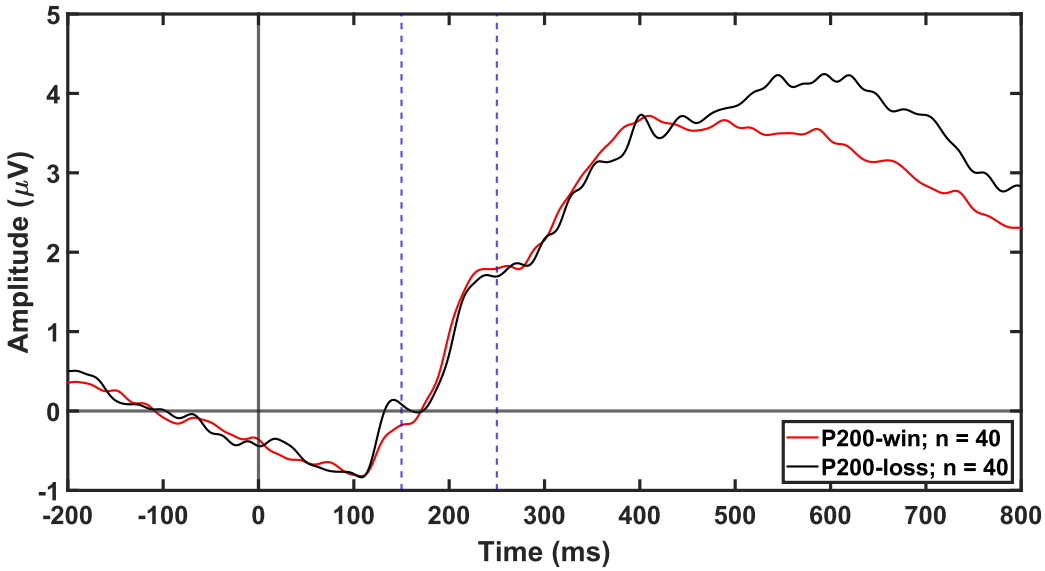
Reference lines illustrate time periods during which component was isolated (P300: 350-450).

Figure 19: Feedback-locked P200 by feedback type (win vs. loss) during IGT (fronto-central region)



Reference lines illustrate time periods during which component was isolated (P200: 150-250).

Figure 20: Feedback-locked P200 by feedback type (win vs. loss) during IGT (centro-parietal region)



Reference lines illustrate time periods during which component was isolated (P200: 150-250).

Figure 21: Path Diagram in which Dopamine mediates the Relation between Iron and Abstinence at 4 weeks

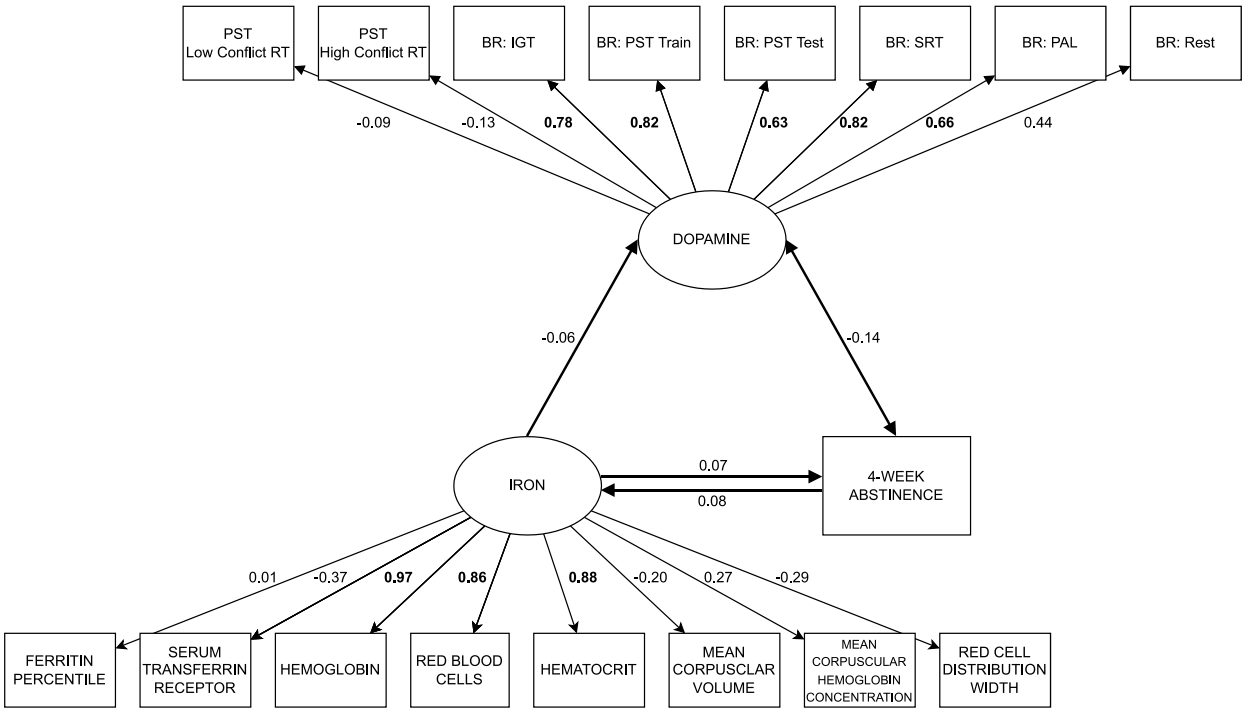


Figure 22: Path Diagram in which Learning mediates the Relation between Iron and Abstinence at 4 weeks

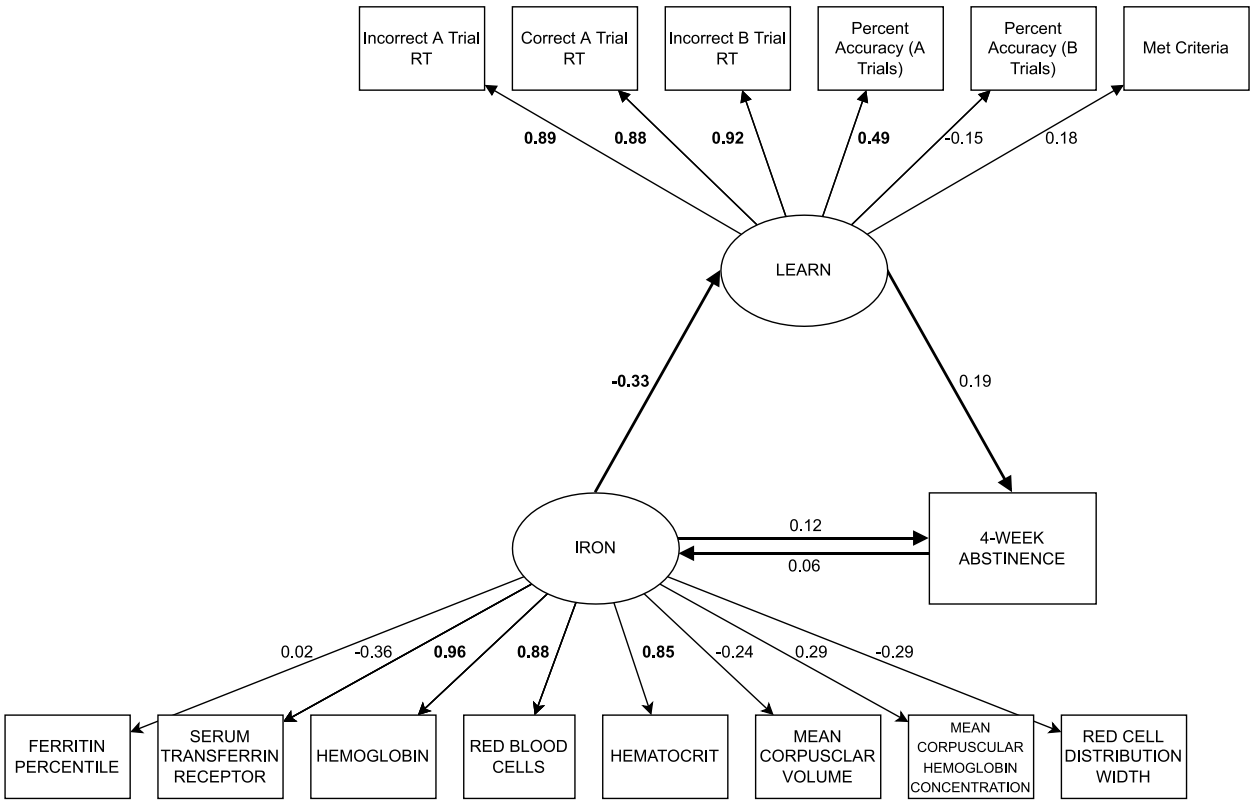


Figure 23: Path Diagram in which Neural Activation mediates the Relation between Iron and Abstinence at 4 weeks

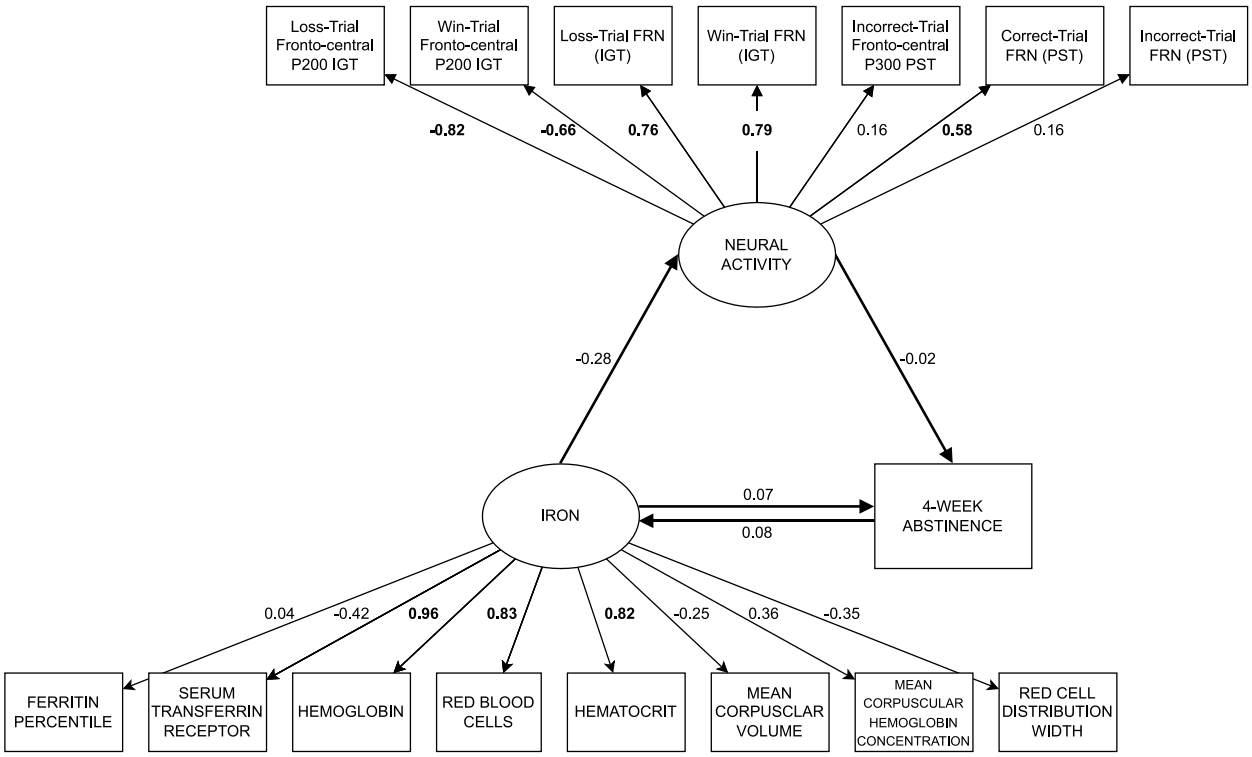


Figure 24: Path Diagram in which Dopamine mediates the Relation between Iron and Abstinence at 12 weeks

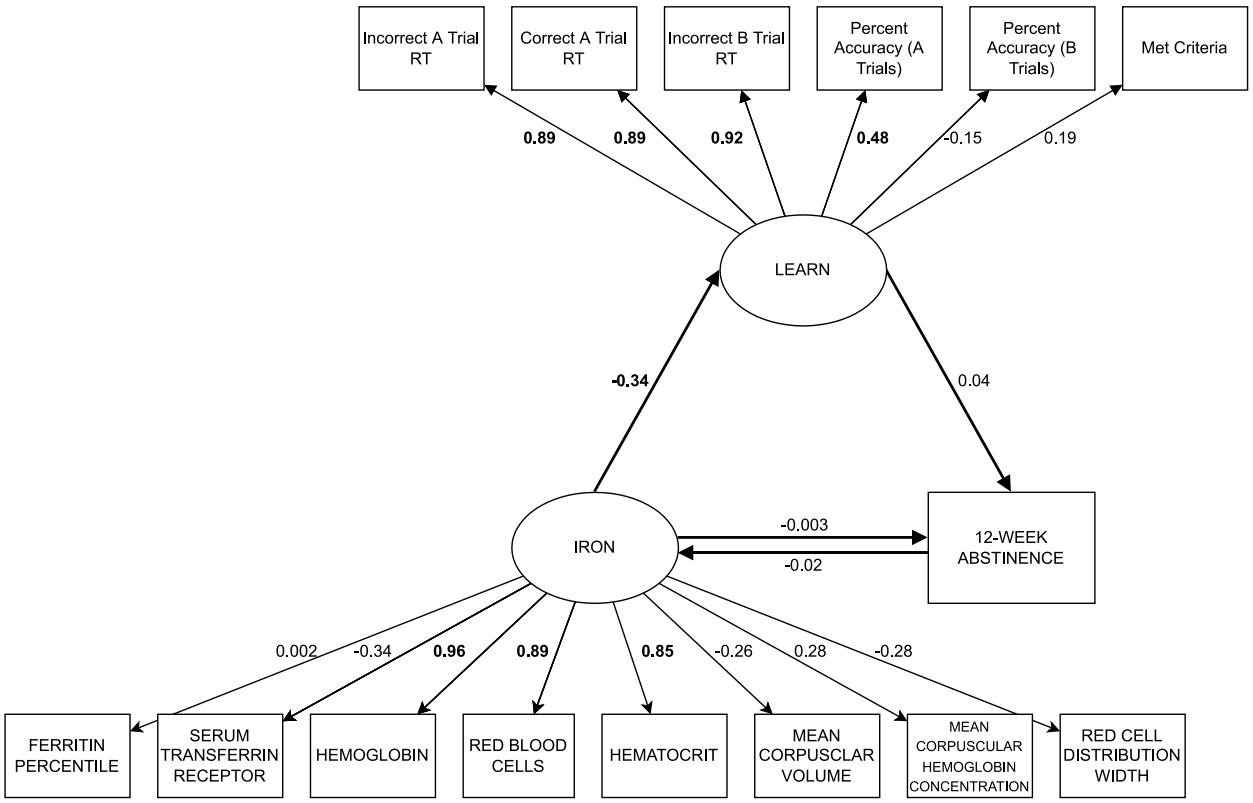


Figure 25: Path Diagram in which Learning mediates the Relation between Iron and Abstinence at 12 weeks

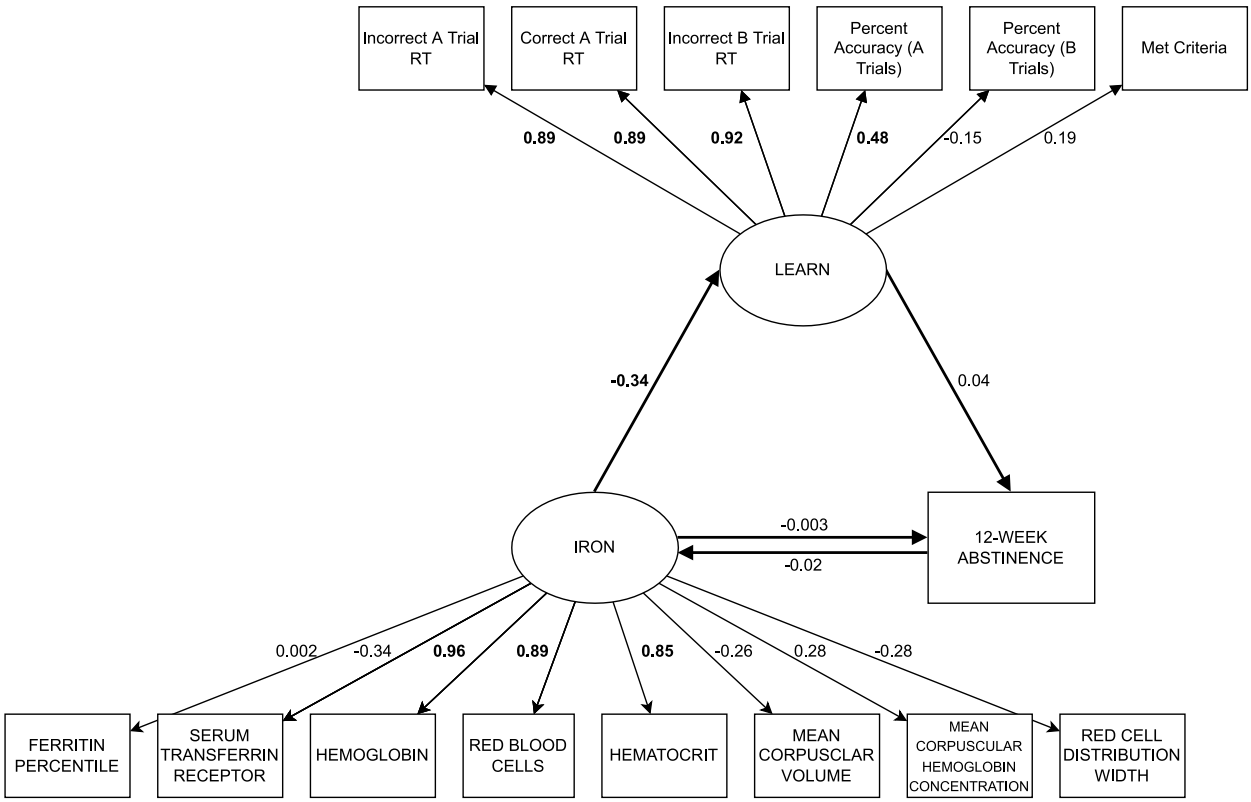
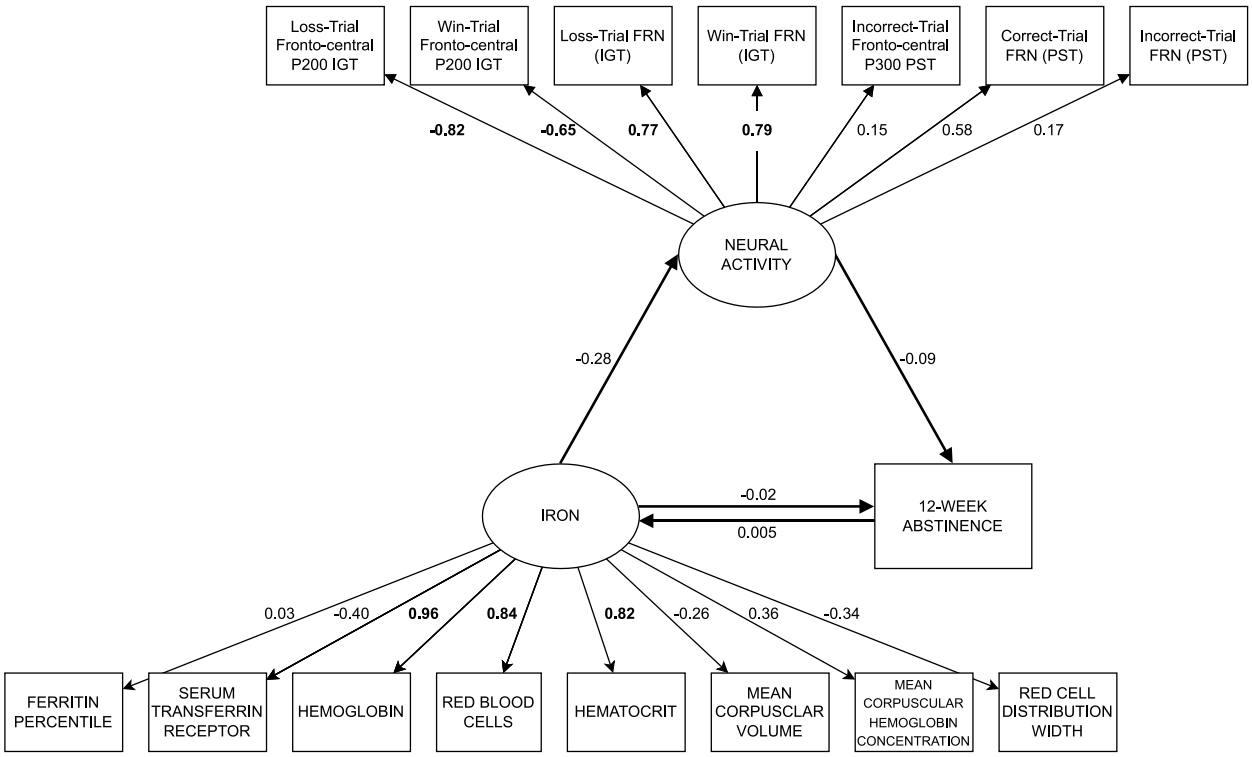


Figure 26: Path Diagram in which Neural Activation mediates the Relation between Iron and Abstinence at 12 weeks



APPENDIX A: PHQ Depression Questionnaire

Question	Asks
Over *the last 2 weeks, how often have you been bothered by any of the following problems?*	
1	Little interest or pleasure in doing things
2	Feeling down, depressed, or hopeless
3	Trouble falling or staying asleep, or sleeping too much
4	Feeling tired or having little energy
5	Poor appetite or overeating
6	Feeling bad about yourself-or that you are a failure or have let yourself or your family down
7	Trouble concentrating on things, such as reading the newspaper or watching television
8	Moving or speaking so slowly that other people could have noticed? Or the opposite-being so fidgety or restless that you have been moving around a lot more than usual

Response options for all questions: 0 = Not at all | 1 = Several days | 2 = More than half the days | 3 = Nearly every day