Abstract

Electrophysiological Indices of Disease Status in Multiple Sclerosis

Background: Multiple Sclerosis (MS) is a disease of the central nervous system, characterized by damage or destruction of axons throughout the brain and spinal cord. Visual impairment is one of the most common symptoms of MS and often the first manifestation of the disease. The current electrophysiological study utilized a battery of short-duration transient visual evoked potential (tVEP) tests, along with clinical measures of depression and fatigue, to assess visual system function and clinical correlates in individuals with relapse-remitting multiple sclerosis (n = 18), and unaffected controls (n = 16).

Methods: tVEPs to a contrast-reversing checkerboard were administered monocularly using small and large checks to examine neural mechanisms and assess differences between groups. Participants completed the Beck Depression Inventory-II and the Fatigue Severity Scale in order to evaluate relationships between clinical measures and tVEPs. In the time-domain, measures of amplitude and latency (peak time) in the tVEP waveform were examined. In the frequency-domain, following a discrete Fourier transform of the time-series data, the following measures of distinct harmonic frequency components and bands of frequency components were examined: amplitude, power, magnitude-squared coherence (MSC), phase, and two novel delay estimates based on phase data were examined. Univariate and multivariate statistics were used to examine patterns of responses and non-parametric tests were used to assess group differences. Linear mixed effects modelling was used to identify significant differences in the outcome variables based on the fixed-effect variables of

group, eye (right vs. left fellow eyes), and stimulus condition (large vs. small checks). Classification accuracy was assessed for all measures using receiver operating characteristic (ROC) curve analysis and the non-parametric measures of area under the ROC curve (AUC) and A'.

Results: Overall, small checks yielded stronger responses and more significant results than did large checks. There was great variability in amplitude within and between groups. There were significant differences in P60-N75 amplitude and mean MSC for Frequency Band 3 (30-40 Hz) between groups (right eye: $p = .02$, left eye: $p = .01$). There were numerous group differences in the time and frequency-domain delay measures. The frequency-domain measures produced the strongest effects and had the greatest classification accuracy for group membership (H24 delay: $p < .01$, Band 2 delay: $p = .01$).

Conclusions: The current study findings support the use of the short-duration tVEP-CR test (Zemon & Gordon, 2018) in the assessment of visual dysfunction in patients with MS. Results support increased latency as a characteristic finding in MS and add to existing literature by providing evidence for the use of novel frequency domain measures, which provide greater classification accuracy and are more sensitive to detecting differences between groups than are time-domain measures. These techniques require further study using larger samples as they have the potential to be used as biomarkers for this disease. Findings that early excitatory input into the cortex is impacted in MS indicate that neural damage (likely associated with demyelination) is occurring in early, precortical levels of visual processing. Notably, the P60 time point is often missed and not evaluated in clinics. Additionally, these findings potentially provide evidence to support the utility of VEPs in the diagnosis, symptom progression, and treatment monitoring of individuals with MS.

Electrophysiological Indices of Disease Status in Multiple Sclerosis.

by

Kasey B. Siegel, M.A.

Submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the Ferkauf Graduate School of Psychology

September 2019

The committee for this doctoral dissertation consists of:

Vance Zemon, Ph.D., Chairperson, Ferkauf Graduate School of Psychology, Yeshiva University

Frederick Foley, Ph.D., Ferkauf Graduate School of Psychology, Yeshiva University

James Gordon, Ph.D., Hunter College of the City University of New York

Acknowledgements

This project would never have been developed or completed had it not been for the mentorship and support of Vance Zemon, Ph.D. Vance, I cannot thank you enough for the time, commitment, and enthusiasm you have put into this study. You have helped make me a scientist in ways I never thought were possible. I will be forever grateful for the role you have played throughout my graduate career and most especially, throughout this process. I would also like to thank all the members of my committee for taking part in this last and very important phase of my graduate career.

Dedication

To my husband, Jared, my mom, Robin, Bampy, and the entire Bovarnick and Siegel families, with immense love and gratitude.

Table of Contents

List of Tables x

List of Figures xi

Chapter I: Introduction 2

Background and Significance 3

Rationale for the Current Study 46

Innovation 47

Aims and Hypotheses 49

Chapter II: Methods 53

Participants 53

Recruitment and Eligibility 54

Measures 54

Procedures 57

Data Analysis 58

Chapter III: Results 64

Chapter IV: Discussion 89

Interpretation 89

Clinical Implications 100

Limitations and Future Directions 100

References 103

List of Tables

Table 1. Demographic and clinical characteristics for study sample by cohort 121

Table 2. Group means, standard deviations, and Mann-Whitney tests for the small check condition, right eye..... 122

Table 3. Group means, standard deviations, and Mann-Whitney tests for the small check condition, left eye..... 123

Table 4. Group means, standard deviations, and Mann-Whitney tests for the large check condition, right eye..... 124

Table 5. Group means, standard deviations, and Mann-Whitney tests for the large check condition, left eye 125

Table 6. Zeroth-order correlations among amplitude and MSC/power frequency-band measures by cohort for controls and patients for the small check condition, right eye..... 126

Table 7. Zeroth-order correlations among amplitude and MSC/power frequency-band measures by cohort for controls and patients for the small check condition, left eye..... 127

Table 8. Zeroth-order correlations among amplitude and MSC/power frequency-band measures by cohort for controls and patients for the large check condition, right eye..... 128

Table 9. Zeroth-order correlations among amplitude and MSC/power frequency-band measures by cohort for controls and patients for the large check condition, left eye..... 129

Table 10. Zeroth-order correlations among time and frequency-domain latency measures by cohort for controls and patients for the small check condition, right eye..... 130

Table 11. Zeroth-order correlations among time and frequency-domain latency measures by cohort for controls and patients for the small check condition, left eye.....131

Table 12. Zeroth-order correlations among time and frequency-domain latency measures by cohort for controls and patients for the large check condition, right eye.....132

Table 13. Zeroth-order correlations among time and frequency-domain latency measures by cohort for controls and patients for the large check condition, left eye.....133

Table 14. Receiver operating characteristic curve analysis for the small check condition...134

Table 15. Receiver operating characteristic curve analysis for the large check condition....135

List of Figures

Figure 1. Contrast-reversing checkerboard stimulus pattern used to elicit tVEPs.....136

Figure 2A. Representative data from a control participant for the small check condition...137

Figure 2B. Representative data from a control participant for the large check condition... 138

Figure 2C. Representative phase versus frequency plot with Band 2 delay estimate from a control participant for the large check condition.....139

Figure 3A. Representative data from an MS participant for the small check condition.....140

Figure 3B. Representative data from an MS participant for the large check condition.....141

Figure 3C. Representative phase versus frequency plot with Band 2 delay estimate from an MS participant for the large check condition.....142

Figure 4. Representative data from an MS participant with optic neuritis for the small check condition..... 143

Figure 5. Representative data from a control participant with medications affecting GABA for the small check condition.....144

Figure 6A. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the small check condition, right eye.....145

Figure 6B. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the small check condition, left eye.....145

Figure 6C. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the large check condition, right eye.....146

Figure 6D. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the large check condition, left eye.....146

Figure 7A. Power (microvolts^2/band) plotted as a function of frequency band for the small check condition, right eye..... 147

Figure 7B. Square root of power plotted as a function of frequency band for the small check condition, right eye.....147

Figure 8A. Mean MSC plotted as a function of frequency band for the small check condition, right eye.....148

Figure 8B. Mean MSC plotted as a function of frequency band for the large check condition, right eye.....148

Figure 8C. Mean MSC plotted as a function of frequency band for all conditions (small and large check) and eyes (OD and OS) 149

Figure 9. Scatterplot of P60-N75 amplitude by square root of power in Band 3 for the small check condition, left eye.....150

Figure 10. Scatterplot of N75-P100 amplitude by square root of power in Band 2 for the small check condition, right eye.....151

Figure 11A. Mean peak time estimates (ms) by time- and frequency-domain delay measures for the small check condition.....152

Figure 11B. Mean peak time estimates (ms) by time- and frequency-domain delay measures for the large check condition.....152

Figure 11C. Mean peak time estimates (ms) by time- and frequency-domain delay measures collapsed across conditions and eyes.....153

Figure 12A. Scatterplot of P100 Peak Time by Harmonic 24 Delay for the small check condition, right eye.....154

Figure 12B. Scatterplot of P100 Peak Time by Harmonic 24 Delay for the small check condition, left eye.....155

Figure 12C. Scatterplot of P100 Peak Time by Harmonic 24 Delay for the large check condition, right eye.....156

Figure 12D. Scatterplot of P100 Peak Time by Harmonic 24 Delay for the large check condition, left eye.....157

Figure 13. Receiver operating characteristic (ROC) curves of square root of power in Band 3 and MSC in Band 3 (bottom) for small checks..... 158

Figure 14. ROC curves of peak times for small and large checks.....159

Figure 15. ROC curves of Band 2 delay and H24 delay measures for small and large checks.....160

Figure 16A. Interocular scatterplot of MSC in Band 3 for the small check condition.....161

Figure 16B. Interocular scatterplot of square root of power in Band 3 for the small check condition.....162

Figure 17A. Scatterplot of H24 delay by Band 3 MSC for the small check condition, right eye.....163

Figure 17B. Scatterplot of H24 delay by Band 3 MSC for the small check condition, left eye.....164

Figure 18A. Classification accuracy of composite frequency domain variable for the small check condition..... 165

Figure 18B. Classification accuracy of P100 variable for the small check condition.....166

Chapter I: Introduction

Multiple sclerosis (MS) is a complex chronic disease of the central nervous system (CNS) that causes inflammation and damage to neural pathways in the brain and spinal cord. MS is an immune-mediated disease; it impacts the CNS through the body's immune system attacking its own tissues, or healthy cells. MS works by destroying the myelin sheath that surrounds and protects axons in the brain and spinal cord. This results in suppression or blocking of communication between neurons. The prevalence of MS is increasing rapidly (Wallin et al., 2019). Current estimates suggest that MS affects close to 1 million people in the United States or 362.6 cases for every 100,000 people (Wallin et al., 2019). It is widely accepted among researchers that multiple mechanisms contribute to the development of MS, however, specific causes are still unknown. Visual impairment is a key symptom in MS, and often the first symptom reported.

The visual evoked potential (VEP) is an electrophysiological response that has been used to assess the functional integrity of neural mechanisms and pathways in a variety of neurologic conditions, including MS. Collecting VEPs through electroencephalographic (EEG) recording is a non-invasive, objective way to examine the underlying neural processes within the visual system. Techniques have been developed in our laboratory to examine the integrity of neural pathways that contribute to the VEP. The bulk of existing VEP research on MS has been conducted using the conventional contrast-reversing checkerboard stimulus. Using this stimulus, conclusions can be drawn based on amplitudes and peak times in the waveform (Regan, 1989; Zemon et al., 1995). However, the conventional VEP is limited in that it elicits responses from a host of neural mechanisms and thus lacks the resolution

needed to identify specific information about select neural pathways and mechanisms of dysfunction (Zemon et al., 1995). Additionally, VEP abnormalities caused by lesions may exist anywhere along the retino-geniculo-cortical pathway, thereby further limiting information regarding neuroanatomic specificity (Beh, Frohman, & Frohman, 2013). Using novel VEP techniques, many of which were developed in our laboratory, we can examine more specific mechanisms and pathways than can conventional approaches.

The current study aimed to assess the effects of the disease state on visual processing in individuals with MS as compared to healthy controls. We expect to identify differential patterns of visual responses to various stimuli between the groups that reflect select deficits in distinct neural mechanisms. Two of the most common conditions experienced by individuals with MS are fatigue and depression (Giovannoni, 2006; Siegert & Abernethy, 2005). The high incidence of these conditions is understandable, given the multitude of physical, sensory, and cognitive impairments in MS, which can easily lead to fatigue and depression. Given the high rates of comorbidity between MS, fatigue, and depression, the current study utilized measures of fatigue and depression to assess their associations with the visual functions measured.

Background and Significance

Background on Multiple Sclerosis

The MS disease process works by damaging and destroying oligodendrocytes, the cells in the CNS that make myelin, and myelin, the insulating substance that surrounds and protects nerve fibers (axons) in the brain and spinal cord (Garg & Smith, 2015). The damaged myelin forms scar tissue (known as sclerosis). As a result of damage or destruction to the myelin sheath, nerve impulses that typically transmit information throughout the CNS

become interrupted or distorted. This causes the production of a wide range of symptoms, which will be discussed in detail in later sections.

Epidemiology. Multiple sclerosis currently affects 400,000 people in the United States and approximately 2.5 million people around the world. MS has been diagnosed at all ages but is most commonly diagnosed between the ages of 20 and 50, with an average age of onset at 30 (Milo & Miller, 2014). Women are at least two-three times more likely than men to develop MS and those with a family history of MS are also at an increased risk (Milo & Miller, 2014; National MS Society, 2018). MS is more prevalent in industrialized countries and countries further from the equator, although the latter finding has recently been called into question (Ha-Vinh et al., 2016; Milo & Miller, 2014). Interestingly, studies on migration patterns indicate that exposure to an environmental agent before puberty (e.g., before age 15) may create a predisposition for MS. This is based on the finding that a pre-pubescent migrant will inherit the MS risk-level of their new home region (Kurtzke, 2000).

Racial/Ethnic Differences in MS. MS is most common in Caucasians with northern European ancestry, although it does occur in most other ethnic groups including Hispanics/Latinos, Asians, and African Americans. Some studies suggest that the disease course in African Americans may be more active and rapidly progressive than in other ethnic groups, thereby leading to greater disability in this group (Johnson et al., 2010; Kister et al., 2010). Research suggests that an individual's race/ethnicity may differentially impact the MS disease process, symptom presentation, and other characteristics of the disease (e.g., diagnosis, treatments, biomarkers).

Buchanan et al. (2010) compared demographic and clinical characteristics of individuals with MS across different racial/ethnic groups using a database that included

26,967 Caucasians, 715 Latinos, and 1,313 African Americans. Results indicated that Caucasians tend to be older at the onset of symptoms and time of diagnosis (30.1 years and 37.4 years respectively) than Latinos (28.6 years and 34.5 years respectively) or African Americans (29.8 years and 35.8 years respectively). Additionally, a higher proportion of Caucasians reported issues with mobility and bladder/bowel function than Latinos. Notably, larger proportions of African American (45.8%) and Latino (44.2%) individuals reported experiencing depression compared to Caucasians (38.7%) (Buchanan et al., 2010).

Al-Kawaz et al. (2017) explored the differential impact of the MS disease process on cortical and gray matter structures in African Americans (n = 44) and Caucasian Americans (n = 54). Groups were evenly matched according to age, gender, disease, and treatment duration. Results revealed significant differences between groups. African American participants reported greater disability, as measured by the Expanded Disability Status Scale (EDSS) scores, and lower cortical thickness in multiple regions, including the occipital lobe. Caucasian Americans exhibited reduced thalamic volume compared to African Americans. The groups did not differ in terms of T2 hyperintense lesion volume. Finally, both groups had a strong negative relationship between disability and total thalamic volume, such that as disability increases, thalamic volume percentage decreases (Al-Kawaz et al., 2017).

Lichtman-Mikol et al. (2019) conducted a study on Caucasian Americans (n = 32) and African Americans (n = 27) with RRMS to evaluate how race impacts MS pathobiology. Correlations were analyzed between MRI to assess gray matter and OCT to assess neuroaxonal health (by measuring the thickness of the ganglion cell inner plexiform layer). MRI and OCT measures were significantly correlated for Caucasian Americans but not for

African Americans. The authors argue that more large-scale research is needed to identify biomarkers in MS that are reliable for all racial groups (Lichtman-Mikol et al., 2019).

Etiology. The specific causes of MS are still unknown. It is widely accepted among researchers that multiple mechanisms contribute to the development of MS. However, an accumulation of decades of research suggests that it may result from an abnormal immune response to an infectious or environmental trigger in an individual who is already genetically susceptible. Current research into MS etiology is focused intensively on investigation of immunology, genetics, demographics, infections, and the environment (e.g., industrial toxins, diet, trace metal exposures) and how these relate to the development of MS (Garg & Smith, 2015; Milo & Miller, 2014; National MS Society, 2018). Known risk factors for MS include smoking, low vitamin D levels, and childhood obesity (Hedström, Olsson, & Alfredsson, 2016; Munger et al., 2017). MS is not inherited directly, but genetics do play a role. According to the National MS Society (2018), the risk of a person in the general population developing MS is 1 in 750 while the risk of a person with a direct relative (parent, sibling, child) is 1 in 40. The risk for identical twins is 1 in 4 (Ebers, 2013; Patsopoulos et al., 2011). Researchers have surmised that infectious agents may trigger MS because viruses are well known to cause demyelination and inflammation. No virus or bacteria has been positively identified as an MS trigger at this time. However, more than a dozen infectious diseases have been or are currently being studied, including measles, canine distemper, human herpes virus-6, Chlamydia, pneumonia, and Epstein-Barr (Milo & Miller, 2014; National MS Society, 2018).

Diagnosing MS. The current diagnostic criteria for MS, the McDonald criteria, require evidence of two or more lesions or damage in the CNS, distributed in space and time

(by at least one month), through the use of two or more objective methods of assessment (e.g., MRI, spinal fluid analysis, EEG) (McAlpine, 1972; McDonald et al., 2001; Milo & Miller, 2014; Polman et al., 2005; Polman et al., 2011; Thompson et al., 2018). Additionally, part of the diagnostic process involves ruling out any other possible causes.

MS phenotypes. CIS is often a precursor to MS and is characterized by a single episode of neurologic symptoms, caused by inflammation or demyelination in the CNS, lasting at least 24 hours (Milo & Miller, 2014; National MS Society, 2018). Approximately 80% of individuals who have CIS will go on to develop MS. There are three MS subtypes: relapsing-remitting (RRMS), secondary-progressive (SPMS), and primary-progressive (PPMS). The most commonly occurring disease course (85% at the time of diagnosis) is relapsing-remitting, characterized by clearly defined acute attacks or worsening neurologic function, followed by full or partial recovery periods (Milo & Miller, 2014; National MS Society, 2018). During these periods of remission, symptoms may continue, improve, or disappear entirely, and there is no disease progression between attacks. Depending on the state of disease course, RRMS can be classified as active (with relapses and/or evidence of new activity as measured by MRI) or inactive, and worsening (a specified period of time marked by increased disability following a relapse) or not (Milo & Miller, 2014; National MS Society, 2018). SPMS initially looks similar to RRMS but becomes more progressive. Most people with RRMS will eventually be diagnosed with SPMS. SPMS is characterized by fewer inflammatory changes (relapses and/or new MRI indicated inflammation in the CNS) and a progressive decline in neurologic function over time. SPMS is also classified as active or not and with or without progression (evidence of worsening over time). PPMS is characterized by a steady worsening of symptoms, especially neurologic function, and the

absence of clearly defined attacks from the outset. There is an accumulation of deficits and a progression of disability that continues for months or years, although brief periods of minor improvement or stability can be observed.

Symptoms. Symptoms of MS are unpredictable and can vary greatly from one person to another. Visual impairment is a characteristic symptom manifestation in MS, and often the first reported symptom (McAlpine, 1972). One early warning sign of MS is optic neuritis (ON), typically occurring unilaterally, which is an inflammation of the optic nerve. ON often results in temporary blurring or loss of vision and often painful eye movement. ON may also cause a scotoma (blind spot) in the center of one's visual field. Other visual symptoms commonly experienced include diplopia (double vision) and nystagmus. Dizziness and vertigo are also reported.

Fatigue and bladder dysfunction are two of the most common symptoms and are reported by as many as 80% of individuals with MS. Depression is often comorbid with MS and is more common than in the general population. Greater than 50% of individuals with MS will experience a major depressive episode.

Cognitive functioning is significantly impacted in MS, including processing speed, new learning and memory, and executive functions. Progression of disease in MS is strongly associated with cognitive decline across various domains. Approximately 65% of individuals diagnosed with MS will experience some form of cognitive impairment (LaRocca, 2016; Sumowski & Leavitt, 2013).

Motor symptoms are also very common in MS and include ambulation problems such as a lack of coordination, difficulty walking, or maintaining balance. Sensory symptoms include tingling, numbness or weaknesses in the limbs (typically occurring on one side of the

body at a time), and neuropathic pain in the face or body (McAlpine, 1972). Tremors can manifest as uncontrollable shaking of the limbs, trunk, voice, eyes or head. Spasticity, which includes feelings of stiffness and a wide range of involuntary muscle spasms, can range from mild to severe. Dysarthria is a speech disorder caused by muscle weakness and is characterized by slow or slurred speech or decreased volume. Sexual dysfunction is also experienced by many individuals with MS in the form of impaired arousal, erectile dysfunction, reduced vaginal lubrication or sensory changes (Sanders, Foley, Larocca, & Zemon, 2000).

Secondary symptoms can arise from primary symptoms, for example a bladder infection or decreased bone density due to inactivity. For some primary and most secondary symptoms, treatment options are available and will be mentioned briefly.

Prognosis and Treatment Options in MS. While there is no cure for MS, there are treatment options available for managing relapses, and primary and secondary symptoms as well as disease-modifying therapies (DMTs) to help slow the progression of the disease. Evidence indicates that the earlier that treatment begins in the disease process, the greater likelihood of positive long-term prognosis. For all individuals with MS, a comprehensive multi-disciplinary treatment plan is needed. Treatments for MS can fall into one of several subgroups: disease and symptom management, psychosocial support, and rehabilitation. Disease and symptom management involves the use of medications (DMTs) to modify disease course, manage symptoms, and treat relapses. The US Food and Drug Administration (FDA) has approved DMTs that have been tested in clinical trials and were found to reduce the number of relapses, limit new disease activity, and delay progression of disability (as evidenced by MRI) (Freedman et al., 2013). DMTs can be administered in injectable,

infused, or oral form. Injectable medications include interferon beta-1a, interferon beta-1b, glatiramer acetate, and peginterferon beta-1a (National MS Society, 2018). Oral medications include teriflunomide, fingolimod, dimethyl fumarate, cladribine, and siponimod (National MS Society, 2018). Infused medications include alemtuzumab, mitoxantrone, ocrelizumab, and natalizumab (National MS Society, 2018).

For treatment of severe relapses, such as severe ambulatory problems or loss the vision, high-dose, intravenous corticosteroids followed by an oral prednisone are often prescribed in order to reduce inflammation and shorten the length of a relapse (Freedman et al., 2013; National MS Society, 2018). However, corticosteroids do not show evidence of any long-term benefit in treating MS (National MS Society, 2018). There are also numerus medications and treatment options available to treat symptoms of MS, including bladder problems, psychological dysfunction, fatigue, sexual problems, pain, bowel dysfunction, spasticity, tremors, ambulatory issues, dizziness and vertigo (National MS Society, 2018; Van Kessel, 2008). In addition, rehabilitation (occupational, physical, speech, etc.), increased psychosocial supports, and an improved health and wellness model for daily living are essential components to comprehensive treatment in MS (National MS Society, 2018).

Prognosis for MS is variable from person to person but continued dysfunction in the CNS logically leads to increased disability and worsening of disease course over time (Milo & Miller, 2014). Milo and Miller (2014) report that life expectancy in MS is reduced by seven to ten years, on average, and that medical complications are the cause of death in 50% of cases, while the remaining causes include suicide and other common causes seen in the general population. Certain prognostic factors have been found to predict worse prognosis and these include age of onset (after 40), gender (male), ethnicity (Asians or African

Americans), motor, and cerebellar or sphincter symptoms, which are present at initial diagnosis (Milo & Miller, 2014). In addition, characteristics of attacks are indicative of prognosis such that prognosis is worse if there are more frequent attacks during the early years post-diagnosis, a shorter interval between the first two attacks, incomplete recovery from initial attacks, rapid progression of disability in the early years, and progressive disease and cognitive impairment from onset (Milo & Miller, 2014). Finally, presence in cerebrospinal fluid (CSF) of oligoclonal immunoglobulins and initial MRI indications of high disease burden or gadolinium (Gd) enhancement are also factors that increase the likelihood of poorer prognostic outcomes (Milo & Miller, 2014).

MS and Visual Function. Visual impairment is a key component of the MS disease process (Galetta & Balcer, 2013). MS lesions can affect any part of the neural networks involved in vision and therefore, can cause a variety of neuro-ophthalmic manifestations. While acute optic neuritis (AON) is often a precursor or primary manifestation, visual impairments characterized by structural, axonal, and other neuronal loss often occur in patients without a history of ON. Blurring and visual distortions are two of the most commonly reported visual abnormalities reported in MS. Other commonly reported symptoms include pain, visual field loss, and flashes of light. At later stages of the disease, MS can lead to complete blindness. Compromised components of the visual system have been associated with worsening of the disease. Treatment of visual deficits has been found to improve quality of life in MS patients. The frequency with which the afferent visual pathway is involved in the disease process of MS makes it a very useful model to study the pathophysiological mechanisms involved (Galetta & Balcer, 2013). Visual evoked potentials elicited through electrophysiological recording have been used to study the integrity of neural

mechanisms and pathways in MS since the earliest stages of research into the disease and will be used in the current study to advance the current knowledge of early-stage visual processing in MS (Duwaer & Spekreijse, 1978; Halliday, McDonald, & Mushin, 1973a; Regan; 1989).

Magnetic Resonance Imaging as a Diagnostic Tool in MS. Magnetic resonance imaging (MRI) is considered one of the most sensitive, objective, non-invasive methods of imaging the brain, spinal cord, and body. MRI techniques are used as primary diagnostic tools in MS. MRI is a radiological measure, which uses magnetic fields and radio waves to measure the relative water content in tissues throughout the body. The quality of images produced by MRI is dependent on the strength of the magnet used (measured in Tesla, *T*). Most conventional machines use 1.5*T* or 2.0*T* while MRI machines used in research often use even stronger magnets (Takemura et al., 2016). Newer MRI techniques, using even higher-level scanners (e.g. 3*T*), have improved recognition of cortical lesions affected in MS (Graham & Kilstorner, 2016).

MRI studies on white matter integrity support reduced cortical excitatory activity in MS. White matter (WM) is a tract of myelinated axons and WM signaling is known to involve multiple neurotransmitters, including glutamate, ATP, GABA, and norepinephrine (Butt, Fern, & Matute, 2014). Thus, WM integrity modulates the signaling activity of neurotransmitters. Butt, Fern, and Matute (2014) published a review of neurotransmitter signaling and included results from studies assessing the activation of nonglutamatergic receptors in WM. Results provide evidence for reduced excitatory activity associated with activation of all receptors listed (e.g. GABA, dopamine, 5-HT, adreno-receptors). Given this association with attenuated excitatory activity, damage (or demyelination) to WM, as seen in

MS, would thus logically lead to an attenuation of excitatory cortical processing (Butt, Fern, & Matute, 2014).

Various MRI scanning methods exist and each type measures human tissue in a different way. A characteristic pathophysiological process in MS is myelin damage, which is also a hallmark of disease status. The MRI is particularly useful in detecting demyelination in the CNS (cerebral white matter and spinal cord), making it a powerful diagnostic tool in MS. The myelin layer protecting nerve cell fibers is typically fatty and repels water. However, damage to myelin results in stripping away the fat, permitting the damaged location to hold more water. The resulting MRI scan will appear as either a bright white spot or a darkened area, depending on which scan is used. T-1-weighted images, with gadolinium injected intravenously to enhance scanning sensitivity, provide information about current disease activity by highlighting areas of active inflammation. T-2-weighted MRI images supply information about disease burden or lesion load. *Fluid attenuated inversion recovery* (FLAIR) images are used to better identify brain lesions associated with MS. Magnetic resonance spinal cord imaging can be used to identify pathology in the spinal cord, which can help establish an MS diagnosis by demonstrating damage that has occurred in different parts of the CNS (DIS) at different times (DIT). Magnetic resonance *diffusion tensor imaging* (DTI) provides quantitative measures of structural changes in the brain and permits the localization of lesions within white matter tracks. DTI permits evaluation of axonal and myelin damage in MS (Naismith et al., 2012).

MRI is particularly useful for individuals diagnosed with CIS, as the number of lesions on an initial MRI can help assess risk for developing a second attack (Leocani, Rocca, & Comi, 2016). Additionally, some treatment options have shown efficacy in

delaying the onset of a second neurologic episode. Thus, MRI can be a critical method for early diagnosis and for monitoring treatment efficacy. Once a diagnosis has been established, MRI is often used annually as a way to track disease progression. Despite its obvious strengths in research and clinical applications of MS, MRI is not as conclusive and broadly useful as people previously thought (Graham & Kilstorner, 2016). Approximately 5% of individuals with clinically definite MS will not display lesions on MRI scans at the time of diagnosis (National MS Society, n.d.). Additionally, many lesions identified on MRI may be located in areas of the brain that do not produce symptoms, which makes it difficult to find correlations between imaging results and an individual's clinical disease presentation. Thankfully, there are additional tools and techniques that are used to assess MS.

Additional Techniques for Diagnosing & Assessing MS. When diagnosing MS, individuals will typically provide a thorough report of their medical history (most notably to ensure rule-outs of other medical conditions), receive a thorough neurologic examination, undergo spinal fluid analysis and several other tests, including blood tests. Spinal fluid analysis (SFA) involves obtaining spinal fluid using a lumbar puncture inserted into the lower back in order to assess CSF, which is typically impacted by MS. Evaluation of the CSF in MS typically shows the existence of a specific group of proteins called oligoclonal bands and an elevation of the level of protein, which is indicative of an abnormal immune response in the CNS and can be caused by myelin damage. Notably, the results of the SFA are not definitive of MS, as abnormal CSF activity has been found in many other diseases and 5-10% of individuals with clinically definite MS do not show any CSF abnormalities at all.

There are other techniques used to study structure and function in the CNS in MS. Optical coherence tomography (OCT) is an extremely useful structural biomarker of disease

activity in MS (Balcer, Miller, Reingold, & Cohen, 2015; Di Maggio et al., 2014). It is a method for quantifying and tracking axonal loss in individuals with MS (Di Maggio et al., 2014). Results from OCT studies indicating visual system dysfunction in MS will be discussed later in this section.

Magnetoencephalography (MEG) is another method that has been used to assess abnormalities in the CNS in MS (Kotini, Anninos, & Tamiolakis, 2007; Tewarie et al., 2017). The MEG measures fluctuations in magnetic field strength in the cerebral cortex, which are produced by movements of cerebral ions throughout the neuronal cell body creating a current dipole, which follow changes in membrane potential. The orientation of the current dipole affects magnetic field measurements and is therefore a critical component to study. Results from MEG studies indicate abnormal cell activity in the cortex of individuals with MS. MEG researchers who prefer this method over EEG recording argue that EEG is subject to interference from tissues and fluids that lie between the cortex and the scalp (Kotini et al., 2007). Some researchers argue that there is value in collecting MEG and EEG simultaneously, while proponents of EEG recording argue that MEG is not as simple, short-duration, or cost effective as the VEP and, therefore, is less useful with clinical populations and in clinic settings (Goldenholz, 2009).

Evoked potentials (visual, auditory, somatosensory, motor, multimodal) are often used as functional measures of analysis in MS. They are especially helpful in the diagnostic process in that they can identify clinically silent lesions that do not show up with other techniques (Gronseth & Ashman, 2010). Additionally, VEPs represent brain activity on the order of milliseconds and are therefore a uniquely sensitive measure of real-time neurologic function.

Overview of The Visual System.

The human visual system is comprised of lateral and direct-through pathways, which carry inhibitory and excitatory information through the cerebral cortex, respectively (Ratliff & Zemon, 1984; Zemon & Ratliff, 1982, 1984).

Retina. The human visual system is a part of the CNS that begins at the retina and extends to the occipital lobe, primary visual cortex (V1) and beyond. Incoming visual information is coded in the retina by several types of cells including photoreceptors, bipolar cells, horizontal cells, amacrine cells, and ganglion cells. Information is first received by the retina in the form of light, which is then converted into electrical signals which are transmitted to the brain via the optic nerve. The initial signals are processed by photoreceptor cells, which are comprised of rods and cones. Rods assist in low-level light perception and provide monochromatic information while cones aid in the perception of detailed form and color.

Retino-Geniculo-Cortical Pathway. The primary human visual system processes information via the retino-geniculo-cortical pathway. The majority of signals that exit the retina travel to the dorsal lateral geniculate nucleus (LGN) of the thalamus, and on to V1 in the occipital lobe. The LGN is the visual information relay center in the brain. All information coming into the LGN is monocular, and this information remains segregated at the level of the LGN. There are three main types of ganglion cells that project to the LGN: M, P, and K cells. Each cell type synapses in different layers of the LGN and form different, parallel pathways. The LGN consists of six neuroanatomical layers: magnocellular (M) pathway is comprised of the two ventral layers, parvocellular (P) pathway is inclusive of the

four dorsal layers, and koniocellular (K) pathway is formed at the borders of each layer. The K pathway processes information from short-wavelength cones (Kaplan, 2004).

Parallel pathways. The M and P pathways are two important parallel pathways which convey information from their respective ganglion cells (Purves et al., 2001). They begin at the retina and remain separate through V1 (Hubel & Livingstone, 1987; Liu et al., 2006; Livingstone & Hubel, 1987; Merigan & Maunsell, 1993; Tootell & Nasr, 2017). These pathways process and respond differently to light and have distinct roles in visual perception. The M pathway plays a critical role in the perception of brightness, form, depth, and movement. This pathway is sensitive to low-contrast stimuli, has low spatial resolution, high temporal resolution, and is characterized by transient responses (Liu et al., 2006). The M pathway projects to the dorsal stream of the visual system. M ganglion cells have larger receptive fields and axons and therefore they have faster conduction velocities than do P cells (Purves et al., 2001). The P pathway plays an essential role in the perception of color, visual acuity, spatial details and object identification. This pathway is sensitive to high-contrast stimuli, has high spatial resolution, low temporal resolution, and is characterized by sustained responses (Liu et al., 2006). The P pathway projects to the ventral stream of the visual system (Butler et al., 2007; Calderone et al., 2013; Kaplan, 2004; Kaplan & Shapley, 1986; Livingstone & Hubel, 1988; Tootell et al., 1988; Tootell & Nasr, 2017). Given that these pathways are selective in response to distinct visual stimuli, they are useful in the identification of visual processing deficits (Hartline, 1938; Purves et al., 2001; Schiller, 1982).

Brightness and Darkness Perception. The M and P pathways can be further divided into another parallel subsystem within the visual cortex: the ON/OFF system, which

determines brightness and darkness perception via activation of ON or OFF cells, respectively (Hartline, 1938; Schiller, 1982; Zemon & Gordon, 2006). These ON/OFF responses are shaped by excitatory and inhibitory mechanisms. ON cells become activated by positive contrast or light objects on dark backgrounds (e.g., white chalk on a blackboard) whereas OFF cells are activated by negative contrast or dark objects on light backgrounds (e.g., black ink on white paper).

Visual Evoked Potentials (VEPs). The visual evoked potential (VEP) is an electrophysiological response, elicited by a visual stimulus, that has been used to assess functional integrity of neural mechanisms and pathways (Butler et al., 2001; Conte & Victor, 2009; Duwaer & Spekreijse, 1978; Greenstein, Seliger, Zemon, & Ritch, 1998; Lachapelle, Ouimet, Bach, Ptito, & McKerral, 2004; Regan, 1989; Siper et al., 2016; Weinger et al., 2014; Zemon et al., 2008; Zemon & Gordon, 2018). It reflects excitatory and inhibitory postsynaptic activity in the visual cortex. VEPs are measured from the surface of the head and originate primarily in the cerebral cortex. Three electrode sensors are placed on the head to measure the signal, which is first generated in the retina, modified in and relayed from the LGN, and finally arrives in the visual cortex. The VEP is extracted from ongoing brain waves (continuous EEG recording) by means of signal averaging or methods of frequency analysis (Ratliff & Zemon, 1982; Regan, 1989).

The VEP is a uniquely useful measure of visual function in part because it reflects real-time brain processes on the order of milliseconds, unlike other techniques (e.g., MRI, OCT), thereby allowing for the quantitative evaluation of the functional integrity of connections and mechanisms within the visual system. VEPs are frequently used in the assessment of neural dysfunction in a variety of clinical populations, including multiple

sclerosis, schizophrenia, traumatic brain injury, autism, glaucoma, and epilepsy, among others (e.g., Regan, 1989; Duwaer & Spekreijse, 1978; Zemon, 1984; Lachapelle et al., 2004; Kim, Zemon, Pinkhasov, Gordon & Marks, 2002; Kim, Zemon, Saperstein, Butler et al., 2001,; Greenstein et al., 1998; Zemon et al., 2008; Ratliff & Zemon, 1984; Conte & Victor, 2009; Weinger et al., 2014; Siper et al., 2016).

The visual system is comprised of linear and nonlinear mechanisms, which can be examined via the VEP. VEPs are elicited by simple luminance stimuli, such as a flash of light, or more complex stimuli, such as a spatial pattern that is contrast-reversed in time (pattern stimulation) (Odom et al., 2010). Different pathways and mechanisms within the visual system are elicited based on choice of stimulation and analysis conditions. Mechanisms assessed include lateral-inhibitory circuitry (e.g., GABAergic ($GABA_A$) intracortical inhibition), frequency-selective mechanisms, and contrast gain control processes (Zemon et al., 1980, 1986; Zemon & Gordon, 2006; Zemon & Gordon, 2018).

Electrogenesis of the Visual Evoked Potential. VEP responses collected via EEG recording are measures of extracellular activity. Specifically, VEP's reflect the sum of excitatory and inhibitory postsynaptic potentials (EPSPs and IPSPs), which produce responses at the top of the head that are volume conducted through neural tissue and non-neural layers surrounding the brain. In detail, the VEP reflects extracellular currents produced by EPSP and IPSP activity occurring on apical dendrites of pyramidal neurons in superficial layers of the neocortex (Zemon et al., 1980; Zemon & Gordon, 2018). Excitatory and inhibitory signals received from pyramidal cells are modulated by apical dendrites. Apical dendrites possess dipole-like properties and are often located perpendicular to the pial surface, thereby creating a large net extracellular current flow. EPSPs create surface

negativity and IPSPs create surface positivity (Zemon et al., 1986). These structures dominate the superficial layers of the visual cortex and provide major input to "complex" neurons. These factors are all components essential to the generation of the visual evoked potential (Zemon et al., 1986).

Role of GABA in electrogenesis of VEPs. Evidence from animal experiments has demonstrated that $GABA_A$ plays a significant mediating role in generating the shape of VEP waveforms (Zemon et al., 1980, 1986). These studies also provide support for the electrogenesis of the VEP and for the excitatory/inhibitory mechanisms involved in early VEP deflections. Studies using bicuculline to selectively block $GABA_A$ receptors in the visual cortex of cats have demonstrated an enhancement of the negative wave and attenuation or elimination of the subsequent positive wave in the VEP response (Zemon et al., 1980, 1986).

Transient and Steady-State Visual Evoked Potentials. VEPs are classified as *transient* (tVEP) or *steady-state* (ssVEP) based on the stimulus frequency and temporal waveform used to elicit the response (Regan, 1966, 1989). Transient VEPs (tVEPs) are produced by abruptly modulating the contrast or luminance of a stimulus (e.g., uniform field of light or spatial pattern) at low frequencies (e.g., 1 Hz). This produces a conventional VEP waveform with a series of positive and negative deflections in the first few milliseconds of the response. This activity is typically over by the next stimulus change, resulting in the transient nature of the response. Transient VEP data are analyzed using magnitude (peak-to-trough amplitude) and latency (time from stimulus change to a peak or trough) measures for early peaks and troughs in the waveform. These time-series data can also be represented in the frequency domain by means of applying a Fourier transform. This results in frequency

components expressed in terms of amplitude (microvolts, μV) and phase (degrees) values. The current study utilizes tVEPs.

Steady-state VEPs are elicited by a moderate to high frequency of temporal modulation (typically ≥ 3.5 Hz), such that the response to one stimulus change is not complete when the next change occurs (Regan, 1989; Zemon & Ratliff, 1982, 1984). The result is an overlap of successive responses, yielding a smoother oscillatory waveform, with less distinct character. The shape of responses produced by ssVEPS is approximately sinusoidal and power in the response is concentrated in the first few harmonic frequency components with some additional power in a few later frequency components. The response at the stimulus frequency is referred to as the fundamental component (also known as the first harmonic – a harmonic is an integer multiple of the stimulus frequency). These frequency components are extracted from the ssVEP by means of Fourier analysis and quantified in terms of amplitude and phase data (or equivalently, in terms of cosine and sine coefficients).

Neural Mechanisms. The stimulus conditions used in the current VEP study were designed to examine specific neural processes. These stimuli and the mechanisms they aim to measure are described below.

Transient VEP to contrast-reversing checkerboards (tVEP-CR). Transient VEPs are typically elicited using a conventional checkerboard pattern of high contrast (85-100%) with light and dark elements reversed abruptly at the start of the stimulus cycle and again halfway through the cycle (1 Hz square-wave temporal signal). The stimulus has high contrast in order to produce a clear response relative to noise in the recording. Traditionally, contrast-reversal checkerboard stimuli are presented over 60-second runs. In this study, runs collect

only ~2 seconds of EEG data. Ten short runs are collected, which yield independent estimates of the neural response. Given independent responses, the noise level decreases as the square root of the number of responses averaged. The ten responses collected are averaged to obtain a mean value for the overall response, represented in a mean response waveform. These processes allow for proper statistical estimation, collected on the entire response, and presented in the frequency domain. One long-standing role of this response is to aid in diagnosis and monitoring of multiple sclerosis (Fuhr et al., 2001). Results from such studies have consistently indicated that individuals with MS display a delayed P100 (P_1) peak time compared to healthy controls (Galetta & Balcer, 2013; Kilstorner et al., 2012; Thurtell, et al, 2009).

The tVEP-CR is valuable as a functional measure because so much is known about its physiological/biochemical origins, however, it is complicated by the fact that the response is a composite of activity from multiple parallel streams (Zemon & Gordon, 2018). As mentioned previously, precortical parallel pathways are present throughout the visual system (e.g., ON/OFF, M/P) and these contribute to the cortical responses in the tVEP-CR. While techniques have been developed to examine separate contributions of ON and OFF pathways, contrast-reversal stimuli drive them both (Zemon & Gordon, 2018). Bright checks (e.g., when the luminance of the check is higher than the luminance of the background) elicit ON pathway responses, while dark checks (e.g., when the check is lower in luminance than the background) elicit OFF pathway responses. (Zemon, Gordon, & Welch, 1988; Zemon & Gordon, 2006; Zemon et al., 1995). Similar to the ON/OFF dichotomy, the tVEP-CR strongly drives both the magnocellular and parvocellular pathways (Zemon & Gordon, 2018). This stimulus is also believed to drive the koniocellular pathway (Kaplan, 2004).

Thus, the strength of this technique lies in the extensive research that exists on the relationship of its time-domain features to excitatory and inhibitory processes in the cortex. A drawback of this technique is the multiple contributions from different pathways and the interaction of these contributions in the recording (Zemon & Gordon, 2018).

Transient VEP waveforms. The transient VEP waveform is analyzed through assessing magnitude and latency of positive and negative peaks and troughs, which reflect excitatory (glutamate) and inhibitory (GABA, gamma-amino-butyric acid) activity in the cortex. Evidence from numerous studies (MEG, EEG in conjunction with fMRI, invasive EEG recordings, etc.) indicates that prominent early deflections in transient waveform originate in primary visual cortex (V1) (Zemon & Gordon, 2018). The positive and negative peaks and troughs found in a VEP waveform are strongly believed to represent different cellular activities (Zemon & Gordon, 2018). The earliest component in the waveform (P_0), the initial positive deflection, is thought to reflect initial depolarization of neurons in the recipient cortical areas of the primary visual cortex from the LGN and occurs around 60 ms in healthy controls (Zemon et al., 1986; Zemon & Gordon, 2018). The subsequent negative deflection (N_0; occurs around 75 ms) reflects excitatory post-synaptic activity with spreading of activation in the supragranular layers (2 and 3) of V1. The second positive peak, occurring around 100ms (P_1) reflects the sum of IPSP activity at V1 (e.g. GABAergic activity and hyperpolarization) (Zemon et al., 1980, 1986; Zemon & Gordon, 2018). The second negative deflection (N_1, occurs around 135 ms) originates in later EPSP activity. It is worth noting that activity from these peaks and troughs do not reflect contributions from isolated mechanisms, as there is considerable overlap in the time domain (Zemon & Gordon, 2018).

Frequency domain measures of the tVEP. The conventional method of analysis is to examine latency and amplitude of peaks and troughs in the time domain in a complex and possibly noise waveform. This requires subjective selection of a few time points in the record to measure these values. Additionally, these analyses do not provide information on the statistical significance of the response and many additional response details cannot be explained or parsed out. Novel techniques have been developed in our laboratory to analyze these values in ways that permit greater confidence in the subjective interpretations and in the statistical significance of the response, while also allowing for additional information to be extracted from the response. One technique, magnitude-squared coherence (MSC), can be used to determine signal power relative to total power in the response. This statistic, which ranges from 0 to 1, provides information about the strength of the response for each frequency mechanism of interest. Another technique, the 24th harmonic delay estimate, can be used to calculate an estimate of signal transmission time related to P100 peak time based on phase data in the frequency domain (e.g., the 24th harmonic frequency component) (Zemon & Gordon, 2018).

The MSC statistic was developed and validated for VEP analysis by Zemon and colleagues (1999, 2009). These researchers conducted extensive repeated testing on individual participants in order to examine the mechanisms contributing to the overall VEP response. They also completed a principal component analysis (PCA) in order to identify the relevant frequency bands (Zemon et al. 2009; Zemon & Gordon, 2018). Comparisons between PCAs and time-domain measures indicate that power found in high-frequency bands can be used as objective measures of excitatory input into the visual cortex (Zemon & Gordon, 2018). Another benefit of MSC in data analysis of the tVEP-CR technique is that it

permits comparisons of responses between individuals by providing a normalized adjustment for different scales of response obtained between individuals (Zemon & Gordon, 2018).

Zemon and Gordon (2018) recently conducted two studies in order to identify objective measures that capture the entirety of information contained in the response through transformation of data from the time domain to the frequency domain via a discrete Fourier transform (DFT). These studies aimed to further assess the relevant frequency mechanisms involved in the generation of the response (MSC and power) and to provide evidence for estimations of signal transmission delay. Time delay estimates were calculated based on linear regression of phase vs. harmonic frequency plots and from phase of a single harmonic component (24^{th} harmonic) (Zemon & Gordon, 2018).

The overarching purpose of the two studies was to test the validity of a novel short-duration 2-second stimulus presentation to be used as a rapid and objective clinical tool. Study 1 involved extensive tVEP-CR testing of a small sample of healthy controls ($N = 10$, $M_{age} = 22.3$ yrs) using a long-duration (32-second) stimulus repeated 40 times and it was designed to identify and characterize neural mechanisms contributing to the response. Study 2 involved testing a much larger sample ($N = 89$, $M_{age} = 21.7$ yrs) using the short-duration stimulus (2-s condition repeated 10 times) and a conventional long-duration stimulus (60-s condition). Study 2 was designed to replicate Study 1 findings from the smaller sample while also providing evidence for the reliability and validity of the short-duration stimulus.

Results supported the reliability and validity of the short-duration stimulus. Study 1 yielded more accurate measurements of the driven frequency components. Study 2 restricted the frequency range to 2-48 Hz, because of the limited amount of data collection in the 2-s test and because there is a wider noise bandwidth associated with the 2-s EEG epochs

(Zemon & Gordon, 2018). This choice of restriction was supported by analysis of MSC values, which indicated a lack of significant estimates in the high-frequency range. Results from Study 2 yielded four distinct frequency bands, defined as: Band 1, 6-12 Hz; Band 2, 14-28 Hz; Band 3, 30-40 Hz; Band 4, 42-48 Hz (Zemon & Gordon, 2018). The 2nd and 4th harmonic components were excluded because they did not fall into a specific band (Zemon & Gordon, 2018). The current study will utilize these four frequency bands for evaluation of the tVEP-CR response in the frequency domain.

Comparison of waveform analyses in 2-s vs. 60-s conditions indicated significant differences in responses collected between the long-duration and short-duration runs (Zemon & Gordon, 2018). Zemon and Gordon (2018) note that the long-duration stimulus distorts the character of an individual's response and that adaptation effects are involved in this longer condition. Study 2 analysis revealed shorter P100 peak times (by 2-3 ms), increased amplitudes and higher power estimates in the 2-s condition. The ten epochs collected in the 2-s condition are separated in time and can be considered independent estimates of the population response, whereas the 60-s condition, in which the 60-s EEG epoch is divided into 10 6-s epochs, includes correlations among epochs, which can lead to an overestimation of coherence (Zemon & Gordon, 2018).

Analysis of amplitude and phase vs. frequency plots indicated extremely low amplitudes above 40 Hz, which the researchers hypothesize may indicate a lack of significant response activity present in the high frequency range (Zemon & Gordon, 2018). Additionally, the researchers discovered that regression equations restricted to each segment result in a better fit than a single regression line fit to the entirety of phase data. The slope of each phase vs. frequency plot yielded a time delay estimate. This is due to the fact that fixed

transmission time added to the system produces a linear phase shift vs. frequency as frequency increases and that different linear segments with varying slopes may represent distinct neural mechanisms. Results also indicated that most participants displayed steeper slopes at lower frequencies, suggesting that the response at these low frequencies corresponds to later cortical activity. Although harmonic frequency components were low in amplitude above 40 Hz, the researchers note that MSC peaks were consistently observed in the high-frequency range (Zemon & Gordon, 2018).

Zemon and Gordon (2018) also analyzed associations of time- and frequency-domain magnitude measures. Results indicated that using the square root of power in Band 2 yields accurate and objective estimates of N75-P100 amplitude, with approximately 90% of variance explained by the linear relation. Results also revealed correlations between P60-N75 amplitude with Bands 2, 3, and 4, while P100-N135 amplitude correlated best with Band 1. MSC correlated best with power within each band (Zemon & Gordon, 2018).

Zemon and Gordon (2018) also attempted to identify a frequency-domain measure that could be used as an objective estimate of P100 peak time. P100 is the time-domain measure relied on most often in clinical settings, where extensive testing is not feasible and subjectivity in waveform analysis is risky due to noise-related variability. The authors used the 24^{th} harmonic frequency component to calculate a time (phase) delay measure to estimate P100 peak time. To use this equation, the millisecond value calculated is obtained from the phase value of the 24^{th} harmonic divided by 360^{o} and added to multiples of this harmonic's period (41.67ms) to yield a value between 80 and 110 ms for healthy adult observers (Zemon & Gordon, 2018). The authors also note that they obtained a high correlation coefficient as well as exceptional absolute agreement between the time delay estimate derived from the 24^{th}

harmonic and the peak time in the waveform (Zemon & Gordon, 2018). The current study will utilize this estimate of P100 time delay based on the 24th harmonic.

Measures of neural noise and magnitude of responses. Techniques are utilized in this study to evaluate neural noise and the strength of a response relative to noise. Neural noise refers to random variability in neural activity and noise related to non-neural factors (e.g., movement artifacts and environmental noise). Noise is estimated by the amount of variability in a participant's repeated VEP responses to the same stimulus. Signal-to-noise ratio (SNR) is a measure of relative strength or magnitude of a response and is computed as the ratio of amplitude to noise at a given frequency. An SNR greater than one is indicative of a significant response at an alpha level of .05. The T_{circ}^2 statistic is a multivariate statistic, calculated using sine and cosine coefficients from a frequency component of interest, in order to estimate the variability (noise) in a set of responses at that frequency (Victor & Mast, 1991), and for a single frequency component it is algebraically equivalent to MSC (Dobie and Wilson, 1993). This statistic is represented as the radius of a circle in a sine-cosine plot which surrounds a vector-mean response indicating a 95% confidence region for the harmonic component of interest. If the circle includes the origin in the sine-cosine plot, the response is not significant at the .05 level. Another technique used in this study is Fourier analysis. Through the use of a discrete Fourier transform (DFT), data can be transformed from the time-domain to the frequency-domain. Frequency-domain techniques allow for more detailed analysis of the functional integrity of neural pathways and mechanisms as well as a more objective analysis of the responses collected.

MS and Visual Function

Afferent Visual Pathway as Model of MS Pathophysiology. The afferent visual pathway receives, relays, and processes visual stimuli. It consists of the eye, optic nerves, chiasm, optic tracts, lateral geniculate nucleus (LGN), optic radiations, striate cortex, and visual association cortices. MS lesions can affect any point along the afferent visual pathway. Disturbances along this pathway occur in 80% or more of individuals with MS at some point during their disease course (Beh et al., 2016). Lesions in the afferent pathway often result in symptoms that are prominent and impairing and, therefore, cause the individual to seek medical guidance. This process often results in a diagnosis of MS, after the individual has visited a neurologist. Selective deficits have also been found in M and P streams in MS, suggesting that the disease process may differentially affect these neural subsystems (Thurtell et al., 2009).

Optic Neuritis and MS. Optic neuritis (ON) is an inflammation of the optic nerve. Common symptoms include blurring vision or blindness in one eye and pain when moving the eyeball (Chilinska, Ejma, Turno-Krecicka, Guranski, & Misiuk-Hojlo, 2016; Galetta & Balcer, 2013). A dark spot may appear in the center of the visual field. Most people who have a single episode of ON will recover their vision. Steroid medications have proven to be an effective treatment for most individuals. Approximately 15-20% of individuals who develop ON will later receive a diagnosis of MS and approximately 50% of individuals diagnosed with MS will also later develop ON (Balcer, 2006; Graham, & Kilstorner, 2015). In some cases, ON may be the first indication of MS (e.g., a clinically isolated syndrome).

The following imaging and recording methods are used to assess visual dysfunction along the afferent visual pathway in individuals with MS.

Optical Coherence Tomography as a visual measure of MS. Evidence from OCT studies point to axonal and other neuronal degeneration in the anterior visual pathway as important contributors to visual dysfunction in MS, even in individuals without a history of acute ON (Beh et al., 2016; Galetta & Balcer, 2013; Grecescu, 2014). Approximately 70% of individuals with MS experience thinning of the retinal nerve fiber layers (RNFL) (Beh et al., 2016). Thinning of the RNFL is related to retrograde degeneration, resulting from optic nerve damage (Galetta & Balcer, 2013). One study found a mean difference in RNFL thickness of 11.8 μm between MS participants (with and without a history of ON) and disease-free controls (Galetta & Balcer, 2013). Individuals with MS with a prior history of acute optic neuritis (AON) have even lower RNFL thickness than individuals with MS without a history of AON. On average, an AON attack results in a 20% decrease in RNFL thickness (Beh et al., 2016). In comparison, healthy eyes only lose 0.017% of RNFL thickness each year (Galetta & Balcer, 2013). Additional AON episodes result in more severe RNFL thinning. On average, individuals with MS lose approximately 2 μm of RNFL each year, regardless of AON history (healthy controls had a mean of 104.5 μm) (Beh et al., 2016; Galetta & Balcer, 2013).

High-resolution spectral-domain OCT (sdOCT) is a reliable and sensitive measure of the integrity of the retina (Graham & Kilstorner, 2015). Historically, OCT was used to evaluate the retinal ganglion cells (RGC), which form the optic nerve (Galetta & Balcer, 2013). RGC axons in the retina are unmyelinated until passing through the lamina cribosa and, therefore, structural thickness can be measured prior to demyelination, after which it can be compared as a measure of neurodegeneration and, possibly, repair (Galetta & Balcer, 2013). RGC's are also of particular interest because they represent typical neurons and

therefore are analogous to brain grey matter, while their axons are equivalent to brain white matter (Graham & Kilstorner, 2015). Technological advances improved OCT resolution so much that high-resolution spectral domain OCT now permits objective measurement of all nine layers of the retina (Beh et al., 2016). Additionally, recent research indicates that results from sdOCT have higher test-retest, inter-rater, and intra-rater reliability and reproducibility than results from temporal domain OCT (Galetta & Balcer, 2013).

Recent research suggests that thickness of the ganglion cell layer (GCL) and inner plexiform layer (IPL) are stronger correlates with functional and radiological measures of vision, compared to thickness of the RNFL (Beh et al., 2016; Galetta & Balcer, 2013). Additionally, research suggests that GCL and IPL atrophy may be a more accurate measurement of damage from attack of AON because, unlike RNFL, GCL and IPL thickness are unaffected by edema in the acute stages of AON (Galetta & Balcer, 2013). To summarize, OCT is multifaceted, it can define and document changes over time in the optic nerve head and in the retinal nerve fiber layers (RNFL) (Graham & Kilstorner, 2015). Research has found that thinning of the RNFL, particularly the GCL and IPL, are common in MS and are associated with reductions in visual function, retinal function, and self-reported quality of life. While OCT is essential in understanding the neural processes involved in MS, studies have shown that the VEP may be a more sensitive measure of visual dysfunction in this population (Chilinska et al., 2016; Naismith et al., 2009). Other studies support the critical utility of both measures and encourage the simultaneous use in the assessment and monitoring of individuals with MS (Hamurcu et al., 2017).

Electroretinography as a visual measure of MS. Retinal integrity may also be compromised in individuals with MS. In some cases, it can be difficult to discern if the

pathological disease process is impacting the retina or the optic nerve. In such cases, electroretinography (ERG) can be used for clarification. The conventional ERG is useful in evaluating the integrity of rods and cones, which are located in the outer retina. However, this limits assessment of the inner retina and optic nerve, where retinal ganglion cells (RGC) play a large role. Pattern electroretinograms (PERGs) provide promising measures for early detection in MS, as they can detect early losses of RGC's in optic neuropathies (Luo & Frishman, 2013; Porciatti & Ventura, 2004). The current study protocol included ERG and PERG stimuli (but the focus of this thesis is the tVEP-CR).

Visual Evoked Potentials as a visual measure of MS. The VEP has been shown to be of value in aiding in the diagnosis of multiple sclerosis (MS) and in monitoring disease progression. The VEP is a particularly useful measure in that it can identify underlying neural disruption, specifically demyelination and axonal damage, at different places and at different times; necessary requisites for an MS diagnosis (Balcer et al., 2015; Milo & Miller, 2014). The VEP measures the functional integrity of the retino-geniculo-cortical pathway. The optic nerve, a component of the visual pathways, is a primary site for demyelinating plaques and evidence supports the existence of compromised electrical responses along the afferent and, less commonly, efferent visual pathways, and within the primary visual cortex (Beh et al., 2016).

Clinical usefulness of VEPs in MS. There are several different ways that VEPs can be clinically useful in the field of MS. VEP studies have indicated their usefulness as a measure of disease status, disease progression, treatment efficacy, and as a potential biomarker of MS (Hardmeier et al., 2014; Hardmeier, Leocani, & Fuhr, 2017; Leocani et al., 2016). VEPs have been utilized to identify demyelination and axonal damage in the CNS,

indicative of an MS diagnosis, prior to the report of subjective symptoms (Balnyte et al., 2011; Janáky, Janossy, Horvath, Benedek, & Braunitzer, 2017; Klistorner et al., 2013). Thompson et al. (2018) recently published a revised version of the 2010 McDonald criteria for the diagnosis of MS. In their report, based on a meeting of the International Panel on Diagnosis of Multiple Sclerosis, they proposed updated criteria and recommended revisions. Thompson et al. (2018) addressed the inclusion of assessment of the anterior visual pathway in diagnostic assessment of MS, including the usefulness of VEPs. Their conclusion was that, while evidence exists to suggest that this method of analysis may be useful as an essential component of the diagnostic process, more research is needed in order to support its inclusion (Thompson et al., 2018).

VEPs have also been used as a measure of disease progression in MS (Leocani et al., 2016). Some research suggests that VEPs are most useful in evaluating disease status and progression in MS when used in conjunction with other evoked potential (EP) modalities (Hardmeier et al., 2014; Hardmeier et al., 2017). Treatment efficacy studies have also demonstrated the utility of VEPs to assess functional changes in the brain in MS, which can be attributed to medication (Meuth, Bittner, Seiler, Gobel, & Wiendl, 2011). Numerous studies support the use of VEPs and other EPs in conjunction with structural measures of analysis for the most comprehensive picture of disease status in MS.

Visual Acuity at Low Contrast in MS. Visual acuity abnormalities are common at low contrast in individuals with MS, most significantly in individuals with ON (Beh et al., 2016). Therefore, low-contrast acuity is a sensitive measure of visual function (Galetta & Balcer, 2013). Low-contrast vision likely reflects MS disease in the optic nerves and other anterior visual pathway structures (Galetta & Balcer, 2013). Low-contrast letter acuity scores

are well correlated with brain MRI lesion burden, VEP amplitudes, health-related quality-of-life, and RNFL axonal and neuronal loss, as measured by OCT (Beh et al., 2016; Galetta & Balcer, 2013). Low-contrast visual acuity scores are also strongly correlated with abilities such as driving, facial recognition, and other activities of daily living (Galetta & Balcer, 2013).

Research has been conducted using OCT methods to assess the impact of axonal and neuronal degeneration on visual acuity, specifically low-contrast (LC) letter acuity, in MS (Galetta & Balcer, 2013). Results consistently indicate that individuals with MS display significantly reduced LC acuity scores compared to healthy controls (Beh et al., 2016; Galetta & Balcer, 2013). The correlation between LC acuity scores and vision-specific quality of life connect visual impairment to loss of function in individuals with MS. Therefore, LC visual acuity scores might be important sources of information regarding physical abilities and limitations of individuals with MS (Galetta & Balcer, 2013). If LC is as sensitive of disease process in MS as previous research suggests, then the LC VEP measures used in the current protocol (but not included in this thesis) should be particularly useful in studying early-stage visual processing in MS, which other imaging techniques are not able to do (Beh et al., 2016; Galetta & Balcer, 2013).

Studies have also evaluated VEPS in comparison to OCT and MRI methods in order to see how they differ as well as complement one another (Chilinska et al., 2016; Galetta & Balcer, 2013; Naismith et al., 2009; Sakai, 2011). Literature reviews of visual structure and function in MS consistently report that there is a complementary relationship between many of the methods in terms of patterns of responses and correlations to neurologic dysfunction (Chilinska et al., 2016; Galetta & Balcer, 2013; Naismith et al., 2009; Sakai, 2011). Several

reviews also report that VEPs can be more sensitive than OCTs in detecting certain visual impairments (Chilinska et al., 2016; Naismith et al., 2009; Sakai, 2011).

The relationship between ON, MS, and VEPs. As mentioned previously, ON, an inflammation of the optic nerve, is very common in MS (Balcer et al., 2014; Chilinska et al., 2016; Galetta & Balcer, 2013). ON is often an initial indicator of MS, but visual deficits, such as damage to/or destruction of axons and dysfunction of neural mechanisms, also occur in individuals with MS who have no history of ON (Balcer et al., 2014). ON is typically unilateral in adults with MS but can also occur in both eyes. ON leads to reduced visual acuity, more significantly to low contrast stimuli and color vision than to high contrast stimuli. ON also results in relative afferent pupillary defect (RAPD) in the affected eye (or in the more severely affected eye of bilateral ON). Two-thirds of individuals with ON exhibit seemingly normal optic discs during the acute phase, as assessed by direct ophthalmoscopy (Balcer et al., 2014). VEPs and OCTs have therefore been used as alternative measures in assessing the presence of ON or, the dysfunction caused by its occurrence, in MS.

Recall that most people who have a single episode of ON will recover their vision and steroid medications have proven to be an effective treatment for most individuals. Despite return of visual acuity to normal, some individuals with MS with a history of ON will continue to display VEP abnormalities (Halliday, McDonald, & Mushin, 1973b). These findings provide support for the VEP as a sensitive marker of visual dysfunction in MS.

Recent research suggests that using low-contrast images, particularly low-contrast reversing-pattern stimuli, may be sensitive for detecting abnormalities in visual function in MS and in detecting occult optic neuropathy that may otherwise remain undetected (Beh et al., 2016; Kupersmith et al, 1984; Thurtell et al., 2009). Studies using pattern-reversal VEPs

have also revealed the unique usefulness of these stimuli in assessing ON damage in individuals diagnosed with MS (Chilinska et al., 2016; Janáky et al., 2017). One study revealed that increased response latency accurately identified demyelinating lesions in the optic nerve pathways 90% of the time, while other studies reported results of 81% and 83% accuracy (Chilinska et al., 2016; Naismith et al., 2009). These studies also consistently found that VEPs were more sensitive measures of visual dysfunction in MS than OCT (Chilinska et al., 2016; Naismith et al., 2009).

Janáky et al. (2017) studied tVEPs in MS by comparing individuals with ON to those without current or past history of ON. The authors conducted this study in order to show the additional utility of VEPs in MS, aside from evaluation of P100 latency. Results indicated that VEP abnormalities were present in more MS participants than expected and these abnormalities were not necessarily linked to a history of ON or worsening visual acuity. The authors concluded that VEPs can therefore be useful in the detection of visual dysfunction in MS in the absence of subjectively reported symptoms and without a history of ON (Janáky et al., 2017; Kilstorner et al., 2012).

VEPs to contrast-reversing checkerboard in MS. The bulk of existing VEP research on MS has been conducted using tVEPs to the conventional contrast-reversing checkerboard stimulus. Using this basic contrast-reversing checkerboard, we can make conclusions based on amplitudes and peak times in the waveform (Zemon et al., 1995). In particular, an increased P100 latency in the tVEP waveform has become a hallmark presentation of MS in the VEP response (Halliday, McDonald, & Mushin, 1973a; Duwaer & Spekreijse, 1978). However, the conventional VEP is limited in that it elicits responses from a host of neural mechanisms and thus lacks the resolution needed to identify specific information about select

neural pathways and mechanisms of dysfunction (Zemon et al., 1995). Additionally, VEP abnormalities caused by lesions may exist anywhere along the retino-geniculo-cortical pathway, thereby further limiting information regarding geographic specificity (Beh et al., 2016).

Research using tVEPs to evaluate visual function in MS have revealed several consistent findings related to the disease process using conventional checkerboard stimuli. Past research has demonstrated a consistent increase in response latency (e.g., P60 and P100) in MS, most commonly at the P100 marker (Chilinska et al., 2016; Galetta & Balcer, 2013; Halliday et al., 1973a, 1973b; Kilstorner et al., 2012; Sakai et al., 2011; Thurtell et al., 2009). The P100 latency measure is so commonly used that it has been referred to as a diagnostic sign of optic nerve demyelination in MS (Janáky et al., 2017). Recently, studies have begun to more closely examine additional peaks and troughs in the VEP waveform (P60, N75, N135), as interest grows in the utility of the VEP as a sensitive functional measure of underlying neural dysfunction in MS (Behbehani, Ahmed, Al-Hashel, Rousseff, & Alroughani, 2017; Hardmeier et al., 2014; Janáky et al, 2017).

Studies have also reported decreases in VEP amplitude in MS, which has been attributed to axonal loss (Balcer et al., 2015; Chilinska et al., 2016; Galetta & Balcer, 2013; Hamurcu, Orhan, Sarıcaoğlu, Mungan & Duru, 2017; Kilstorner et al., 2012; Thurtell et al., 2009; Sakai et al., 2011). Sakai et al. (2011) published a review on the relationship between vision, MRI, and VEP in MS. The authors report that P100 latency with normal amplitude has long been reported as characteristic of MS, however, axonal loss reduces the amplitude of the VEP as well. They also suggest that stimuli, particularly of low contrast, might be most sensitive in detecting visual dysfunction delay in MS (Sakai et al., 2011). In their

review of vision and MS, Balcer et al. (2015) noted that increased P100 latency with a normal looking waveform (e.g., normal amplitude) may be a result of conduction block caused by demyelination while reduced VEP amplitude may also result from conduction block due to either demyelination or damage to/loss of axons.

Novel statistical analyses developed in our laboratory will be used in the current study to further assess the underlying mechanisms within the tVEP CR responses in an MS sample (Zemon & Gordon, 2018; Zemon et al., 2009). The hope is to corroborate previous findings on MS and VEPs and to further identify underlying mechanisms and patterns of responses while simultaneously increasing confidence in the pattern of responses seen in individuals with MS, which are indicative of the dysfunction in those neural mechanisms.

VEPs to radial stimuli in MS. Zemon (1984) reported one of the first case studies using contrast-reversing checkerboards to elicit tVEPs and radial patterns to elicit ssVEPs in an individual with MS. The study was conducted in an attempt to separate out VEP contributions from different functional subsystems using a novel nonlinear systems analysis technique. Two conditions were utilized in this study: a superimposed condition (a spatial pattern contrast reversed at the sum of two sinusoids) and a lateral condition (windmill-dartboard stimulus) where parts of a pattern are modulated by one sinusoidal signal while other parts are modulated by a second, non-overlapping but contiguous, sinusoidal signal). Zemon (1984) studied a healthy control participant and an individual with MS under these conditions. Results from the healthy control indicated differences in amplitude and phase data between the two conditions. Notably, the amplitude peaked around 20 Hz in the superimposed condition and around 10 Hz in the lateral condition. There was a difference of approximately 200 degrees between phase responses between the two conditions at low to

intermediate frequencies, while phase data were almost identical at higher frequencies. Zemon (1984) interpreted these results by suggesting that local interactions (emphasized by superimposed condition) are excitatory, while lateral interactions are inhibitory). Therefore, these conditions provide a way of assessing integrity of local excitatory and long-range inhibitory interactions in the visual pathways of humans (Zemon, 1984).

The MS participant in this study reported subjective complaints of reduced contrast in his affected left eye and had visual acuity of 20/25 in that eye with corrected vision (Zemon, 1984). Results from the tVEP technique indicated a prominent P100 peak with normal range latencies in the unaffected right eye while the affected left eye produced increased P100 latencies (around 26 ms longer) than responses obtained from the fellow eye. Zemon (1984) notes that there is disagreement about the origins of the P100 latency in MS. As mentioned previously, some suggest that P100 latency reflects slowing conduction in the optic nerve caused by demyelination (Halliday et al., 1973b). Other investigators hypothesize that increased latencies may be a result of synaptic malfunction (Bodis-Wollner & Onofrj, 1982). Zemon (1984) notes that if a signal delay was the explanation for the P100 latency, then investigators would expect to see phase lags elicited by sinusoidal modulation, but this did not occur. Zemon (1984) suggests the possibility that amplitude loss at intermediate frequencies may produce increased P100 latency in the waveform.

Zemon (1984) emphasizes the difficulties with generalizing these results to the larger MS population due to small sample size and heterogeneity of the MS population. He reports that testing conducted on a second participant with MS revealed similar increases in P100 latency for the affected eye but no amplitude loss at any frequency, thereby supporting the author's argument regarding the sizable heterogeneity in disease process and presentation in

MS (Zemon, 1984). While these data are interesting and noteworthy and results support the use of the P100 peak time in assessing deficits in MS, the sample size was small and new techniques are now available to better quantify obtained tVEP results in an MS sample, which the current study intends to utilize.

Psychological & Physical Conditions Related to MS and Visual Function

Fatigue in MS. Fatigue has been identified as one of the most common symptoms of MS and it is recommended that an MS patient without reported symptoms of fatigue should be closely reevaluated for an MS diagnosis (Penner at al., 2009). Some reports suggest that 50-60% of individuals diagnosed with MS report fatigue as one of the most serious symptoms interfering with daily life activities and, therefore, significantly negatively impacting their quality of life (Giovannoni, 2006; Penner et al., 2009). Neural correlates have been identified which support the subjective experience of fatigue reported by individuals with MS, but more objective measures of fatigue are needed. Early animal studies have indicated that some demyelinated axons fatigue rapidly and are unable to conduct at high firing rates (Regan & Neima, 1984). Emerging evidence also suggests that MS-related fatigue is in some way related to inflammatory disease activity (Giovannoni, 2006). However, MS studies have shown weak correlations between fatigue and markers of systemic inflammation. Giovannoni (2006) cites evidence that MS-related fatigue is not correlated with Gd enhancing lesions on MRI, which is the most widely accepted marker of active inflammation in MS. He therefore points out that MS-related fatigue may be linked to peripheral rather than central inflammation and may be mediated by proinflammatory cytokines (Giovannoni, 2006). While prolonged and repeated testing of VEPs has logically

resulted in the study of a fatigability component, this alone does not explain the relationship between VEPs, fatigue, and MS.

Fatigue and VEPs in MS. Pokryszko-Dragan et al. (2015) conducted a study to assess the relationship between visual and auditory evoked potentials, fatigue, and MS using a group of MS participants ($n = 86$) and a group of healthy controls ($n = 40$). Fatigue measures used in this study were the FSS and the MFIS (Pokryszko-Dragan, Bilinska, Gruszka, Kusinska, & Podemski, 2015). VEPs were collected monocularly and stimuli consisted of a black and white checkerboard pattern with a check size of 36 cm^2 (viewing distance = 100 cm) and a pattern reversal of 1.9 Hz. Experimenters analyzed P100 latencies and amplitudes and then compared these between healthy controls and three subsets of the MS group (non-fatigued, moderately, and severely fatigued). An additional measure of relative P100 latency was computed based on interocular latency differences. Each eye was tested twice (Pokryszko-Dragan et al., 2015). Results indicated multiple asymmetrical abnormalities between the eyes of MS participants with fatigue (Pokryszko-Dragan et al., 2015). P100 latency increased and amplitude decreased in MS participants considered moderately and severely fatigued. Significantly longer P100 mean latencies and relative P100 latencies (interocular difference) were found in the MS group for both eyes as well as within each MS subgroup when compared to controls. Interestingly, mean P100 amplitude was significantly lower for the moderately and severely fatigued MS groups compared to controls but only for the left eye. For the right eye, P100 amplitudes were significant lower in the severely fatigued MS group compared to the MS group without fatigue. In the groups of MS participants considered moderately and severely fatigued, P100 latency for the right eye was significantly longer compared to MS participants without fatigue. The relative P100 latency

parameter, which was a measure of interocular latency difference, tended to correlate with the FSS, although this correlation was not significant. Given there was no significant correlation between relative P100 latency and MFIS results (which is a more detailed fatigue measure than the FSS), the authors state that there is an argument for further exploration of the relationship between VEPs and fatigue through the examination of physical and cognitive aspects (Pokryszko-Dragan et al., 2015). The physical and cognitive components of fatigue are assessed in the FSMC fatigue measure used in the current study.

Regan and Neima (1984) performed a VEP study based on the work of Halliday et al. (1973) to assess fatigability as a product of testing and to refine VEP specificity to measure fatigue in MS. Ten MS participants were evaluated, along with 10 individuals with glaucoma, 10 with Parkinson's disease, 10 with ocular hypertension and 10 healthy controls (Regan & Neima, 1984). The stimuli used were a conventional black and white checkerboard pattern and an altered version of the checkerboard that was contrast reversed, either by irregularly flickering light or by a moving pattern in order to stimulate fatigue. Study results indicated an increase in P100 amplitude in the MS and glaucoma groups but not in the Parkinson's group. All 10 MS participants evidenced delayed VEPs. Seven of the 10 MS patients showed abnormal attenuation when either the flicker or the moving pattern was added to the stimulus display. In total, nine MS patients displayed abnormal VEPs on one of the two fatigue tests. Notably, the one participant who showed no fatigue was the only MS participant without a history of clinical visual symptoms. Fatigue was evidenced in the MS group and the group with ocular hypertension but not in the Parkinson's or glaucoma groups. Regan and Neima proposed the possibility that these fatigue tests revealed the presence of neurons in the visual pathway that are functioning with a "reduced safety factor" and that the

functional integrity was reduced by partial demyelination in MS, whereas in glaucoma or Parkinson's, neurons were either unaffected or completely nonfunctional (Regan & Neima, 1984).

Pokryszko-Dragan et al. (2009) highlighted the results from the Regan & Neima (1984) study in order to discuss how they interpret their own results as originating from MS-related fatigue versus visual pathway dysfunction more generally. Pokryszko-Dragan et al. (2009) argue that VEP abnormalities observed in their study can be attributed to MS while the asymmetrical nature of the abnormalities may best be attributed to fatigue of the visual system more generally. According to Pokryszko-Dragan et al. (2009, p.241): "Asymmetrical damage to CNS pathways interferes with the perception and integration of stimuli of particular modality. To compensate for these dysfunctions, some additional areas of the brain may become activated. This corresponds with the concept of fatigue as a result of excessive load of CNS due to dysfunction of specific areas, as is supported by neuroimaging studies involving MS patients with fatigue."

The current study hopes to elucidate the current understanding of these relationships by including a well-validated measure of fatigue in order to assess correlations amongst VEP measures and fatigue in a group of MS participants and healthy controls.

Depression is commonly linked to fatigue in MS and it is highly recommended that measures of fatigue and depression be included in all MS evaluations (Herndon, 2016). Yalachkov et al. (2019) assessed the relationship between clinical variables, including physical and cognitive impairment, depression and fatigue, and patient-reported quality of life in individuals with RRMS (n = 39) and PMS (n = 16). Depression was measured using the BDI-II and fatigue was measured using the FSMC. Results indicated that depression,

disease duration, degree of education, disease type, and symptoms of psychological distress significantly impacted patient QOL, while fatigue and cognitive impairment demonstrated no significant effects. Given the limited sample size of their study and that fatigue has been found to impact QOL in other studies, these results should be interpreted with caution and further research is needed in this area.

Depression in MS. Given the wide range of impairments in MS that affect quality of life, it is evident that individuals with MS are at an increased risk for depression (Beck, Steer, & Brown, 1996a, 2000; Rao, Leo, Bernardin, & Unverzagt, 1991). Some reports suggest that depression is the most common psychiatric disorder in MS (Sacco et al., 2016). It is estimated that MS has a prevalence of 20% in MS and is 3-10 times more common in the MS population than in the general population (Sacco et al., 2016). The Beck Depression Inventory – Second Edition (BDI-II) is a useful tool in research and clinical practice to screen for depressive symptoms in individuals with MS (Benedict et al., 2006; Sacco et al., 2016). A recent study sought to validate measures of mood status in individuals with MS (Watson, Ford, Worthington & Lincoln, 2014). Results from this study support the utility of the BDI-II and revealed that the optimum cutoff score of 23 for the BDI-II yielded moderate sensitivity (85%) and specificity (76%) in the MS sample (Watson et al., 2014). Another study examined interrelations between depressive symptoms, disability, and disease course using a sample of 1,011 MS patients (Solaro et al., 2016). Using the BDI-II as the measure of depressive symptoms, the study found that 30% of the sample met criteria for clinically significant depression (Solaro et al., 2016). Results from this and other studies have also revealed correlations between the BDI-II with disability status and disease course in MS (Benedict et al., 2006; Sacco et al., 2016; Solaro et al., 2016).

Depression and VEPs in MS. Previous research provides evidence that the presence of depressive symptoms can affect VEP responses (Bubl, Kern, Ebert, Bach & Tebartz van Elst, 2010; Bubl et al., 2015; Normann, Schmitz, Fürmaier, Döing & Bach, 2007). Bubl et al. (2015) conducted a VEP study to assess whether depression has a cortical correlate, as previous research has shown that the PERG is affected in depressed participants. The study sample was comprised of 40 patients with a current depressive episode or diagnosis of *major depressive disorder* and 28 healthy controls. The Beck Depression Inventory (BDI) and the Hamilton Rating Scale for Depression (HRSD) were used as self-report measures to assess severity of depression. VEP stimulus was a black and white contrast-reversing checkerboard (0.51° check size, 12.5 reversals/sec, contrast of 3-80%). Amplitudes at all contrast levels were compared between the two groups and were correlated with severity of depression, as measured by the HRSD and BDI. Results indicated significantly reduced VEP amplitudes (reduced to 75%) in participants with major depression at all levels of contrast when compared to controls (p =.029). Additionally, VEP amplitude at all contrast levels correlated with the BDI and the HRSD. When assessed at individual contrast levels, correlations became weaker as contrast increased. Results suggest that depression affects the cortical response in major depression to a lesser degree than the retinal responses but still in a significant way. The study authors hypothesize that this may be a result of a moderation effect on losses in the retina, through modified contrast adaptation in the LGN or cortex (Bubl et al., 2015).

Normann, Schmitz, Fürmaier, Döing and Bach (2007) studied VEPs in a group of participants with major depression (n = 40) and healthy controls (n = 74) using black and white checkerboard patterns for stimuli. Study results showed that early VEP amplitudes

were decreased in participants with depression. Additionally, VEP amplitudes increased after chronic intake of an anti-depressant, which supports the usefulness of VEPs in tracking treatment progression for depression (Normann et al., 2007). The current study will employ the FSMC and the BDI-II, as subjective measures of fatigue and depression.

Rationale

VEPs are electrical potentials generated in the primary visual cortex and are a primary measure of visual function in the nervous system. This sensitive electrophysiological measure of cortical responses taps real-time brain dynamics in an objective, valid, and reliable manner (Zemon, 1984; Zemon, Pinkhasov & Gordon, 1993; Zemon & Gordon, 2018). VEPs are useful for evaluating visual abnormalities in a variety of disorders and have been shown to be a useful measure of disease state in individuals with MS. Visual impairment is a key symptom and often the first reported symptom of MS. This study will utilize objective methods of data analysis and uniquely chosen stimuli in an attempt to identify the visual system correlates of MS and how these neural processes are affected. The hope is that this research will lead to faster, more efficient methods for accurately diagnosing and monitoring the progression of the disease. Additionally, as new treatment options are introduced to target MS symptoms, VEPs may prove a useful measure of treatment efficacy.

Compromised axonal and neuronal integrity are important markers of disease status and process in MS. It is possible that functional measures may be more sensitive of MS disease status because functional changes can be taking place before gross structural changes. Previous studies have identified increased latency and decreased amplitudes in the waveform of conventional transient VEPs to contrast-reversing checkerboards, but the reason for this result is still unclear. Some researchers interpret it as delay in the system, but this is not

necessarily true. The proposed study will employ sophisticated analytic and frequency-domain techniques in order to parse out unique neural mechanisms involved in obtained responses and hopefully clarify misunderstandings.

Innovation

Traditional methods of neuroimaging in MS provide information on neural structure and related measures of neural activity. However, they are unable to capture cellular functioning on the level of milliseconds, which the VEP enables us to do. The current study hopes to add to existing knowledge of the structure-function relationships within the visual system in MS. Prior research has begun investigating the relationship between VEPs and MS but there is still more work to be done. VEP techniques will be used in this study that are not commonly used in the study of visual dysfunction in MS but have shown significant results when studied with an MS sample (Zemon, 1984).

Furthermore, the current study utilizes additional novel techniques, which were developed in our laboratory, to analyze entire responses objectively rather than relying only on the traditional method of subjective interpretations made by the examiner for tVEP data (Zemon & Gordon, 2018). These new analytic techniques that involve transformation of data from the time domain to the frequency domain and permit calculations based on frequency components enable us to obtain more accurate and precise information about what is contained within the response than ever before.

Repeat testing over extended periods have demonstrated the high reliability of VEP responses within an individual over time (Zemon & Gordon, 2018). However, responses vary greatly across individuals. Therefore, such comparisons can be difficult, especially in the time domain. Transforming data to the frequency domain and using magnitude squared

coherence normalizes the responses and allows for interindividual comparisons, which the current study intends to utilize.

A primary goal of the current study was to contribute to the existing literature regarding visual system impairment in multiple sclerosis through the measurement of early-stage visual processes via the VEP. Utilizing novel objective techniques to obtain patterns of data that are consistent with prior studies, will provide support for previous findings regarding VEP abnormalities in MS (e.g., reduced amplitudes and delayed latencies).

A secondary goal of this study is to explore associations between VEPs and measures of depression and fatigue in an MS sample. The intention is that the information gained from this study will enable greater understanding of how these factors may influence VEP presentation and/or how VEPs may be useful in tracking these symptoms over time in individuals with MS, potentially even to measure treatment response. There is hope that greater understanding of how VEPs are impacted by MS-related depression and fatigue will increase the clinical usefulness of VEPs in the MS population.

In summary, the current study hopes to fill knowledge gaps in the field of MS research in order to improve the accuracy and efficiency of early identification of the disease even when symptoms may not be present, processes used to monitor progression of the disease, as well as tracking the efficacy of new potential treatments. Previous studies have yielded significant results, which show promise of new methods of diagnosis, improved understanding of disease etiology, and treatment monitoring. It is expected, therefore, that the current study will elucidate further the nature of the disease process and the involvement of key neural functions, thereby supporting the utility of the study of VEPs to assess neural dysfunction in MS.

Aims and Hypotheses

Aim 1: Compare conventional time-domain and novel frequency-domain measures (i.e., the four MSC/power frequency bands , and phase delay estimates) of the tVEP-CR between groups.

Hypothesis 1A: We expect to find variability in peak-to-trough amplitudes between individuals in both groups. We hypothesize that MS participants will show weaker responses on measures related to early excitatory input to the cortex (e.g. P60-N75 amplitudes and MSC and power values in Band 3 (30-40 Hz). We expect MSC values to yield larger effects of group differences than amplitude and power measures, given that MSC reduces some of the inter-individual variability in the response that cannot be accounted for by power measures.

Empirical evidence from healthy control participants indicates that VEP amplitudes vary greatly between individuals and remain relatively consistent within an individual over time (Zemon & Gordon, 2018). Past research on MS has also found considerable variability in inter-participant amplitude (Halliday et al., 1973a; Sakai et al., 2011; Balcer et al., 2015). Thus, latency is considered a more reliable indicator of disease status in MS. Given the inter-participant variability in amplitude, delay measures (explored in Hypothesis 1B) are expected to result in more significant differences between groups than strength of response measures.

Despite its exclusion from the ISCEV protocol (Odom et al., 2016), P60 will be assessed in the current study as this early deflection in the waveform reflects afferent neural activity in V1. Given that the frequency domain measures extract additional information from the recording, it is predicted that the objective frequency domain measures will be more

sensitive than subjective time-domain measures at characterizing this activity (Zemon & Gordon, 2018).

Additionally, the current study utilized a short-duration (2-second) tVEP-CR condition as opposed to the conventional long-duration (1-minute) condition. We hypothesize that results from healthy controls on tVEP-CR measures will be consistent with results found by Zemon and Gordon (2018).

Hypothesis 1B: We hypothesize that MS participants will exhibit significantly longer peak times (latencies) in the time-domain and greater delay estimates in the frequency-domain compared to control participants. We anticipate that P100 peak time will result in the most significant difference between cohorts in terms of time-domain measures, consistent with previous research. Numerous studies have reported prolonged latencies in individuals with MS compared to healthy controls (e.g., Chilinska et al., 2016; Galetta & Balcer, 2013; Halliday et al., 1973a; Kilstorner et al., 2012; Sakai et al., 2011; Thurtell et al., 2009; Zemon, 1984). Accordingly, in the time-domain, P100 peak time is a critical measure of interest. However, this study will also assess latency of additional time points in the tVEP-CR waveform in order to assess whether these earlier or later peaks and troughs are also affected in MS. We hypothesize that additional peak times in the waveform (e.g. P60, N75, N135) will exhibit significant differences between groups.

We predict that frequency-domain delay measures (i.e., H24 delay to predict P100 peak time and Band 2 delay) will be strong predictors of their time-domain counterparts. Zemon and Gordon (2018) demonstrated that power in Frequency Band 2 is a strong predictor of the conventional N75-P100 amplitude measure in healthy controls, and we expect to replicate this finding in the current study. In addition, based on the work of Zemon

and Gordon (2018), we hypothesize that a time delay estimate based on the 24th harmonic (H24 delay) will be strongly predictive of P100 peak time. We expect to find larger effect sizes in the novel frequency domain estimates of time delay than in conventional peak times, which are measured subjectively in the time-domain waveform, because the frequency domain measures are more objective and reflect more of the informational content in the response (Zemon & Gordon, 2018).

Aim 2: Compare predictive power/classification accuracy of conventional time-domain and novel frequency-domain measures (i.e., the four MSC/power frequency bands , and phase delay estimates) of the tVEP-CR for group membership.

Hypothesis 2: We anticipate that the novel frequency-domain measures will yield larger predictive power than the conventional time-domain measures, and therefore, offer greater classification accuracy of group membership. The frequency-domain measures are based on more information in the signal (i.e., harmonic frequency components) than are time-domain measures, which rely on only a few single points in the time-series data, and therefore, are highly sensitive to noise and likely to exhibit greater variability (Zemon & Gordon, 2018). Additionally, the phase delay measure (i.e., H24 delay to predict P100 peak time) and the MSC frequency bands (i.e., MSC Bands 1 to 4) may serve as complementary measures that provide greater predictive power in a combined model for identifying individuals with MS. Thus, the objective frequency-domain measures may eventually be useful as a biomarker for MS.

Aim 3: Explore the relationships between VEP measures and measures of fatigue and depression in MS and in healthy controls.

Hypothesis 3: We hypothesize that MS participants will display positive relationships between depression and fatigue measures and that these correlations will be stronger than for controls. Additionally, given the inherent connection between the MS disease process and increased motoric disability, as well as the knowledge of cognitive decline in MS over time, we hypothesize that MS participants will report greater motor fatigue, based on the FSMC-motor subscale, and cognitive fatigue, based on the FSMC-cognitive subscale, than controls. Due to the heterogeneity of the MS disease process and existing evidence supporting the impact of fatigue and depression on VEPs, we anticipate that participants' psychological functioning will likely moderate differences in brain activity (Luo & Frishman, 2013; Porciatti & Ventura, 2004). We expect visual abnormalities (e.g., decreased amplitudes and increased latencies) to be correlated with measures of depression and fatigue in both groups. Specifically, we hypothesize that decreased amplitudes and increased latencies in the tVEP-CR responses will be negatively and positively correlated, respectively, with higher scores on the FSMC and the BDI-II.

Chapter II: Methods

Participants

A total of 40 adults were recruited for this study. Twenty healthy control participants and 20 participants with a diagnosis of relapse-remitting multiple sclerosis were recruited from Holy Name Medical Center and around the New York Tri-State area.

Data from four control participants were removed. One participant had a history of seizures, a second participant had a history of optic neuritis, and a third was removed due to medications that affect GABAergic inhibition. A fourth control participant was removed after waveform analysis indicated an abnormality that likely reflects a deficit in magnocellular function. Data from two MS participants were removed because of suspected glaucoma and a lack of any usable response measures. The final sample was comprised of 18 adults with MS and 16 healthy controls.

Table 1 summarizes the demographic information and clinical characteristics of the sample. A Mann-Whitney U test revealed no significant difference in age for control participants and MS participants ($U = 107$, $p = .20$, two-tailed test). A chi-square test for independence found that there was no significant difference in gender between groups ($\chi^2 = .06$, $p = .80$). There was no group difference in visual acuity at a viewing distance of 65 cm for either eye. Each group yielded a median value of 1.0 ($IQR = 0.2$). Acuity was measured by using the decimal form of acuity (e.g., 20/40 = 0.5) For 14 patients who provided information regarding length of illness, disease duration ranged from 1 to 17 years, with a mean of 6.9 years. All but two participants in the patient group ($n = 16$) reported taking one or more medications at the time of testing. MS participants were on a variety of medications including Cyclobenzaprine, Lemtrada, Natalizumab, Copaxone, Abagio Topiramate,

Pilocarpine, Modafinil, Phenytoin, Ampyra, Baclofen, Ocrevus, Rituxan, Lexapro, Adderall, Nexplanon, Synthroid, Myrebetriq, Tizanidine, Acyclovir, and Levothyroxine.

Recruitment and Eligibility

Adults between the ages of 20 and 40 were eligible to participate. All clinical participants met 2010 McDonald criteria for relapse-remitting MS. Exclusion criteria were the same for all participants and included diagnosis of an active seizure disorder, visual acuity less than 20/30, and visual diagnoses or medications that impact the VEP response (e.g., glaucoma, medications affecting GABAergic inhibition). Visual acuity was measured at time of testing from the viewing distance for each participant. Participants best-corrected visual acuity was used such that participants were asked to use eyeglasses or contacts to improve acuity if appropriate. Patients with ON were included and their status documented. Stratified sampling was used for sex in order to ensure comparable proportions of men and women in each group. This is necessary because women have a higher prevalence of MS than men. Participants with MS were recruited directly through the Multiple Sclerosis Center at Holy Name Medical Center. Recruitment flyers were distributed and posted at Holy Name Medical Center, Hunter College, and Yeshiva University in order to recruit controls. Control participants were recruited through personal contacts, emails, and word of mouth. Informed consent was obtained from all participants according to the Declaration of Helsinki. The Albert Einstein College of Medicine's Institutional Review Board and Holy Name Medical Center approved the current study.

Measures

VEP Equipment. All equipment and supplies were provided by Holy Name Medical Center. An EvokeDx system (Konan Medical USA) was used for stimulus presentation, data

collection, storage, and analysis. An infrared system built into the EvokeDx system tracked eye movements throughout testing in order to ensure that participants were attending to the stimuli. The researcher also monitored the gaze of each participant to ensure steady fixation. An organic light-emitting diode (OLED) display was used for stimulus presentations. The background luminance of the screen was approximately 50 cd/m^2. The frame rate was 60 Hz. An isolated differential amplifier, with a gain of 20K and a bandpass filter of 0.5-100 Hz, was used to protect against electrical shock. Visual acuities were collected monocularly from the viewing distance of stimulus presentation (65 cm) using the EvokeDx system. The stimulus field subtended a visual angle of 17° by 17° with a screen resolution of 1024 x 1024 pixels. The EEG signal was recorded synchronized to the display's frame rate, amplified, digitized at ten samples per frame, and stored in the computer.

Stimuli. The tVEP-CR responses were elicited by a conventional high contrast black and white checkerboard pattern (Figure 1). The stimulus was presented for approximately 3 seconds, with a 1 s adaptation period and 2 s of data collection for each of ten runs. The small check condition had a 64 x 64 checkerboard (check width = 16') while the large check condition utilized a 16 x 16 checkerboard (check width = 64'). Stimuli were presented with ~85% Michelson contrast and contrast-reversed with a square-wave signal at 1 Hz.

Demographic Questionnaire. A questionnaire was created for this study in order to collect demographic information, including gender, age, race, ethnicity, and years of education for each participant. Participants' medical and psychiatric history, including medical and psychiatric diagnoses (past or present) as well as any current medications, were also recorded. Participants were asked about current visual symptoms including

colorblindness, current or past ON, glasses/contacts, or other visual problems or hearing impairments.

Depression Measure. The Beck Depression Inventory (BDI-II) was administered as a brief screening measure of depression (Beck et al., 1996a). The BDI-II is a 21-item self-report questionnaire designed to assess the severity of a range of depressive symptoms over the past two weeks. There are four response options for each item ranging from "not present" (0) to "severe" (3). The BDI-II results in a summary score ranging from 0 to 63, which falls within one of six categories, defined by increasing depressive severity, ranging from "normal ups and downs" (0-10) to extreme depression (> 40). A score of 21 or above is indicative of clinically significant depression.

The BDI-II has strong psychometric properties, with coefficient alpha reliability estimates of .92 for outpatients with depression and .93 for a nonclinical sample (Beck, Steer, & Brown, 1996b). The BDI-II also has strong concurrent validity to other depressive measures. It exhibited a moderately high correlation with the Hamilton Psychiatric Rating Scale for Depression-Revised (r = .71) in a sample of psychiatric outpatients (Beck et al., 1996b). Previous research has indicated that the presence of depressive symptoms can result in reduced VEP amplitudes (Bubl et al., 2010; Bubl et al., 2015). Thus, we might expect reduced VEP amplitudes to be positively correlated with high depression scores on the BDI.

Fatigue Measure. Fatigue is an essential factor to record when testing individuals with MS over a sustained period (Balcer et al., 2015). Studies indicate that between 75%-95% of patients are affected by MS-related fatigue (Penner et al., 2009; Regan & Neima, 1984). The Fatigue Scale for Motor and Cognitive Functions (FSMC) is a 20-item self-report questionnaire that assesses fatigue in healthy controls and clinical populations and is used in

this study to evaluate MS-related cognitive and motor fatigue (Penner et al., 2009). The FSMC uses a 5-point Likert scale ranging from "does not apply at all" (1), "does not apply much" (2), "slightly applies" (3), "applies a lot" (4), and "applies completely" (5). The FSMC provides an overall fatigue score ranging from 0 to 100 as well as subscales of cognitive (range 0 – 50) and motor (range 0 – 50) fatigue. All three scores are associated with qualitative ranges indicating mild, moderate, or severe fatigue.

Penner et al. (2009) performed a validation study using the FSMC with a group of MS patients (n = 309) and healthy controls (n = 147). Results indicated high sensitivity and specificity of the FSMC in detecting fatigue in MS patients with high internal consistency (Cronbach's α >.91) and test-retest reliability (r > .80) (Penner et al., 2009). Subscale scores differentiated significantly between healthy controls and MS patients (p < .01). The FSMC also has strong convergent validity with other fatigue measures. High intercorrelations were obtained between overall scores and subscales of the FSMC, the Fatigue Severity Scale (FSS), and the Modified Fatigue Impact Scale (FMIS), however, logistic regression indicated the FSMC had higher levels of sensitivity and specificity when compared to the other fatigue measures (Penner et al., 2009). The FSMC was chosen for this study due to its strong psychometric properties for healthy controls and the MS population.

Procedures

Testing occurred at Holy Name Medical Center in Teaneck, New Jersey. Testing occurred over a single session and lasted approximately 1.5 hours. Participants began by reading and signing the consent form, completing the demographic questionnaire, and the depression and fatigue screening measures, followed by ERG and EEG testing. ERG testing took approximately twenty minutes while EEG testing took approximately one hour. Visual

stimuli were viewed monocularly in a dimly lit room. For VEP testing, participants wore an eye patch to cover the untested eye and were subsequently tested with their fellow eye. Half of the study participants had their right eye tested first and the other 50% had their left eye tested first. A three-electrode-lead, single-channel recording system was used to measure EEG activity. Two surface electrodes (*Oz,* active, and *Cz,* reference) were attached to the scalp using a water-soluble electrode paste, and a disposable ground electrode (*Pz*) was placed on the forehead, in accordance with the International 10-20 System (Jasper, 1958). Participants sat at a viewing distance of 65 cm from the display and were directed to fixate on the crosshairs in the center of the monitor. The VEP stimuli used in the current study were part of a larger battery of ten VEP test conditions and two ERG test conditions. Each condition was run until ten valid runs were collected per condition per eye.

Data Analysis

The current study was designed to investigate relationships among tVEP-CR measures and MS status using a short-duration stimulus and novel frequency-domain analysis techniques. Measures of depression and fatigue were explored as correlates of the electrophysiological measures and their interrelation as well. First, representative data for each cohort are illustrated, followed by univariate descriptive statistics for each measure of interest and bivariate descriptive statistics for relations between pairs of variables (scatterplots, correlation matrices). Next, inferential statistics (i.e., linear mixed-effects modeling [LMM], non-parametric Mann-Whitney tests) were applied to assess for significant fixed effects of group, stimulus, and eye condition.

Aim 1. To compare conventional time-domain and novel frequency-domain tVEP-CR measures between groups, tVEP-CR data were averaged over ten runs for each condition for

each participant. tVEP-CR measures of interest were extracted from the EEG data by first applying a discrete Fourier transform (DFT) to the time-series data and then filtering the transformed response in the frequency domain. SPSS v.25 and Microsoft Excel were used to further explore relationships among variables. Initially, by filtering out the odd- and retaining the even-harmonic frequency components from 2-100 Hz, as well as removing 60 Hz noise, we substantially reduced overall noise in the response, as the odd components contain no real response and 60 Hz is known to contain line noise. The signal strength for each frequency component was estimated by computing its amplitude in microvolts and its power in microvolts-squared per Hz. Relative response power synchronized to the stimulus presentation was measured by computing magnitude-squared coherence (MSC) for each harmonic component (e.g., Dobie & Wilson, 1989), and mean MSC values for each of four frequency bands identified through principal component analysis were also calculated (Zemon & Gordon, 2018). The four frequency bands were defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), and Band 4 (42-48 Hz) (Zemon & Gordon, 2018). Total power and its square root were also computed and analyzed for each band. Given that the prominent character of the tVEP-CR waveform is composed of the frequency components in these frequency bands (mechanisms), the time-domain waveform was reconstructed via an inverse discrete Fourier transform using only even harmonics from 6-48 Hz.

In the time domain, peak-to-trough amplitudes were computed and analyzed for P60-N75, N75-P100, and P100-N135. Latencies (peak times) of P60, N75, P100, and N135 were also analyzed for each eye and each condition.

For frequency-domain measures of latency, phase vs. harmonic frequency was plotted for each condition and eye for each participant. These plots included a regression line that was fitted for each band to calculate slope, which provides an estimate of time delay in the visual system in milliseconds. This delay estimate is calculated after obtaining vector-mean amplitude and phase data for all even harmonic components from 2 – 48 Hz. When the phase of a single harmonic component is greater than 180 degrees from that of the neighboring harmonic, an adjustment was applied (adding or subtracting 360°) to correct for a "wrap around" phenomenon and produce a smaller phase difference. The calculation for a delay estimate is as follows: (slope [in degrees/Hz]/360°) x 1000 = delay estimate in ms). Given that much of the informational content in the response exists in the N75 and P100 deflections, and that Band 2 largely characterizes this peak-to-trough deflection in the waveform, this band will be the frequency band of interest for this measure in the current study. The validity of Band 2 delay estimates relies on significant MSC values. The critical MSC value for each band depends on the number of harmonics in that band. Thus, for Band 2 (14 – 28 Hz), estimates were removed if the MSC values were below a critical value for significance (MSC < .145, $\alpha = .10$).

An additional frequency-domain, time delay estimate of P100 latency based on the 24th harmonic component was computed for each stimulus and eye condition for each participant (Zemon & Gordon, 2018). To use this equation, the millisecond value calculated from the obtained phase value of the 24th harmonic (H24) is divided by 360° and added to multiples of this harmonic's period (41.67 ms) to yield a value in the expected range of P100. This value is considered valid if the corresponding MSC value is above the critical value for a single harmonic component (i.e., MSC > .23, $\alpha = .10$). H24 estimates with corresponding

MSC values below the critical value for significance were removed from further analyses. Interestingly, the estimates that were removed came from participants who were excluded from all study analyses for one of the reasons mentioned above. Thus, the sample size used for 24th harmonic analyses remained intact (Control: $n_{small} = 32$, $n_{large} = 32$; MS: $n_{small} = 36$, $n_{large} = 36$).

Pearson bivariate correlation coefficients (two-tailed) were calculated to examine zeroth-order correlations among variables: for example, the relationship between conventional time-domain and novel frequency-domain measures of the tVEP. Scatterplots were used to detect bivariate outliers in the sample and to test for linearity of the data.

Given the small sample size and repeated measures within an individual, assumptions of normality and independence were unmet. Thus, the Mann-Whitney test statistic was calculated to determine differences between the MS and control groups for time-domain and frequency-domain amplitude and phase measures.

Linear Mixed-Effects Modeling. After the above analyses were completed, the database was restructured to perform linear mixed-effects modeling (LMM). LMM permitted exploration of the intraclass correlations (ICCs) within individual observers as well as interaction effects and fixed effects of group membership and the other factors within the study. ICC's computed for the *null model* in LMM were used to determine if data were correlated within an individual. Fixed effects were added to the models and to examine between-group differences in amplitude, power, MSC, and delay measures.

Aim 2. Scatterplots were reviewed to assess the relationships among strength measures (peak-to-trough amplitudes, MSC, and power) and latency measures split by group. Receiver operating characteristic (ROC) curve analyses were used to assess classification

accuracy by measuring the area under the ROC curve (AUC). Specifically, ROC curve analysis was used to test the hypothesis that novel frequency-domain measures of the tVEP have greater classification accuracy than do conventional time-domain measures. An additional measure referred to as A' was used to obtain a better estimate of AUC (Grier, 1971; Macmillan & Creelman, 2005; Pollack & Norman, 1964).

Aim 3. The final aim was focused on exploring relationships between clinical measures and VEP measures between groups. Total raw scores were collected for the BDI-II and the FSMC, as well as cognitive and motor subscale scores for the FSMC. Descriptive statistics were reviewed for all scores to assess normality and outliers. Given the small sample size, assumptions of normality were unmet. Thus, non-parametric Spearman's rho correlation coefficients and Mann Whitney tests were used to explore relationships between psychological measures and between psychological and VEP measures.

Power Analysis

Given that an intent of this study is to develop diagnostic measures and measures that can monitor disease progression, it is critical to obtain a large effect in whatever measure is of value. Prior studies using a range of sample sizes have found significant differences and large effects between MS and control groups using the characteristic P100 latency measure (Chirapapaisan et al., 2015; Duwaer & Spekreijse, 1978; Halliday, McDonald, & Mushlin, 1973b; Hamurcu et al., 2017; Thurtell et al., 2009;). Thurtell et al. (2009) used a sample size of 15 MS patients and 15 controls and reported a significant difference in P100 latency ($p <$ .01), with mean responses 20-30 ms greater for patients. Balynte et al. (2011) studied VEPs in 63 MS patients and 63 controls and found significantly increased P100 latencies in the MS group in both eyes, with mean latencies (ms) of 122.76 ± 14 and 122.60 ± 12.52, with effect

sizes (Cohen's *d*) of 1.46 and 1.54 ($p < .001$) for the right and left eyes, respectively. Chirapapaisan et al. (2015) conducted a longitudinal study of 35 individuals with "suspected" MS who had evidence of a clinically isolated lesion, an incidence of optic neuritis, or a clinical symptom indicative of possible MS disease development. After two years, 12 participants met criteria for clinically-definite MS, while 23 participants did not. Responses to tVEP-CR's were recorded for all participants, and P100 latencies (P100 ≥ 120 ms) were significantly prolonged in the MS group, with a mean latency (ms) of 119.08 ± 15.72 and an effect size of 1.36 (Chirapapaisan et al., 2015). The large effects observed in past research using this measure led us to expect that the current study will generate large effect sizes under monocular viewing conditions. To achieve a large effect size of $d = 1.0$ with $\alpha = .05$ and power = .80 using the Mann-Whitney test, a sample size of 14 per group is needed, which is surpassed in the current study which had 18 MS and 16 control participants (G*Power v3.1.9.2).

Part III: Results

Representative tVEP-CR Waveforms

Representative tVEP-CR data from a control participant are presented in Figure 2. It is notable that many participants in both cohorts displayed a significant and measurable response at P60, indicating that cortical activity can be measured at this early time point.

Control. Representative fellow-eye data from a control participant for the small check condition as displayed in the EvokeDx output are shown in Figure 2A. Both time- and frequency-domain measures are reported. The waveforms obtained for right (OD) and left (OS) eyes are consistent and each one has clearly defined peaks and troughs as labeled in the figure. P60, N75, P100, and N135 deflections are prominent with peak times within expected ranges (Zemon & Gordon, 2018; Odom et al., 2016). Frequency-domain measures are given on the right side of the output and they also yield consistent results for fellow-eye data. Individual MSC values at each harmonic frequency and mean MSC values for each frequency band are displayed to the right of the waveforms. In the current study, Bands 3 and 4 are defined slightly differently than those given here, but the first two bands are defined identically. Mean MSC values are greatest for Band 2. Critical values for .05 significance based on a normal approximation are shown as green horizontal lines. For purposes of the current study, measures for only the first four bands are analyzed. Given the limited extent of data collection during a run, many participants do not yield significant high-frequency responses, which are contained in Bands 5 and 6. Amplitude and phase plots for harmonic

frequency components are displayed below the MSC plots. Amplitudes decrease dramatically at frequencies above 30 Hz.

Figure 2B depicts data from the large check condition for the same participant. Response characteristics are similar to the small check condition; the waveform on the left exhibits prominent peaks and troughs, with slightly reduced amplitudes and earlier peak times compared to the small check condition. The signal strength is weaker for the large check conditions than small check conditions as only MSC Bands 1 – 3 reach significance.

Figure 2C depicts a phase versus harmonic frequency plot for the small check condition, left eye. This plot provides an estimate of time delay based on the square root of power in each frequency band in milliseconds (ms). The representative plot displays slopes and delay estimates for Bands 1 – 3. The different slopes for the separate bands indicate timing of different frequency mechanisms in the brain. All slopes have strong linear relations and, as expected, Band 1 (low frequency components) has the longest transmission time while Band 3 (high frequency components) has the shortest delay. However, Band 2 is the focus of the current study, which represents the predominant contribution to N75 and P100 deflections in waveform. Band 2 delay estimate for this participant is consistent and within the expected range for a healthy control.

MS. Figure 3A depicts representative fellow-eye data from an MS participant for the small check condition as displayed in the EvokeDx output. Similar to the control participant, the waveforms obtained for right (OD) and left (OS) eyes are consistent. The peaks and troughs are visible and defined, although N75 produced reduced amplitudes as compared to the control participant. This participant exhibits characteristically prolonged latencies: N75 peak times are approximately 20 ms longer and P100 peak times are approximately 30 ms

longer than controls. Frequency-domain measures are again given on the right side of the output and do not yield as consistent results for fellow-eye data as the control observer. Mean MSC values are greatest for Band 2. There is a clear deficit in response strength for this participant for bands other than 1 and 2, as reflected in the MSC plot. Critical values for .05 significance based on a normal approximation are shown as green horizontal lines. Amplitude and phase plots for harmonic frequency components are displayed below the MSC plots. Amplitudes again decrease dramatically at frequencies above 30 Hz.

Figure 3B depicts data from the large check condition for the same participant. Results are similar across conditions. Peak times are shorter for this condition and power in MSC bands is reduced, as expected. Overall, there is good agreement between eyes and mean MSC values are greatest for Band 2.

Figures 3C depicts a phase versus harmonic frequency plot for the small check condition for this participant. As expected, the slopes were steeper, and peaks occurred later for this participant. Similar to the control participant, Band 1 has the longest delay and Band 3 has the shortest, although Band 3 has very low amplitudes, so the delay estimate is not considered reliable. The Band 2 delay estimate is quite late, as expected.

MS Participant with Optic Neuritis. Figure 4 depicts fellow-eye data for the small check condition as displayed in the EvokeDx output for an MS participant with ON in the left eye. The effect of ON is visibly evident in the waveform. The right eye has clearly defined peaks and troughs and a prominent P100 peak while the left eye exhibits a much choppier waveform with less distinct peaks and troughs, and no prominent P100. Frequency-domain measures are again given on the right side of the output and, as expected, do not yield as consistent results for fellow-eyes. Mean MSC values are greatest for Band 1, although the

right eye reaches significance for Band 2. There is a clear deficit in response strength for all other bands. Amplitudes are reduced for the left eye compared to the right eye and decrease dramatically at frequencies above 30 Hz.

Control Participant with Medications Impacting GABA. Data from the small check condition for a control participant who is taking benzodiazepine medication is presented in Figure 5. Benzodiazepines enhance the effect of GABA in the cortex and the main positive deflection in the waveform (P100) is thought to be dependent on GABAergic inhibition (Zemon, Kaplan, & Ratliff, 1980). This is illustrated clearly in the waveform where there are distinct, early, and prominent P100 peaks.

Aim 1

The first aim in this study was to compare conventional time-domain and novel frequency-domain measures (i.e., power and MSC values for four distinct frequency bands and phase delay estimates) of the tVEP-CR between patients and controls. Short-duration tVEP-CR responses were collected under two stimulus conditions (checkerboards with small or large checks) and under monocular viewing conditions (right [OD] and left [OS] eye).

The number of participants whose tVEP-CR responses were recorded and deemed as valid along with cohort means and standard deviations for tVEP-CR measures are reported for both conditions and eyes in Tables 2 - 5. A total of 34 participants were included in these analyses (controls: $n = 16$, MS: $n = 18$).

Descriptive Analysis for Strength of Response Measures

To address Hypothesis 1A, strength of response measures (amplitude, power, MSC) were analyzed in the time- and frequency-domain. Histograms were reviewed to assess normality of the tVEP-CR data for all strength of response measures. Across amplitude

measures, substantial inter-individual variability in responses within groups, as well as overlap in amplitudes between groups was observed. This is expected given that past research studies have reported large inter-individual variability in amplitudes.

Amplitude. Figure 6 (A – D) depicts bar graphs of peak-to-trough amplitudes for both conditions (small and large check) and eyes. On visual inspection, there were no substantial differences observed between cohorts for N75 P100 or P100-N135 amplitudes, although the MS cohort did appear to display smaller amplitudes compared to controls. For the small check condition, for both eyes, the greatest difference between groups was for P60-N75 amplitudes. The large check condition, (Figure 6 C and D), yielded weaker responses overall and little between-group differences compared to the small check condition.

Power. Power was computed ($\mu V^2/Hz$) for each harmonic component and total power (power summed over harmonic components) was computed for each of the four frequency bands. Group differences were negligible, and signal power dropped off dramatically at higher frequencies (e.g., Band 4). Therefore, the square root of power in each frequency band was computed and used for all future analyses. Square root of power is useful because it enables an evaluation of more information at higher frequencies in a single plot and it is directly comparable to amplitude. To illustrate this difference, a clustered bar graph of power in the four frequency bands for the small check condition, right eye was created (Figure 7A) and there is a noticeable lack of power in Band 4. In comparison, Figure 7B displays a clustered bar graph using the square root of power for the same small check condition and there is a noticeable amplification of values for Bands 3 and 4. Square root of power graphs for small check conditions indicate that bands are relatively similar between groups, except for Band 3, which shows the largest difference between groups. This is to be

expected as Band 3 reflects excitatory input in the cortex and is consistent with the difference exhibited in P60-N75 amplitudes between groups, which also reflects excitatory activity in the cortex.

MSC. MSC was used as a measure of relative power, which adjusts for some of the inter-individual variability and non-neural factors in the response. Lower mean MSC values reflect weaker relative signal power in the response. Figure 8A displays a clustered bar graph of mean MSC values in the four frequency bands for the small check condition, right eye. MSC response patterns were matched for both eyes within and between groups in this condition. Similar to power graphs, MSC graphs for the small check condition depicts groups with similar strengths across bands with the exception of Band 3, which shows a sizable difference between groups and is much closer to noise level in the MS group. In the large check conditions, the difference between groups for Band 3 is negligible, as displayed in Figure 8B. Figure 8C depicts an aggregate bar graph of mean MSC values for the four frequency bands with both stimulus conditions and eyes collapsed. Based on this figure, the greatest difference between groups is observed for Band 3.

Scatterplots

Figure 9 displays a scatterplot of P60-N75 amplitude with square root of power in Band 3 for the small check condition, left eye. Analysis of scatterplots revealed a linear relationship between Band 3 and early excitatory input into the cortex (P60-N75) for both groups. In general, controls displayed greater increases in amplitude and power compared to controls. Interestingly, there was a subset of the MS cohort that displayed very low amplitudes with simultaneous increases in power. There were no extreme deviations from the linear relations seen in either group. Figure 10 depicts a scatterplot of N75-P100 amplitude

with square root of power in Band 2 in the small check condition, right eye. This scatterplot revealed strong linear relations for both groups and an even stronger relation for the MS group, thus supporting results from Zemon and Gordon's (2018) study.

Correlational Analysis

Correlation matrices for strength of response measures are displayed in Tables 6 - 9. Pearson correlation matrices were computed to assess how well the time-domain measures (amplitudes for P60-N75, N75-P100, P100-N135) and frequency-domain measures (mean MSC values and square root of power for Bands 1 to 4) relate to one another and to evaluate the strength of the relationships for each cohort with each measure.

Time-domain and frequency-domain strength of response measures by cohort. To further assess relationships among the time-domain and frequency-domain measures, correlational analysis was computed again but was split by cohort. As expected, there were differences in the values of correlation coefficients between groups.

Amplitude. There were correlations within the conventional time-domain peak-to-trough amplitude measures, as expected. In all conditions for controls and MS, there was a positive relationship between amplitudes of P60-N75 with N75-P100 and N75-P100 with P100-N135 with moderate to strong effect sizes. Interestingly, in the small check condition, the MS cohort had moderate positive correlations for P60-N75 amplitude with P100-N135 amplitude, which was not seen with controls.

Power. Positive correlations were exhibited for power measures in both groups. Overall, the MS group had more significant correlations than did the control group. In the small check condition, controls had moderate positive correlations for Band 2 with Band 3, with additional moderate correlations between bands that were not consistent in both eyes. In

the large check condition, both groups had one moderate positive correlation for Band 3 with Band 4 in both eyes.

Power vs. Amplitude. In the small check condition, in both eyes, the MS group exhibited moderate to high positive correlations for P60-N75 amplitude with power in Bands 1 – 3, the highest with Band 2. Correlations were stronger for MS patients than for controls overall. Both groups showed moderate to strong correlations for Bands 1 - 3 with N75-P100 amplitude. The highest correlation for this time-domain measure was with Band 2 in both groups, as expected and the MS group exhibited higher correlation than controls in both eyes on this measure. This higher correlation is likely related to greater dispersion in the MS group, due to the heterogeneity of the disease. Notably, the MS group again exhibited stronger correlations for all bands with P100-N135 amplitude than did controls.

The large check conditions exhibited weaker correlations and less significant results overall. For the MS group, Band 2 exhibited a strong positive correlation with N75-P100 (r = .793, $p < .01$) while the control group notably displayed no significance for Band 2, although the correlation was trending (r = .479, p = .06). For P100-N135 amplitude, both groups had strong to moderate positive correlations with Bands 1 and 2. The MS cohort exhibited slightly stronger correlations than the controls.

MSC. Within the frequency-domain measures, we found positive relationships among the four MSC frequency bands and these relationships differed by cohort. Overall, the MS group displayed more significant correlations that the control group. In the control group for the small check condition, Band 1 had no correlations with other bands. In the small check condition, right eye, the MS group had correlations with all three bands for each band, except for Band 1 with Band 3. Correlations within the MS group are likely enhanced due to disease

process. The large check condition produced few consistent correlations in both eyes in either group. In the control group, there were no significant correlations that were consistent in both eyes and positive correlations exhibited in the right eye were not strong. In the MS group, there were moderate positive correlations observed in both eyes.

MSC vs. Amplitude. There were significant positive correlations observed in both groups. For the MS group, in the small check condition, right eye, there were correlations among all measures, ranging from moderate to strong. Similar correlations were observed for the left eye but some did not reach significance. There was a strong correlation for N75-P100 amplitude with MSC Band 2. Correlations ranged from moderate to strong for P60-N75 amplitude with MSC Band 3. The control group exhibited less significant correlations in the small check condition. In both eyes, there were moderate to strong correlations for P60-N75 amplitude with Band 3, N75-P100 amplitude with Band 2, and P100-N135 amplitude with Band 1.

The large check condition again resulted in greater interocular variability and weaker correlations in both groups. There were some notable findings in this condition. The MS group had significant positive correlations in both eyes for N75-P100 amplitude with MSC Band 2. The control group had moderate to strong correlations in both eyes for P100-N135 amplitude with Band 1.

MSC vs. Power. Correlations were computed to assess relationships among frequency-domain measures in both groups. Overall, the MS cohort exhibited more positive relationships in both eyes than controls. For the small check condition, controls displayed more significant correlations in the left eye compared to the right eye. In the large check

condition, both groups had positive moderate to strong correlations between MSC and power in Band 1 and MSC and power in Band 2.

Non-parametric Analysis between Groups

Mann-Whitney tests were conducted to compare scores on strength of response measures between groups. Mann-Whitney test statistics, including *p*-values indicating the degree of significance for each measure are listed in Tables 2 - 5. There were several significant differences between groups. Notably, a Mann-Whitney *U* tests revealed significant differences between groups for MSC Band 3 in the small check condition in both eyes, with MSC in Band 3 greater for controls than for MS participants. There was also a significant difference in P60-N75 amplitude between groups in the small check condition, right eye and there was a trend in the left eye for the same condition.

Descriptive Analysis for Measures of Delay

To address Hypothesis 1B, latency measures were analyzed in the time- and frequency-domain. In all conditions, the MS group exhibited longer latencies than did the control group. Control participants displayed shorter peak times than typically observed under this condition when using the conventional long-duration (60-s) stimulus presentation. However, peak times for controls were consistent with Zemon and Gordon's (2018) results using the 2-s condition and a healthy control sample.

Latency. In the small check condition (OD) MS participants exhibited increased peak times that ranged from 10 ms for P60 to 15 ms for P100 compared to controls. The range was even greater in the left eye, with increased peak times ranging from 7 ms for P60 to 17 ms for P100. Similar results were obtained for the large check condition, with the MS group

displaying increased peak times (range: 5-12 ms) with slightly shorter peak times than the small check condition, as expected.

Band 2 Delay. As mentioned previously, the validity of Band 2 delay depends on the significance of MSC values (MSC < .145, $\alpha = .10$). For the small check condition, in the control group, two delay estimates were removed, while six estimates were removed for the large check condition. The MS group had six estimates that were below the critical value for the small check condition and 11 estimates in the large check condition. Of 64 possible estimates for the control group (32 small and large check estimates including the left and right eye), 56 were included in Band 2 delay analyses ($n_{small} = 30$, $n_{large} = 26$). Of 72 possible estimates for the MS group (36 each for small and large check estimates), 55 were included ($n_{small} = 30$, $n_{large} = 25$).

In the frequency domain, across conditions and measures, Band 2 delay produced the greatest differences between groups. In the small check condition, the MS group had a 17-18 ms longer Band 2 delay compared to controls for right and left eye data, respectively. In the large check condition, delays were 20 ms longer for MS than control participants.

H24 Delay. The H24 delay measure, which is predictive of P100 peak time, produced increased delays for MS participants that ranged from 11-14 ms longer than controls in all conditions. In the small check condition, P100 peak time was 2-5 ms longer than the predictive H24 delay measure. In the large check condition, P100 peak time and H24 delay were approximately the same (102-103 ms).

Aggregate bar graphs of mean peak times for each delay measure are presented in Figure 11 (A – C). Figures 11 A and B display aggregate graphs for small and large check conditions, respectively. Typically, we expect to obtain better results from small checks than

large checks and this is apparent for P100 and H24 delay measures. Although Band 2 delay shows a greater difference for large checks than for small checks, overall, small checks do produce better results. Finding significant differences between groups at P60 is notable here, as it provides evidence for demyelination occurring in cells prior to the cortex. The difference between groups increases in size as the signal travels further through the cortex, except at N135, where there is no significant difference. This is possibly due to greater inter-individual variability at later stages in the cortex. Figure 11C depicts all delay measures collapsed across conditions and eyes. Overall, the frequency-domain measures produced greater differences between groups than the time-domain measures, as expected. Band 2 delay appears to show the biggest difference between groups; however, it should be noted that the error bars for H24 delay are smaller than those for Band 2 delay, indicating that the magnitude of the difference between groups for these measures are similar. All latency measures exhibit significant differences in this graph except for N135.

Scatterplots

Scatterplots of the P100 peak time with H24 delay for the small and large check conditions and both eyes are depicted in Figure 12 (A-D). The relationship between P100 peak time and H24 delay is much stronger in the MS group in the small check condition, right eye, while controls display a slightly stronger relationship in the left eye, as evidenced by these plots. Scatterplots of the large check conditions replicate the finding that the MS group has a much stronger linear relation between these measures than the control group. This result is possibly related to the heterogeneity of the disease, which may cause greater dispersion of values, leading to stronger correlations. Alternatively, the control group has a

more restricted range of values, as expected, and therefore, has a weaker correlation in most conditions.

Correlational Analysis

Correlation matrices for latency measures for both eyes and conditions are displayed in Tables 10 – 13. Pearson correlation matrices were computed to assess relationships among time-domain measures (P60 latency, N75 latency, P100 latency, N135 latency) and frequency-domain measures (Band 2 delay and H24 delay) and to evaluate the strength of relationships.

Time-domain and frequency-domain latency measures by cohort. In the small check condition, the control group had strong positive correlations between P60 and N75, N75 and P100, and P100 and H24 delay, as expected. The strongest correlation for the right eye was for N75 and P100, while the left eye showed a stronger correlation for P100 and H24 delay. These results did not hold up for the large check condition, except for a strong positive correlation between P100 and H24 delay in one eye (OD). There were several correlations that were significant for one eye and one condition, but these results should be considered with caution, given there were no significant differences in visual acuity for participants. Band 2 delay had a positive moderate correlation with P100 in the small check condition for the right eye only. Interestingly, in the large check condition, Band 2 had a moderate negative correlation with N135 latency. It is notable that Band 2 delay was not significantly correlated with any other measures in the small check condition. However, this is likely due to the small sample size.

The MS group exhibited more significant correlations among measures than did controls. In all conditions, the MS group had consistent positive correlations: moderate

correlations were found between P60 and N75, strong correlations between N75 and P100 (strongest correlation for Small-OD), and moderate to strong correlations between N75 and N135 and N75 and H24 delay. P100 had strong positive correlations with N135 (strongest correlation for Small-OS), H24 delay, and Band 2 delay. Band 2 delay had a consistent and positive relationship with H24 delay in all conditions that ranged from moderate (small checks) to strong (large checks).

Band 2 Delay vs. Latency. As mentioned previously, Band 2 delay measure had a reduced sample size (controls: $n_{small} = 30$, $n_{large} = 26$; MS: $n_{small} = 30$, $n_{large} = 25$) because estimates were removed if MSC values were below the critical value for significance (MSC < .145, $p < .10$, $z = 1.282$). Despite a lack of significant correlations for the control group between N75 latency or P100 latency with Band 2 delay (except for those mentioned above) a review of peak times for latency measures in Table 1 reveals that the Band 2 delay estimate falls between N75 and P100 peak times across all groups and conditions, as expected. It is notable that the MS cohort exhibited strong positive correlations between N75 and Band 2 delay and between P100 and Band 2 delay in all conditions, with one exception for N75 and Band 2 delay in the large check condition (OD).

Harmonic 24 Delay vs. P100 peak time. Harmonic 24 delay (H24 delay) is an objective frequency-domain estimate of P100 peak time (Zemon & Gordon, 2018). Zemon and Gordon's extensive testing of this measure revealed that all estimates fell within the range of 80-110 ms as well as excellent agreement between the P100 peak time- and the 24^{th} harmonic frequency-domain measure. However, the H24 delay estimate is only considered significant if its corresponding MSC value is out of the noise (MSC > .23, $p = .10$). Several participants' delay estimates were removed for this reason. Interestingly, those estimates

were from participants who were removed entirely from the analyses for one reason or another and therefore, the sample size used for H24 delay analyses remained intact in both groups ($n_{control}$ = 32, n_{MS} = 36). Consistent with Zemon and Gordon's (2018) findings, control participants displayed almost identical P100 peak times and H24 delay estimates and strong positive correlations were observed for all conditions except for the large check condition in the left eye. The MS cohort also exhibited almost identical P100 peak times and H24 delay estimates, although they differed by 1–4 ms in the small check condition. The MS group displayed even stronger correlations between the two measures than the controls and correlations were strongest for large checks.

Non-parametric Analysis between Groups

Mann-Whitney *U* tests were conducted to test for significant differences on latency measures between groups. Results for the test for each measure are listed in Table 2. There were significant differences between groups. In terms of time-domain delay measures, as predicted, N75 and P100 latency were longer for the MS group than for controls in small and large check conditions ($p < .01$). Notably, P60 latency was longer in the MS group than the control group in the small check condition (OD) and the large check condition (OS). There was a significant difference in N135 latency in both groups for the small check condition (OS). In the frequency domain, for the small check condition, the MS group had longer delay estimates than controls on both measures. This significant finding held for the large check condition in the right eye but not in the fellow eye.

Linear Mixed-Effects Modeling

Linear mixed-effects modeling (LMM) was performed separately with amplitude, power, MSC, and latency as outcome variables. Power as referred to here is actually the

square root of power. Models were applied to the data using maximum-likelihood estimation with cohort, stimulus condition (check, small vs. large checks), eye (right vs. left), frequency band (1 – 4), amplitude measure (P60-N75, N75-P100, and P100-N135), and delay measure (P60, N75, P100, N135 peak times, Band 2 delay, H24 delay) as fixed-effect factors, including interaction terms. Models were built step-by-step and each began with the "null model." Additional variables and interaction terms were added individually.

Amplitude. First LMM was computed with amplitude as the outcome measure and cohort, check, eye, and amplitude measure as fixed effects. In Model 1 (null model), for intercept, participant was entered as both a fixed and random effect. Participant accounted for a significant proportion of the variance in amplitude in this model, with an intraclass correlation coefficient (ICC) of .40 ($p < .01$), indicating that data are highly correlated within individuals, as expected, and therefore, application of LMM is necessary. In Model 2, cohort was added as a fixed effect but did not achieve significance. In Model 3, Check size (small vs. large check) was included as a factor and was a significant predictor of amplitude (F (1, 371.15) = 49.63, $p < .01$) with greater amplitudes for small than large checks (t (371.15) = 7.05, $p < .01$). In Model 4, the interaction term of Cohort x Check was not significant. In Model 5, type of amplitude (amplitude measure) was added as a fixed effect and there was a significant effect of amplitude measure (F (1, 370.94) = 155.53, $p < .01$). In Model 6, Eye (right vs. left) was added as a fixed effect but it was not significant. In Model 7, interaction terms were added for Cohort x Amplitude Measure, Check x Amplitude Measure and Cohort x Check x Amplitude Measure. None of the interaction effects were significant, however, the three-way interaction of Cohort x Check x Amplitude Measure approached significance (F (2, 370.94) = 2.49, $p = .08$). In Model 8, the four-way interaction term of Cohort x Check x

Amplitude Measure x Eye was added, and it was also not significant. A final model included only the significant effects of check and amplitude measure. Notably, cohort was not a significant effect for amplitude.

MSC. LMM was then computed using MSC as the outcome measure with cohort, check, eye, and frequency band as fixed effects. In Model 1 (null model), for intercept, participant was again entered as both a fixed and random effect. Participant accounted for a significant proportion of the variance in MSC (ICC = .17, $p < .01$), indicating that data are highly correlated within individuals, as expected, and therefore, application of LMM is necessary. In Model 2, cohort was added as a fixed effect but did not achieve significance. In Model 3, check size (small vs. large check) was added as a fixed effect and was a significant predictor of MSC (F (1, 510) = 17.24, $p < .01$). Small checks produced larger MSC values than large checks (t (510) = 4.15, $p < .01$). In Model 4, the interaction of Cohort x Check was not significant, indicating that check size does not vary between cohorts. In Model 5, frequency band was added as a factor and there was a significant effect (F (3, 510) = 129.17, $p < .01$). In Model 6, interaction terms of Cohort x Frequency Band, Check x Frequency Band and Cohort x Check x Frequency Band were added as fixed effects. There was a significant interaction for Cohort x Frequency Band (F (3, 510) = 3.34, $p = .02$). Check x Frequency Band and Cohort x Check x Frequency Band did not achieve significance. In Model 7, Eye (right vs. left) was added as a fixed effect but it was not significant. In Model 8, two-way, three-way, and four-way interactions were added as fixed effects for Eye with all other variables, but none achieved significance. A final model included only the significant effects of check, frequency band and Cohort x Frequency Band. Notably, similar to LMM for amplitude, cohort was not a significant effect for MSC.

Power. LMM was next computed using the square root of power as the outcome measure with cohort, check, eye, and frequency band as fixed effects. All steps used above for MSC were repeated for power. In Model 1 (null model) for intercept, participant was entered as both a fixed and random effect (ICC = .09, $p < .01$). In Model 2, cohort was added as a factor but did not achieve significance. In Model 3, check was added as a fixed effect and achieved significance (F (1, 510) = 10.20, $p < .01$), indicating that small checks resulted in greater power than large checks (t (510) = 3.19, $p < .01$). In Model 4, the interaction of Cohort x Check was not significant. In Model 5, frequency band was added as a fixed effect and achieved significance (F (3, 510) = 416.51, $p < .01$), indicating frequency band (1-4) is a significant predictor of power. In Model 6, the interaction of Check x Frequency Band was significant (F (3, 510) = 6.76, $p < .01$). It is worth noting that the significance of this interaction was observed using the power measure but not using the relative power measure (MSC). Conversely, In Model 7, the interaction of Cohort x Frequency Band did not achieve significance for power, while this interaction was a significant effect for MSC. This is due to the fact that MSC adjusts for some of the within-individual variability and non-neural factors in the response and is therefore more sensitive to inter-individual variability than the power measures. In Model 8, Eye and interactions of Cohort x Eye, Frequency Band x Eye and Cohort x Frequency Band x Eye were added as fixed effects but did not achieve significance. In Model 9, the four-way interaction term of Cohort x Check x Frequency Band x Eye achieved significance (F (15, 510) = 1.83, $p = .03$). A final model included only the significant effects of cohort, check, frequency band, Check x Frequency Band and Cohort x Check x Frequency Band x Eye. It is notable that two and three-way interactions for cohort, frequency band, and eye were not significant until check was added to the interaction. While

these results are intriguing, caution should be exercised when generalizing to the population at large. It is also worth noting that cohort was not a significant effect for any strength of response measure.

Latency. Finally, LMM was computed with Latency as the outcome measure and cohort, check, eye, and delay measure as fixed effects. In Model 1 (null model), for intercept, participant was entered as both a fixed and random effect, Participant accounted for a significant proportion of the variance in amplitude in this model, indicating that data are highly correlated within individuals, as expected, and therefore, application of LMM is necessary (ICC = .08, Wald Z = 20.93, $p < .01$). In Model 2, cohort was added as a fixed effect and was significant (F (1, 33.17) = 13.97, $p < .01$). In Model 3, check was added as a factor and was significant (F (1, 754.32) = 5.53, $p = .02$). Small checks produced greater differences than large checks, but this effect was not as significant as the effect of cohort. In Model 4, an interaction of Cohort x Check did not achieve significance, indicating that check size does not vary between cohorts. In Model 5, delay measure was added and there was a significant effect of delay measure on latency (F (5, 753.19) = 1075.70, $p < .01$). In Model 6, interactions of Cohort x Check, Cohort x Delay, Check x Delay, and Cohort x Check x Delay were added as fixed effects. The interaction of Cohort x Delay (F (5, 753.17) = 4.89, $p < .01$) achieved significance, with Band 2 delay exhibiting the greatest difference overall, and P100 peak time exhibiting the greatest difference for time domain delay measures. No other interaction terms achieved significance. In Model 7, Eye was added as a fixed effect but did not achieve significance. A final model included only the significant effects of cohort, delay, check, and Cohort x Delay. Overall, latencies were longer for the MS cohort than controls and responses were shorter for large checks compared to small checks. Band 2 delay had a 20

ms difference between groups and time domain measures show the most significant difference for P100 peak time.

Aim II

Aim 2 focused on assessing the predictive power for group membership using time- vs. frequency-domain measures. Classification accuracy was assessed by measuring area under the receiver operating characteristic (ROC) curve (AUC), and through the use of A' calculations, and the results for small and large check conditions are given in Tables 14 and 15. Overall, latency measures (N75 and P100 peak time and frequency-domain measures) with small checks exhibited higher classification accuracy than did strength of response measures and measures with large checks. Additionally, greater predictive accuracy was obtained using novel frequency-domain latency measures (e.g., Band 2 and H24 delay) than conventional time-domain measures (Band 2 Delay AUC (OD) = .84, 95% CI [0.66, 1], $p <$ .01; H24 Delay AUC (OD) = .85, 95% CI [0.72, .98], $p < .01$). Figure 13 depicts ROC curves for power and MSC in the small check condition. Figure 14 and 15 depict ROC curves for time-domain latency measures for small and large checks and frequency-domain latency measures for small and large checks, respectively.

Assessment of interocular differences using scatterplots revealed similar responses for fellow eyes for all square root of power and MSC bands, as evidenced by Figure 16 A and B respectively. Additionally, analysis of AUC for latency measures revealed strong agreement between fellow eyes. Thus, the final ROC curves were comprised of aggregated fellow eye data for small checks and for large checks.

Notably, the MSC Band 3 scatterplot for the small check condition revealed that most of the MS group clustered below the MSC value of .2 for both eyes. Band 3 is related to early

excitatory input to the cortex, which is thought to be impacted in MS. Additionally, MSC measures are known to produce greater between group differences than are power measures. Thus, additional classification accuracy analyses were focused on this band. Scatterplots of latency measures by MSC Band 3 revealed a subgroup of MS participants who exhibited long latencies for N75 and P100 peak times and H24 delay with almost no Band 3 activity. Figure 17 (A and B) displays scatterplots of H24 delay with MSC Band 3 in the small check condition to illustrate this point. Given that increased P100 latency is a characteristic finding in MS and that H24 delay is predictive of P100 peak time, additional classification analyses were focused on the P100 latency and H24 delay measures. Composite variables were created for P100 latency as well as H24 delay with MSC Band 3 for the small check condition (OD).

Composite Variables. An H24 delay "long" variable was created for this condition with cutoff criterion of H24 delay values greater than or equal to 101 ms. A second variable was created to add an additional specification of MSC Band 3 less than or equal to .2. Individuals were categorized as to whether both dichotomous variables had a value of 1 or not. Using this composite frequency-domain measure to assess classification accuracy, specificity was 87.5% and sensitivity was 77.8%. Results of classification accuracy using the above specifications are depicted in Figure 18A. A P100 "long" variable was also created for the small check condition and participants were included if P100 latency was greater than or equal to 104 ms (based on ROC coordinates). Again, individuals were categorized as to whether the P100 peak time variable had a value of 1 or not, using this cutoff criterion. Using this P100 measure, specificity = 75% and sensitivity = 77.8%. Results of classification accuracy using this time-domain variable are depicted in Figure 18B. Based on these results,

the frequency-domain measure provided greater classification accuracy of group membership than the time-domain measure.

A' Estimation. Given that small check data produced greater classification accuracy than did large check data, and that P100 peak time and H24 delay are comparable time- and frequency-domain measures, we further evaluated the new small check measures (composite P100 "long" variable and composite H24 delay with MSC Band 3 variable) using A' calculations. For the composite H24 with MSC Band 3 variable, A' produced a classification accuracy of 89.6% with sensitivity of 77.8% and specificity of 87.5%. The P100 composite measure resulted in an A' value of 84.6% with sensitivity of 77.8% and specificity of 75%. Overall, results support our hypotheses that the frequency-domain measures produce greater classification accuracy than do time-domain measures.

Aim III

In order to address Aim 3, correlational analyses were conducted to analyze relationships among measures, and Mann-Whitney tests were used to test for differences between groups. Due to the small sample size used in the current study, the final aim is exploratory, and results should be evaluated with caution. Results presented here can be used to better understand the current sample and to inform future studies focused on exploring the relationship between VEPs, fatigue and depression. No adjustments were made for multiple correlations in this analysis, which should be considered when interpreting the following results. Preliminary examination of correlation matrices revealed moderate effect sizes, most of which did not reach significance due to the small sample size.

Analysis of frequency histograms revealed that the MS group had an approximately normal distribution while the distribution for control participants was skewed left, as many of

the control participants had scores clustered around zero for both measures. Given that the assumption of normality was not met, with different distributions in both groups and small sample sizes, non-parametric tests were used to evaluate relationships among variables. Analyses indicated a strong positive relationship between measures of depression and fatigue in the MS cohort, as expected.

Descriptive Statistics

Descriptive statistics for clinical measures are presented in Table 1. Sixteen control participants and 17 MS participants were included in BDI analyses. The FSMC analyses were comprised of 16 control participants and 15 MS participants. BDI scores ranged from 0 to 13 for controls and from 0 to 35 for MS participants. The control group had scores that were clustered close to zero, except for two control participants who had scores of 13 and the next highest score was 5. The MS group had scores that were more dispersed. FSMC total scores ranged from 20 to 76 for controls and from 20 to 80 for patients. Notably, one control participant was an outlier with a total score of 76 while the next highest score was 56. Again, the MS group's FSMC scores were more dispersed than controls.

Non-parametric Tests

Mann-Whitney tests were conducted to compare mean scores between groups. MS participants had significantly higher mean BDI scores than control participants ($U = 62.5$, $z = -2.67$, $p < .01$). MS participants also had significantly higher FSMC total scores ($U = 62$, $z = -2.30$, $p = .02$) and FSMC motor subscale scores ($U = 57.5$, $z = -2.48$, $p = .01$). The difference in means for the FSMC cognitive subscale scores was trending but did not reach significance.

Correlational Analysis

Correlation matrices were computed for depression and fatigue measures with time- and frequency-domain tVEP-CR measures. Analyses revealed significant Spearman's rho correlations between all psychological measures, as predicted. In both cohorts, BDI total score had positive correlations with the FSMC total score (controls: $r_s = .788$, $p < .01$, MS: $r_s = .655$, $p < .01$) and the FSMC subscale scores (FSMC-cog controls: $r_s = .755$, $p < .01$; FSMC-cog MS: $r_s = .702$, $p < .01$; FSMC-mot controls: $r_s = .716$, $p < .01$; FSMC-mot MS: $r_s = .524$, $p = .05$). As expected, the three FSMC scores had strong positive correlations with one another in both cohorts.

Correlations were also explored between psychological measures and time- and frequency-domain VEP measures of amplitude and latency. In terms of strength of response measures, in the small check condition (OS), controls had a negative correlation with square root of power in Band 1 and FSMC motor subscale score ($r_s = -.545$, $p = .03$) and with MSC in Band 1 and the FSMC motor subscale score ($r_s = -.556$, $p = .03$). For the MS group, the FSMC motor subscale score was positively correlated with P100 to N135 amplitude ($r_s = .515$, $p = .05$). In the small check condition in the fellow eye (OD), the control group had a positive correlation with MSC in Band 2 and BDI ($r_s = .548$, $p = .03$). Given the lack of consistency between eyes and the small sample size used in the current analyses, results must be considered carefully and interpreted with caution.

Moderate correlations were observed between psychological and latency measures, but most did not reach significance, perhaps due to the small sample size. The control group displayed several significant correlations. In the small check condition (OS), the control group had positive correlations between N75 latency and all FSMC scores (FSMC Tot: $r_s = .608$, $p = .01$; FSMC Cog: $r_s = .583$, $p = .02$; FSMC Mot: $r_s = .646$, $p < .01$). In the large

check condition (OD), the control group had a positive correlation between Band 2 delay and BDI (r = .605, p = .03). In the small check condition (OS), the MS group had a significant negative correlation between Band 2 delay and BDI (r_s = -.514, p = .04). Scatterplots of this condition revealed two distinct MS subgroups: one with short delays and high BDI scores and a second with much longer delays and low BDI scores. Again, significant correlations were not consistent in both eyes for either group; therefore, they are not considered reliable.

Notably, one control participant was identified as an outlier in these analyses, and this is likely due to benzodiazepine medication affecting GABAergic inhibition, as discussed earlier. Unlike another participant with likely strong GABAergic inhibition, who was excluded from the current study, this participant demonstrated strong responses out of the noise and was retained in the current exploratory analyses.

Part IV: Discussion

Interpretation

The current study was designed to explore visual system functioning in MS through electrophysiological assessment. Time- and frequency-domain measures were included and compared between groups. Clinical measures of depression and fatigue were used to explore relationships among these measures and with the tVEP-CR responses.

Based on results from the time-domain and frequency-domain measures, there is strong indication that excitatory input to the cortex is impacted in MS, and this is evidenced most strongly by MSC in Band 3, followed by P60-N75 amplitude. Frequency Band 3 is believed to reflect early excitatory input into cortex and Frequency Band 2 reflects more of the major waveform complex of the VEP (including P100), which is believed to play an inhibitory role on cortical neurons (Zemon & Gordon, 2018). Results from the current study support previous findings of prolonged latencies in MS and provide evidence for the utility of novel frequency-domain measures in detecting larger differences between MS cases and healthy individuals.

Significant P60 results were obtained in amplitude and phase analyses, indicating that there is measurable activity at the initial input to the cortex in the time domain. Despite P60's lack of inclusion in the ISCEV protocol (Odom et al., 2016), it is clearly a valid time point worth examining. The significant difference in this measure between MS cases and controls provides evidence that demyelination is affecting precortical activity.

Analysis of all response measures supported previous findings that, overall, small checks perform better than large checks in this application. Small checks elicited stronger VEP responses and resulted in more significant statistical findings. Although small checks performed better, responses to large checks also produced differences between groups, albeit smaller differences.

Notably, for the small check condition, the largest difference between groups was for P60-N75 amplitudes, while the large check condition yielded no significant difference for this measure. This finding is interesting due to the fact that some investigators question whether P60 even exists, although this is possibly a reflection of over filtering or noisy recordings, rather than a true absence of a P60 peak. MSC graphs for the small check conditions exhibited relatively similar values for the two groups across bands, with the exception of Band 3, which showed significant differences between groups and was much closer to noise level in the MS group. This pattern of results using relative power indicates strong selectivity for only one mechanism. Results suggest there is an impact on early input into the cortex (earliest amplitude measure in time series data is P60-N75). All of these results taken together support the findings of delayed peak times in MS and indicate that the MS group displays an early excitatory input deficit, which is possibly associated with damage and loss of axons in the optic nerve. In the large check conditions, the difference in Band 3 between cohorts disappeared, indicating that only small checks show the effect on this mechanism.

Aim 1. Comparing magnitude measures between groups in time- and frequency-domain.

Mann-Whitney tests revealed significant differences between groups in P60-N75 amplitude and MSC/Power Band 3, as hypothesized. This difference was more significant using the frequency-domain measures versus the time-domain measure, also as hypothesized, as it was consistent in both eyes in the small check condition for MSC and power measures. In contrast, the significant difference between P60-N75 amplitude was only significant in one eye in the small check condition.

Results from review of frequency histograms for normality and bar graphs of peak-to-trough amplitudes support the hypotheses. Significant variability in amplitudes and considerable overlap between cohorts was observed and, therefore, there was less distinction between groups with these time-domain measures. For the small check condition, in both eyes, the greatest difference between groups was for P60-N75 amplitudes, although the large check condition showed less difference for this measure. This observation, as well as Mann-Whitney tests for differences between groups, also supported the hypothesis that cortical activity is occurring at this early time. Bar graphs of power and square root of power showed that square root of power enabled visual comparisons across frequency bands within a single plot. Thus, square root of power was used for all future analyses throughout the current study.

Within the frequency domain, clustered bar graphs of power and MSC supported the hypothesis that the MSC measures result in greater between-group differences compared to power measures (power and square root of power). This is because MSC is a relative power measure that adjusts for individual differences and reduces some of the noise contained in the response. MSC and square root of power graphs revealed substantial differences between groups for Band 3, and the difference was amplified when compared to time-domain

amplitude graphs. This supports the hypothesis that greater differences between groups are expected with frequency-domain versus time-domain measures.

Additionally, scatterplots of N75-P100 amplitude with square root of power in Band 2 depicted strong linear relations for both groups with the relationship strongest for the MS group. This finding was supported by high correlations for this time-domain measure with square root of power in Band 2 in both groups. These results provide support for those from Zemon and Gordon (2018) using the same short-duration (2-s) stimulus condition.

Correlational analyses revealed additional information about relationships among these magnitude measures. Interestingly, in terms of amplitude, the MS cohort displayed strong correlations for both middle and higher frequency bands with P60-N75 amplitude across conditions and eyes, which was not seen with controls. It is likely that the disease process is contributing to this added correlation among components that otherwise would not be expected to correlate much. For the MS group, correlations ranged from moderate to strong for P60-N75 amplitude with MSC Band 3 in fellow eyes, which is believed to represent excitatory input to the cortex, and which is known to be impacted in MS. In terms of MSC, in the small check condition, right eye, the MS group had correlations among all bands with each other. Correlations within the MS group are again likely due to disease process. Overall, responses to large checks exhibited few significant correlations in either cohort. However, frequency-domain measures as defined in the current study were based on PCA analysis using small checks (Zemon & Gordon, 2018). Thus, it is possible that PCA using large checks might result in different frequency components and bands of interest.

Excitatory input into the cortex has been associated with early time-domain (P60-N75) and higher frequency-domain components (e.g., Band 3) (Zemon & Gordon, 2018). In

MS, the disease process is thought to disrupt excitatory information into the cortex via white matter projections and optic radiations traveling to the cortex from the LGN (Qureshi et al., 2014; Milo & Miller, 2014). Given this, it was hypothesized that the MS group would display weaker responses in early-time domain (P60-N75 amplitude) and higher frequency-domain components (MSC/Power in Band 3) compared to controls. This hypothesis was supported by bar graphs, scatterplots, correlational analyses, and Mann-Whitney tests.

Aim 1. Comparing time- and frequency-domain latency measures between groups.

Overall, the MS group exhibited significantly longer latencies compared to controls. Mann-Whitney tests revealed several consistent significant differences between groups in the time- and frequency-domain. In terms of time-domain delay measures, as predicted, N75 and P100 latency were significantly longer for the MS group than for controls across conditions and eyes. P60 latency was significantly longer in the MS group than the control group in the small check condition, right eye and the large check condition, left eye. There was a significant difference in N135 latency in both groups for the small check condition, left eye. In terms of frequency domain measures, the MS group had longer Band 2 and H24 delay estimates than controls on both measures in the small check condition and for one eye in the large check condition.

Analysis of time- and frequency-domain measures of latency supported the hypothesis that frequency-domain measures yield larger differences between groups than time-domain measures. Significant differences between groups were found in time-domain measures other than P100, also supporting hypotheses. In addition, in the current study, P60 peak time was measurable and significant differences were found between groups using this

measure, supporting the hypothesis that activity is taking place at this early time in the cortex.

In the time domain, the difference between groups grew as time after stimulus onset increased and the signal traveled further into the cortex until it reached the latest time point measured (N135), where there was no significant difference between groups. This supports the hypothesis that, in the time domain, P100 peak time would exhibit the greatest difference between groups. Descriptive statistics show that there is a difference in N135 peak time, however, the difference is not significant. It is likely that inter-participant variability at later time points in the cortex is contributing to this finding.

In the frequency-domain, Band 2 delay produced the greatest differences between groups. In the small check condition, the MS group had delays that were 17-20 ms longer than controls in all conditions. In addition, as hypothesized, the 24th harmonic delay estimate (H24 delay) was strongly predictive of P100 peak time across groups. In all conditions, differences between these two measures ranged from 1-5 ms. Scatterplots revealed that the MS group had a much stronger linear relationship between P100 peak time and H24 delay. This finding is likely due to the heterogeneity of the disease, which causes greater dispersion of values, thereby leading to stronger correlations.

Correlations among latency measures were examined for each cohort. It is notable that Band 2 delay was not significantly correlated with any other measure in the small check condition. However, this is likely due to the small sample size used, as participants were removed from Band 2 delay analyses if their MSC values were considered too low. Numerous correlations among measures for the MS group suggests that the visual system in healthy controls is different than in individuals with MS.

Aim 1. Linear mixed-effects modeling

After assessing that data were correlated within an individual (ICC > .05), LMM was performed with amplitude, power, MSC, and latency measures as an outcome variable in order to assess which variable/factor or interaction of factors (check size, cohort, eye, frequency band, amplitude measure, delay measure) is most effective in detecting differences between cohorts. Results from LMM support the hypothesis that there is greater sensitivity to detecting differences using novel frequency-domain measures compared to conventional time-domain measures.

Most of the variance in amplitude and MSC can be accounted for by check condition (small vs. large) and peak-to-trough amplitude measure (P60-N75, N75-P100, P100-N135) or frequency band (1 – 4), respectively. Given the large inter-individual variability in amplitude, and the fact that MSC is related to amplitude, the lack of significance for most fixed effects is not surprising. Additionally, the significance of check condition is expected, given the evidence that small checks are more sensitive to detecting differences than large checks. Significance of the amplitude measure and frequency band is also expected, given the variation in amplitude that exists at different points in the cortex and the different frequency components contained in each band, as a result of inhibitory and excitatory activity. There was a trend for the three-way interaction of Cohort x Check x Amplitude measure. For MSC, there was an interaction of Cohort x Frequency Band (differences in Bands 2 and 3). It is notable that cohort was not significant for amplitude but did achieve significance as an interaction for MSC and was a significant factor for square root of power as the outcome measure. In addition, eye was significant as part of a four-way interaction for power but was not significant on its own or as a three-way interaction with cohort, check, or

frequency band. Clustered bar graphs computed with 95% C.I.'s show Band 3 is clearly significantly different between groups for MSC.

In terms of LMM for latency, longer latencies were observed in the MS cohort and shorter latencies were observed overall for large checks than small checks, as expected (Moskowitz and Sokol, 1983). In the final model for latency, cohort, check, and delay measure were main effects and there was a significant interaction of Cohort x Delay. Review of aggregate graphs and mean peak times revealed that Band 2 delay had a 20 ms difference between groups and, of the time-domain measures, P100 peak time showed the most significant difference between groups. Aggregate graphs calculated for all six peak time measures (frequency and time domain) revealed a difference of ~12 ms, with no overlap in 95% C.I.

Overall, results showed that check size, amplitude measure/frequency band accounted for a significant proportion of the variance in strength of response outcome variables. Cohort, check size, and delay measure accounted for a significant proportion of the variance in latency. Small checks produced larger responses than did large checks, as expected. Patients exhibited longer latencies compared to controls. Frequency Bands 2 and 3 produced the greatest differences between groups, with Band 3 showing the greatest significance. Eye (OD vs. OS) did not account for a significant proportion of the variance in any model and was only significant as a 4-way interaction in LMM for latency.

Aim 2: Classification accuracy.

Aim 2 focused on evaluating tVEP-CR measures in order to assess classification accuracy for group membership, including whether an individual measure or a combination

of measures was most predictive. To address Aim 2, classification accuracy was evaluated with ROC curves for small and large check conditions.

Notably, despite several participants with diagnosed ON, interocular differences were not substantial enough to warrant additional analysis. This may be due, in part, to the small sample size, and future studies with larger samples may find greater significance when comparing eyes. Thus, these results should be interpreted with that in mind.

Based on the work of Zemon and Gordon (2018), the waveforms have been simplified and made more cohesive, frequency bands of interest have been identified and only those harmonics were used to define the time-domain waveform. Thus, much of the noise and variability that might be seen in clinic waveforms were removed from time-domain analyses, thereby strengthening these measures. Overall, latency measures in the time- and frequency-domain (N75 and P100 peak time, Band 2 delay and H24 delay) and small checks exhibited higher classification accuracy than strength of response measures and large checks. Additionally, greater predictive accuracy was obtained using novel frequency-domain latency measures (e.g., Band 2 and H24 delay) than conventional time-domain measures. Thus, while time-domain measures did provide high predictive power, the frequency-domain measures still outperformed the time-domain measures in predicting group membership, as hypothesized.

Results indicated that small checks were more predictive than large checks, so additional analyses were focused on small checks, which is also the most commonly used stimulus presentation in clinical settings. Small checks, in general, produced larger responses than did large checks, and this is likely due to matching better the receptive field center mechanisms of retinal ganglion cells (Dacey and Petersen, 1992). We then chose the best

candidate for predictive power, P100 peak time, which is also the variable that is most often analyzed in clinics, for the A' calculations. A' calculations were computed to compare P100 peak time for the small check condition to novel measures for the same condition, including H24 delay and MSC in Band 3. Results supported the hypotheses, as frequency-domain measures had higher predictive power than time-domain measures, with almost 90% classification accuracy based on the A' measure.

Analyses revealed a subset of the MS group that had almost no MSC in Band 3 and long H24 delays. Despite close inspection of both the MS subgroups, it remains unclear why this is occurring. There is no simple non-neural explanation for this finding. It was observed that there was more racial diversity in this subgroup. Band 3 is related to early excitatory input to the cortex, which is thought to be impacted in MS and, as mentioned above, MSC is known to produce greater between-group differences than power measures. In addition, H24 delay is a frequency-domain measure that is predictive of P100 peak time and P100 latency is a characteristic finding in MS. Thus, this finding supports several hypotheses including: H24 delay as a frequency-domain measure of P100 peak time, P100 peak time delay as a distinguishing finding in an MS participant's VEP response, and the impact of the MS disease process on early excitatory activity in the cortex.

It is possible that there is another explanation for this finding (e.g., disease duration, number and/or severity of relapses) but the current study was limited in its ability to explore this further. It is important to note that, due to the small sample size, these results cannot be considered generalizable to the greater MS population at this time. This finding does warrant additional analyses to explore if there is some variable or combination of variables that distinguishes between these groups.

Aim 3: Clinical measures of depression and fatigue.

Aim 3 was explored using Mann-Whitney tests to assess differences in psychological measures between cohorts. Additionally, correlations were computed to compare psychological measures with each other and with tVEP-CR measures of amplitude and latency.

Participants in the MS group had scores that were more dispersed across measures than did controls. Mann-Whitney tests revealed that self-reported depressive symptoms were higher for participants in the MS group than those in the control group, which supported the hypothesis. MS participants also had significantly higher self-reported fatigue than controls, based on their FSMC total scores, also supporting the hypothesis. Regarding cognitive and physical (motor) fatigue subscores, MS participants reported greater physical fatigue than the control group, while differences in self-reported cognitive fatigue did not reach significance between groups. The significant difference between groups in motor fatigue supports the current understanding of the MS disease process, whereby lesions in the CNS impact signals sent throughout the body and often lead to significant motor dysfunction. It is likely that larger differences were not detected in the current analyses due to the high incidence of depression and fatigue in the general population, as well as in the MS population, and the small sample size of the current study.

Correlations between psychological and tVEP-CR measures were observed but were low to moderate, often only monocular, did not meet hypotheses, and were not consistent enough to permit generalizable interpretations of the results. This is again likely due to the small sample size used in the current study. A notable finding was replicated when exploring the negative correlation for the MS group between BDI and Band 2 delay. Scatterplots

depicting BDI scores and Band 2 delay estimates revealed two distinct MS subgroups. One subgroup exhibited long delays with low BDI scores, while the other subgroup had much shorter delays with high BDI scores. This is reminiscent of the MS subgroup distinction observed in earlier analyses elicited by H24 delay and MSC in Band 3. It is unclear what distinguishes these subgroups and given the small sample size in the current study conclusions cannot be drawn. Therefore, future research is needed to further elucidate the differential nature of the disease process in individuals with MS.

Clinical Implications

The current study is an investigation into the use of short-duration VEP responses to detect the presence of the MS disease process in the visual system. Evidence from this study supports the use of the short-duration stimulus in clinical settings, as it has been shown to produce valid results that are characteristic of the MS disease process (e.g., prolonged latencies). VEPs have also been useful in detecting clinically silent lesions and this method of assessment also supports this. VEPs have the potential to be used as biomarkers for MS and the current study suggests that frequency-domain measures may be most sensitive in detecting the presence of the disease when it does exist. In addition, this short-duration stimulus can be a rapid, objective clinical tool to be used in the study of treatment efficacy and disease progression over time.

Limitations and Future Directions

Given the small sample size of participants used in the current study, generalization to the MS population is limited. A future study using the same measures and a larger sample size would be useful to confirm study findings and perform additional analyses that require larger samples. In addition, this study used a cross-sectional design, collecting data at only

one time point, which limits interpretations regarding the reproducibility of results and changes in results over time. Longitudinal studies using tVEPs to evaluate the disease process in MS could be extremely informative in terms of how these measures may reflect changes in the CNS over time. The small sample size of the current study also limited the racial and ethnic diversity of the sample. Past research suggests that the MS disease process differs across racial and ethnic groups (Al-Kawaz et al., 2017; Johnson et al., 2010; Kister et al., 2010; Lichtman-Mikol et al., 2019). Thus, studying a larger sample with greater racial and ethnic diversity may help further explain certain findings, such as the different MS subgroups that were discovered when data were displayed graphically. Additionally, all but two of the MS participants were taking medication at the time of testing. It is unknown if these medications had an impact on VEP responses.

Research supports the utility of VEPs in MS to study disease status, progression, symptomatology, and treatment efficacy, and there remain opportunities to utilize the current methods of VEP analysis in these different areas. It is expected that the VEP would serve as a sensitive diagnostic tool and detect dysfunction even in cases without a history of optic neuritis. Work remains to be done on the development and application of new EEG techniques that tap specific neural mechanisms of interest related to the study domains mentioned above. In terms of statistical analyses, the current study used frequency bands defined based on a sample of healthy control participants using only the small check condition. Principal component analysis (PCA) could be run on a larger sample of MS participants using small and large checks to see how frequency mechanisms change with spatial conditions and to assess whether frequency bands as currently defined are representative of MS visual system functioning.

This study focused almost exclusively on assessing VEP responses in MS, while also including subjective depression and fatigue measures for exploratory analyses. Future studies would benefit from the inclusion of additional neuroimaging measures (e.g., OCT, fMRI, etc.) and/or neuropsychological measures for comparison with VEP measures, in order to better clarify and understand the underlying neural mechanisms involved in the disease process. Electroretinography has shown promise in the assessment of individuals with MS and this would be a strong complementary method to electrophysiological measures (Luo & Frishman, 2013; Porciatti & Ventura, 2004).

The current study is part of a larger ongoing study taking place at Holy Name Medical Center, which is examining the neurological correlates and interrelationships between EEG, ERG, and neuropsychological measures in individuals with MS. Individuals with *clinically isolated syndrome* (CIS) are an additional population of interest when studying the disease process in MS. This ongoing study is also collecting information on individuals with CIS, 80% of whom will receive a diagnosis of MS within two years of CIS diagnosis (Milo & Miller, 2014; National MS Society, 2018). This study intends to collect cross-sectional and longitudinal data in order to better understand the disease process over time and how measures of visual system function correlate with disease status and progression. This study also involves collecting subjective measures of depression and fatigue (BDI-II and FSMC) as well as a neuropsychological battery to assess cognitive functioning. The hope is to collect enough of these subjective and neuropsychological measures to warrant further analysis of the relationship between these measures and VEP function in individuals with MS.

References

Al-kawaz, M., Monohan, E., Morris, E., Perumal, J. S., Nealon, N., Vartanian, T., & Gauthier, S. A. (2017). Differential impact of multiple sclerosis on cortical and deep gray matter structures in African Americans and Caucasian Americans. *Journal of Neuroimaging, 27*, 333–338. doi: 10.1111/jon.12393

Balcer, L. J. (2006). Clinical practice. Optic neuritis. *New England Journal of Medicine,* (12), 1273-1280. doi:10.1056/NEJMcp053247

Balcer, L. J., Miller, D. H., Reingold, S. C., & Cohen, J. A. (2015). Vision and vision-related outcome measures in multiple sclerosis. *Brain, 138*(Pt 1), 11-27. doi:10.1093/brain/awu335

Balnytė, R., Ulozienė, I., Rastenytė, D., Vaitkus, A., Malcienė, L., & Laučkaitė, K. (2011). Diagnostic value of conventional visual evoked potentials applied to patients with multiple sclerosis. *Medicina, 47*(5), 263-269. doi:10.3390/medicina47050037

Beck, A. T., & Steer, R. A. (1987). *Manual for the Beck Depression Inventory*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996a). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. B., & Brown, G. K (1996b). Review of the Beck Depression Inventory-II. In P. A. Arbisi & R. F. Farmer (Eds.), T*he fourteenth mental measurements yearbook.*

Beck, A. T., Steer, R. A., & Brown, G. K. (2000). *BDI-Fast screen for medical patients: Manual.* San Antonio, TX: Psychological Corporation.

Beh, S. C., Frohman, E. M., & Frohman, T. (2016). Neuro-ophthalmologic manifestations of multiple sclerosis. In B. S. Giesser (Ed.), *Primer on Multiple Sclerosis* (2nd ed.), (pp. 185-212). New York, NY: Oxford University Press.

Benedict, R. H. B., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., & Weinstock-Guttman, B. (2006). Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society, 12*(4), 549–58. doi:10.1017/S1355617706060723

Bodis-Wollner, I. & Onofrj, M. (1982) System diseases and visual evoked potential diagnosis in neurology: Changes due to synaptic malfunction. *Annals of the New York Academy of Sciences. 388,* 327-348.

Bubl, E., Kern, E., Ebert, D., Bach, M., & Tebartz van Elst, L. (2010). Seeing gray when feeling blue? Depression can be measured in the eye of the diseased. *Biological Psychiatry, 68*(2), 205-208. doi:10.1016/j.biopsych.2010.02.009

Bubl, E., Kern, E., Ebert, D., Riedel, A., Tebartz van Elst, L., & Bach, M. (2015). Retinal dysfunction of contrast processing in major depression also apparent in cortical activity. *European Archives of Psychiatry and Clinical Neuroscience, 265*(4), 343-350. doi: 10.1007/s00406-014-0573-x

Buchanan, R. J., Zuniga, M. A., Carrillo-Zuniga, G., Chakravorty, B. J., Tyry, T., Moreau, R. L., … Vollmer, T. (2010). Comparisons of Latinos, African Americans, and Caucasians with multiple sclerosis. *Ethnicity & Disease*, *20*(Autumn), 451–457.

Butler, P. D., Schechter, I., Zemon, V., Schwartz, S. G., Greenstein, V. C., Gordon, J. … Javitt, D. C. (2001). Dysfunction of early-stage visual processing in schizophrenia. *The American Journal of Psychiatry, 158*(7), 1126–1133. doi:10.1176/appi.ajp.158.7.1126

Butler, P. D., Martinez, A., Foxe, J. J., Kim, D., Zemon, V., Silipo, G., Mahoney, J., Shpaner, M., Jalbrzikowski, M., & Javitt, D. C. (2007). Subcortical visual dysfunction in schizophrenia drives secondary cortical impairments. *Brain, 130*, 417-430 doi:10.1093/brain/awl233

Butt, A. M.; Fern, R. F.; & Matute, C. (2014) Review article: Neurotransmitter signaling in white matter. *Glia, 62,* 1762-1779. doi: 10.1002/glia.22674

Calderone, D. J., Martinez, A., Zemon, V., Hoptman, M. J., Hu, G., Watkins, J. E., Javitt, D. C., & Butler, P. D. (2013). Comparison of psychophysical, electrophysiological, and fMRI assessment of visual contrast responses in patients with schizophrenia. *NeuroImage, 67*, 153-162. doi:10.1016/j.neuroimage.2012.11.019

Camisa, J., Mylin, L. H., & Bodis-Wollner, I. (1981). The effect of stimulus orientation on the visual evoked potential in multiple sclerosis. *Annals of Neurology*, *10*(6), 532-539.

Chilinska, A., Ejma, M., Turno-Krecicka, A., Guranski, K., & Misiuk-Hojlo, M. (2016). Analysis of retinal nerve fibre layer, visual evoked potentials and relative afferent pupillary defect in multiple sclerosis patients. *Clinical Neurophysiology, 127*(1), 821-826. doi:10.1016/j.clinph.2015.06.025

Chirapapaisan, N., Laotaweerungsawat, S., Chuenkongkaew, W., Samsen, P., Ruangvaravate, N., Thuangtong, A., & Chanvarapha, N. (2015). Diagnostic value of

visual evoked potentials for clinical diagnosis of multiple sclerosis. *Documenta Ophthalmologica*, *130*(1), 25–30. doi: 10.1007/s10633-014-9466-6

Conte, M. M., & Victor, J. D. (2009). VEP indices of cortical lateral interactions in epilepsy treatment. *Vision Research*, *49*(9), 898-906. doi:10.1016/j.visres.2008.04.030

Coppola, G., Parisi, V., Di Lorenzo, C., Serrao, M., Magis, D., Schoenen, J., & Pierelli, F. (2013). Lateral inhibition in visual cortex of migraine patients between attacks. *The Journal of Headache and Pain*, *14*(1), 20. doi:10.1186/1129-2377-14-20

Dacey, D. M. & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. Proc Natl Acad Sci U S A 89, 9666-70.

Di Maggio, G., Santangelo, R., Guerrieri, S., Bianco, M., Ferrari, L., Medaglini, S., . . . Leocani, L. (2014). Optical coherence tomography and visual evoked potentials: which is more sensitive in multiple sclerosis? *Multiple Sclerosis Journal, 20*(10), 1342-1347. doi:10.1177/1352458514524293

Dobie, R. A. & Wilson, M. J. (1989). Analysis of auditory evoked potentials by magnitude-squared coherence. *Ear Hear*, *10*, 2-13.

Dobie, R. A. & Wilson, M. J. (1993). Objective response detection in the frequency domain. *Electroencephalogr Clin Neurophysiol* 88, 516-24.

Duwaer, A. L. & Spekreijse, H. (1978). Latency of luminance and contrast evoked potentials in multiple sclerosis patients. *Electroencephalogr Clin Neurophysiol* **45**, 244-58.

Ebers, G. (2013). Interactions of environment and genes in multiple sclerosis. *Journal of the Neurological Sciences*, *334*(1), 161–163. doi:10.1016/j.jns.2013.08.018

Freedman, M. S., Selchen, D., Arnold, D. L., Prat, A., Banwell, B., Yeung, M., … Canadian Multiple Sclerosis Working Group. (2013). Treatment optimization in MS: Canadian

MS working group updated recommendations. *The Canadian Journal of Neurological Sciences. Le Journal Canadien Des Sciences Neurologiques*, *40*(3), 307–323. doi:10.1017/S0317167100014244

Galetta, K. M., & Balcer, L. J. (2013). Measures of visual pathway structure and function in MS: Clinical usefulness and role for MS trials. *Multiple Sclerosis and Related Disorders*, *2*(3), 172 182. doi:10.1016/j.msard.2012.12.004

Garg, N., & Smith, T. W. (2015). An update on immunopathogenesis, diagnosis, and treatment of multiple sclerosis. *Brain and Behavior*, *5*(9), 1–13. doi:10.1002/brb3.362

Giovannoni, G. (2006). Multiple sclerosis related fatigue: Editorial commentary. *Journal of Neurology Neurosurgery and Psychiatry, 77*, 2-3. doi:10.1136/jnnp.2005.074948

Goldenholz, D. M., Ahlfors, S. P., Hämäläinen, M. S., Sharon, D., Ishitobi, M., Vaina, L. M., & Stufflebeam, S. M. (2009). Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human brain mapping*, *30*(4), 1077–1086. doi:10.1002/hbm.20571

Graham, S. L., & Kilstorner, A. (2015). Afferent visual pathways in multiple sclerosis: A review. *Clinical & Experimental Ophthalmology, 45*(1), 62-72. doi:10.1111/ceo.12751

Grecescu, M. (2014). Optical coherence tomography versus visual evoked potentials in detecting subclinical visual impairment in multiple sclerosis. *Journal of Medicine and Life, 7*(4), 538-541.

Greenstein, V. C., Seliger, S., Zemon, V. & Ritch, R. (1998). Visual evoked potential assessment of the effects of glaucoma on visual subsystems. *Vision Res* **38**, 1901-11.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *75*(6), 424-429.

Gronseth, G. S., & Ashman, E. J. (2010). Practice parameter: The usefulness of evoked potentials in identifying clinically silent lesions in patients with suspected multiple sclerosis (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, *54*, 1720 1725.

Grose-Fifer, J., Zemon, V., & Gordon, J. (1994). Temporal tuning and the development of lateral interactions in the human visual system. *Investigative Ophthalmology and Visual Science, 35*(7), 2999-3010.

Halliday, A. M., McDonald, W. I., & Mushin, J. (1973a). Visual evoked response in diagnosis of multiple sclerosis. *British Medical Journal, 4*(5893), 661-664.

Halliday, A. M., McDonald, W. I., & Mushin, J. (1973b). Delayed pattern-evoked response in optic neuritis in relation to visual acuity. *Transactions of the Ophthalmological Societies of the United Kingdom, 93,* 315-324.

Hamurcu, M., Orhan, G., Sarıcaoğlu, M. S., Mungan, S., & Duru, Z. (2017). Analysis of multiple sclerosis patients with electrophysiological and structural tests. *International Ophthalmology*, *37*(3), 649–653. doi:10.1007/s10792-016-0324-2

Hardmeier, M., Hatz, F., Naegelin, Y., Hight, D., Schindler, C., Kappos, L., … Fuhr, P. (2014). Improved characterization of visual evoked potentials in multiple sclerosis by topographic analysis. *Brain Topography*, *27*(2), 318–327. doi:10.1007/s10548-013-0318-6

Hardmeier, M., Leocani, L., & Fuhr, P. (2017). A new role for evoked potentials in MS? Repurposing evoked potentials as biomarkers for clinical trials in MS. *Multiple Sclerosis Journal*, *23*(10), 1309–1319. doi:10.1177/1352458517707265

Hartline, H. K. (1938). The discharge of impulses in the optic nerve of Pecten in response to illumination of the eye. *Journal of Cellular and Comparative Physiology, 11*(3), 465-478.

Ha-Vinh, P., Nauleau, S., Clementz, M., Régnard, P., Sauze, L., & Clavaud, H. (2016). Geographic variations of multiple sclerosis prevalence in France: The latitude gradient is not uniform depending on the socioeconomic status of the studied population. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, *2*, 1-14. doi:10.1177/2055217316631762

Hedström, A., Olsson, T., & Alfredsson, L. (2016). Body mass index during adolescence, rather than childhood, is critical in determining MS risk. *Multiple Sclerosis Journal*, 22(7), 878–883. doi:10.1177/1352458515603798

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience, 7,* 3378-3415.

Janáky, M., Jánossy, Horváth, G., Benedek, G., & Braunitzer, G. (2017). VEP and PERG in patients with multiple sclerosis, with and without a history of optic neuritis. *Documenta Ophthalmologica*, *134*(3), 185–193. doi:10.1007/s10633-017-9589-7

Jasper, H. H. (1958). The ten twenty electrode system of the International Federation. *Electroencephalography Journal, 10,* 371-375.

Johnson, B. A., Wang, J., Taylor, E. M., Caillier, S. J., Herbert, J., Khan, O. A., … Oksenberg, J. R. (2010). Multiple sclerosis susceptibility alleles in African Americans. *Genes and immunity*, *11*(4), 343–350. doi:10.1038/gene.2009.81

Kaplan, E. (2004). The M, P, and K pathways of the primate visual system. In L. M. Chalupa & J. S. Werner (Ed.), *The Visual Neurosciences* (pp. 481–493). Cambridge, MA: MIT Press.

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences, 83*(8), 2755-2757.

Kilstorner, A., Garrick, R., Barnett, M., Graham, S. L., Arvind, H., Sriram, P., & Yiannikas, C. (2012). Axonal loss in non–optic neuritis eyes of patients with multiple sclerosis linked to delayed visual evoked potential. *Neurology*, *80*(3), 242–245.

Kim, D., Zemon, V. M., Pinkhasov, E., Gordon, J. & Marks, D. (2002). Lateral interactions in cases of partial-complex epilepsy: a visual evoked potential study. *Investigative Ophthalmology and Visual Science*, 43: E-Abstract 1809.

Kim, D., Zemon, V., Saperstein, A., Butler, P. D., & Javitt, D. C. (2005). Dysfunction of early-stage visual processing in schizophrenia: Harmonic analysis. *Schizophrenia Research, 76*(1), 55–65. doi:10.1016/j.schres.2004.10.011

Kister, I., Chamot, E., Bacon, J. H., Niewczyk, P. M., De Guzman, R. A., Apatoff, B., … Herbert, J. (2010). Rapid disease course in African Americans with multiple sclerosis. *Neurology*, *75*(3), 217–223. doi:10.1212/WNL.0b013e3181e8e72a

Kotini, A., Anninos, P., & Tamiolakis, D. (2007). MEG mapping in MS patients. *Europa Medicophysica*, *43*(3), 345–348.

Krnjević, K. & Schwartz, S. (1967). The action of γ-Aminobutyric acid on cortical neurons. *Experimental Brain Research, 3*, 320-336. doi:10.1007/BF00237558

Kupersmith, M. J., Seiple, W. H., Nelson, J. I., & Carr, R. E. (1984). Contrast sensitivity loss in multiple sclerosis. Selectivity by eye, orientation, and spatial frequency measured with the evoked potential. *Investigative Ophthalmology and Visual Science*, *25*(6), 632-639.

Kurtzke, J. F. (2000). Epidemiology of multiple sclerosis. Does this really point toward an etiology? Lectio Doctoralis. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *21*(6), 383-403.

Lachapelle, J., Ouimet, C., Bach, M., Ptito, A. & McKerral, M. (2004). Texture segregation in traumatic brain injury--a VEP study. *Vision Res 44,* 2835-42.

LaRocca, X (2016). Cognitive impairment and mood disturbances. In B. S. Giesser (Ed.), *Primer on Multiple Sclerosis* (2nd ed.) (pp. 185-212). New York, NY: Oxford University Press.

Leocani, L., Rocca, M. A., & Comi, G. (2016). MRI and neurophysiological measures to predict course, disability and treatment response in multiple sclerosis. *Current Opinion in Neurology*, *29*(3), 243–253. doi:10.1097/WCO.0000000000000333

Lichtman-Mikol, S., Razmjou, S., Yarraguntla, K., Bao, F., Martinez-Santiago, C., Seraji-Bozorgzad, N., & Bernitsas, E. (2019). Racial differences in retinal neurodegeneration as a surrogate marker for cortical atrophy in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *31*, 141–147. doi: 10.1016/j.msard.2019.04.001

Liu, C. S., Bryan, R. N., Miki, A., Woo, J. H., Liu, G. T., & Elliott, M. A. (2006). Magnocellular and parvocellular visual pathways have different blood oxygen level-dependent signal time courses in human primary visual cortex. *American Journal of Neuroradiology*, *27*(8), 1628–1634. Retrieved from http://www.ajnr.org/content/27/8/1628

Livingstone, M. S., & Hubel, D. H. (1988). Segregation of form, color, movcmcnt, and depth: Anatomy, physiology, and perception. *Science, 240,* 740-749.

Luo, X., & Frishman, L. J. (2011). Retinal pathway origins of the pattern electroretinogram (PERG). *Investigative Ophthalmology & Visual Science, 52*(12), 8571-8584. doi:10.1167/iovs.11-8376.

Macmillan, N. A. & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed.) Mahwah, NJ: Lawrence Erlbaum Associates.

Marozas, D. S., & May, D. C. (1985). Effects of figure-ground reversal on the visual-perceptual and visuo-motor performances of cerebral palsied and normal children. *Perceptual and Motor Skills, 60*(2), 591-598.

McAlpine D. (1972) Multiple sclerosis: a reappraisal. Edinburgh: Churchill Livingstone.

McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H. P., Lublin, F. D., ... Wolinsky, J.S. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, 50:121–127. doi:10.1002/ana.1032

Merigan, W. H., & Maunsell, J. H. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16*, 369-402.

Meuth, S. G., Bittner, S., Seiler, C., Göbel, K., & Wiendl, H. (2011). Natalizumab restores evoked potential abnormalities in patients with relapsing-remitting multiple sclerosis. *Multiple Sclerosis Journal*, *17*(2), 198–203. doi:10.1177/1352458510386998

Milo, R., & Miller, A. (2014). Revised diagnostic criteria of multiple sclerosis. *Autoimmunity Reviews*, *13*(4-5), 518–524. doi:10.1016/j.autrev.2014.01.012

Moskowitz, A., & Sokol, S. (1983). Developmental changes in the human visual system as reflected by the latency of the Pattern Reversal VEP. *Electroencephalography and Clinical Neurophysiology*, *56*, 1–15.

Munger, K. L., Hongell, K., Åivo, J., Soilu-Hänninen, M., Surcel, H. M., & Ascherio, A. (2017). 25-Hydroxyvitamin D deficiency and risk of MS among women in the Finnish maternity cohort. *Neurology*, *89*(15), 1578–1583. doi:10.1212/WNL.0000000000004489

Naismith, R. T., Tutlam, N. T., Xu, J., Shepherd, J. B., Klawiter, E. C., Song, S. K., Cross, A. H. (2009). Optical coherence tomography is less sensitive than visual evoked potentials in optic neuritis. *Neurology*, *73*:46–52. doi:10.1212/WNL.0b013e3181aaea32

Naismith, R. T., Xu, J., Tutlam, N. T., Lancia, S., Trinkaus, K., … Song, S. K. (2012) Diffusion tensor imaging in acute optic neuropathies. *Archives of Neurology, 69:*65-71.

National Multiple Sclerosis Society. (2018). *About MS: MS overview, disease courses, epidemiology, and theories of causation.* [online] Available at: https://www.nationalmssociety.org/For-Professionals/Clinical-Care/About-MS [Accessed 6 July 2019]

National Multiple Sclerosis Society. (n.d.). *Disease-modifying therapies for MS.* [online] Available at: https://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/Brochure-The-MS-Disease-Modifying-Medications.pdf [Accessed 7 July 2019]

National Multiple Sclerosis Society. (n.d.). *Magnetic Resonance Imaging (MRI).* [online] Available at: https://www.nationalmssociety.org/Symptoms-Diagnosis/MRI [Accessed 7 July 2019]

Normann, C., Schmitz, D., Fürmaier, A., Döing, C., & Bach, M. (2007). Long-term plasticity of visually evoked potentials in humans is altered in major depression. *Biological Psychiatry, 62*(5), 373-380. doi:10.1016/j.biopsych.2006.10.006

Odom, J., Bach, M., Brigell, M., Holder, G. E., McCulloch, D. L., Mizota, A., & Tormene, A. P. (2016). ISCEV standard for clinical visual evoked potentials (2016 update). *Documenta Ophthalmologica, 133*(1), 1-9. doi: 10.1007/s10633-016-9553-y

Parallel Streams of Information from Retina to Cortex. In Purves D., et al. (Eds.). *Neuroscience* (2nd ed.). Sunderland, MA: Sinauer Associates.

Patsopoulos, N. A., Bayer Pharma MS Genetics Working Group, Steering Committees of Studies Evaluating IFNβ-1b and a CCR1-Antagonist, ANZgene Consortium, GeneMSA, International Multiple Sclerosis Genetics Consortium, … de Bakker, P. I. (2011). Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Annals of neurology*, *70*(6), 897–912. doi:10.1002/ana.22609

Penner, I., Raselli, C., Stöcklin, M., Opwis, K., Kappos, L., & Calabrese, P. (2009). The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new

instrument to assess multiple sclerosis-related fatigue. *Multiple Sclerosis*, 15(12), 1509–1517. doi:10.1177/1352458509348519

Pokryszko-Dragan, A., Bilinska, M., Gruszka, E., Kusinska, E., & Podemski, R. (2015). Assessment of visual and auditory evoked potentials in multiple sclerosis patients with and without fatigue. *Neurological Sciences, 36*, 235-242. doi:10.1007/s10072-014-1953-8

Pollack, I. & Norman, D.A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, *1*, 125-126.

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., & Filippi, M., … Wolinsky, J. S. (2011) Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology*; *69*:292–302. doi:10.1002/ana.22366

Polman, C. H., Reingold, S. C., Edan, C., Filippi, M., Hartung, H. P., Kappos, L., … Wolinsky, J. S. (2005) Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Annals of Neurology*, *58*(6), 840–846. doi: 10.1002/ana.20703 doi:10.1002/ana.20703

Porciatti, V., & Ventura, L. M. (2004). Normative data for a user-friendly paradigm for pattern electroretinogram recording. *Ophthalmology, 111*(1), 161-168. doi:10.1016/j.ophtha.2003.04.007

Rao, S. M., Leo, G. J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology*, *41*, 685–691. doi:10.1212/wnl.41.5.685

Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology,*

20(3), 238-248.

Regan, D. (1989). Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine.

Regan, D., & Neima, D. (1984). Visual fatigue and visual evoked potentials in multiple sclerosis, glaucoma, ocular hypertension and Parkinson's disease. *Journal of Neurology Neurosurgery and Psychiatry, 47*, 673 678. doi:10.1136/jnnp.47.7.673

Sacco, R., Santangelo, G., Stamenova, S., Bisecco, A., Bonavita, S., Lavorgna, L., . . . Gallo, A. (2016). Psychometric properties and validity of Beck Depression Inventory II in multiple sclerosis. *European Journal of Neurology, 23*, 744-750. doi:10.1111/ene.12932

Sanders, A. S., Foley, F. W., Larocca, N. G., & Zemon, V. (2000). The Multiple Sclerosis Intimacy and Sexuality Questionnaire-19 (MSISQ-19). *Sexuality and Disability*, *18*(1), 3–26.

Schiller, P. H. (1982). Central connections of the retinal ON and OFF pathways. *Nature, 297*(5867), 580-583.

Shapley, R. M., & Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology, 285*, 275-298.

Siper, P. M., Zemon, V., Gordon, J., George-Jones, J., Lurie, S., … Kolevzon, A. (2016). Rapid and Objective Assessment of Neural Function in Autism Spectrum Disorder Using Transient Visual Evoked Potentials. *PLOS ONE* 11.

Sokol, S., Zemon, V., & Moskowitz, A. (1992). Development of lateral interactions in the infant visual system. *Visual Neuroscience, 8*(1), 3-8.

Solaro, C., Trabucco, E., Signori, A., Martinelli, V., Radaelli, M., Centonze, D., . . . Truini, A. (2016). Depressive symptoms correlate with disability and disease course in multiple sclerosis patients: An Italian multi-center study using the Beck Depression Inventory. *PLoS One, 11*(9), e0160261. doi: 10.1371/journal.pone.0160261

Sumowski, J. F., & Leavitt, V. M. (2013). Cognitive reserve in multiple sclerosis. *Multiple Sclerosis Journal, 19*(9), 1122 1127. doi:10.1177/1352458513498834

Takemura, M. Y., Hori, M., Yokoyama, K., Hamasaki, N., Suzuki, M., Kamagata, K., . . . Aoki, S. (2016). Alterations of the optic pathway between unilateral and bilateral optic nerve damage in multiple sclerosis as revealed by the combined use of advanced diffusion kurtosis imaging and visual evoked potentials. *Magnetic Resonance Imaging*. doi:10.1016/j.mri.2016.04.011

Tewarie, P., Balk, L. J., Hillebrand, A., Steenwijk, M. D., Uitdehaag, B. M. J., Stam, C. J., & Petzold, A. (2017). Structure-function relationships in the visual system in multiple sclerosis: An MEG and OCT study. *Annals of Clinical and Translational Neurology, 4*(9), 614–621. doi:10.1002/acn3.415

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., & Coetzee, T. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurology, 17*, 162-173.

Thurtell, M. J., Bala, E., Yaniglos, S. S., Rucker, J. C., Peachey, N. S., & Leigh, R. J. (2009). Evaluation of optic neuropathy in multiple sclerosis using low-contrast visual evoked potentials. *Neurology, 73*(22), 1849–1857.

Tootell, R. B., Hamilton, S. L., & Switkes, E. (1988). Functional anatomy of macaque striate cortex: IV. Contrast and magno-parvo streams. *Journal of Neuroscience, 8,* 1594-1609.

Tootell, R. B., & Nasr, S. (2017). Columnar segregation of magnocellular and parvocellular streams in human extrastriate cortex. *Journal of Neuroscience, 37,* 8014-8032.

Van Kessel, K., Moss-Morris, R., Willoughby, E., Chalder, T., Johnson, M. H., & Robinson, E. (2008). A randomized controlled trial of cognitive behavior therapy for multiple sclerosis fatigue. *Psychosomatic Medicine*, *70*(2), 205–213. doi:10.1097/PSY.0b013e3181643065

Victor J. D., & Mast, J. (1991) A new statistic for steady-state evoked potentials, *Electroencephalography and Clinical Neurophysiology*, 78:378-388.

Watson, T. M., Ford, E., Worthington, E., & Lincoln, N. B. (2014). Validation of mood measures for people with multiple sclerosis. *International Journal of MS Care, 16*(2), 105-109. doi:10.7224/1537-2073.2013-013

Weinger, P. M., Zemon, V., Soorya, L., & Gordon, J. (2014). Low-contrast response deficits and increased neural noise in children with autism spectrum disorder. *Neuropsychologia*, *63*, 10–18. doi:10.1016/j.neuropsychologia.2014.07.031

Wallin, M. T., Culpepper, W. J., Campbell, J. D., Nelson, L. M., Langer-Gould, A., Marrie, R. A., … US Multiple Sclerosis Prevalence Workgroup. (2019). The prevalence of MS in the United States: A population-based estimate using health claims data. *Neurology*, *92*(10), e1029–e1040. doi: 10.1212/WNL.0000000000007035

Yalachkov, Y., Soyda, D., Bergmann, J., Frisch, S., Behrens, M., Foerch, C., & Gehrig, J. (2019). Determinants of quality of life in relapsing-remitting and progressive multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *30*, 33–37.

Zemon, V. (1984). The VEP: Analysis of functional subsystems in the brain. *Sixth Annual Conference IEEE Engineering in Medicine and Biology Society*, 426–431.

Zemon, V., Eisner, W., Gordon, J., Grose-Fifer, J., Tenedios, F., & Shoup, H. (1995). Contrast-dependent responses in the human visual system: childhood through adulthood. *International Journal of Neuroscience, 80*, 181-201.

Zemon, V., & Gordon, J. (2006). Luminance-contrast mechanisms in humans: visual evoked potentials and a nonlinear model. *Vision Research*, *46*(24), 4163–80. doi:10.1016/j.visres.2006.07.007

Zemon, V. & Gordon, J. (2018). Quantification and statistical analysis of the transient visual evoked potential to a contrast- reversing pattern: A frequency- domain approach. *European Journal of Neuroscience.* 00:1–24. doi.org/10.1111/ejn.14049

Zemon, V., Gordon, J., O'Toole, L., Monde, K., Dolzhanskaya, V., Shapovalova, V., …Granader, Y. (2009). Transient visual evoked potentials (tVEPs) to contrast-reversing patterns: a frequency domain analysis. *Investigative Ophthalmology & Visual Science*, 50: E-Abstract 5880

Zemon, V., Gordon, J., & Welch, J. (1988). Asymmetries in ON and OFF visual pathways of humans revealed using contrast-evoked cortical potentials. *Visual Neuroscience*, *1,* 145–150.

Zemon, V., Hartmann, E., Gordon, J., & Prünte-Glowazki, A. (1997). An electrophysiological technique for assessment of the development of spatial vision. *Optometry and Vision Science*, *74*(9), 708–716.

Zemon, V., Kaplan, E., & Ratliff, F. (1980). Bicuculline enhances a negative component and diminishes a positive component of the visual evoked cortical potential in the cat. *Proceedings of the National Academy of Sciences*, *77*(12), 7476–7478.

Zemon, V., Kaplan, E., & Ratliff, F. (1986). The role of GABA-mediated intracortical inhibition in the generation of visual evoked potentials. *Frontiers of Clinical Neuroscience*, 3, 287-295.

Zemon, V., Pinkhasov, E., Gordon, J. (1993). Electrophysiological tests of neural models: Evidence for nonlinear binocular interactions in humans. In *Proceedings of the National Academy of Sciences* (90), 2975-2978.

Zemon, V., & Ratliff, F. (1982). Visual evoked potentials: evidence for lateral interactions. *Proceedings of the National Academy of Sciences*, *79*(18), 5723–5726. doi:10.1073/pnas.79.18.5723

Zemon, V., & Ratliff, F. (1984). Intermodulation components of the visual evoked potential: responses to lateral and superimposed stimuli. *Biological Cybernetics, 50*(6), 401-408.

Zemon, V, Tsai, J. C., Forbes, M., Al-Aswad, L. A., Chen, C. M., Gordon, J., … Jindra, L. F. (2008) Novel electrophysiological instrument for rapid and objective assessment of magnocellular deficits associated with glaucoma. *Documenta Ophthalmologia, 117*:233-243.

Table 1.

Demographic and Clinical Characteristics for Controls (n = 16) and MS Cohorts (n = 18).

Variable	*Controls*	*MS*
Age (years), *Mdn* (Range)	30.0 (20-40)	32.5 (24-39)
Gender, *n* (%)		
Male	3 (18.8%)	4 (22.2%)
Female	13 (81.3%)	14 (77.8%)
Race, *n* (%)		
Caucasian	15 (93.8%)	9 (50%)
Hispanic	1 (6.3%)	5 (27.8%)
White/Hispanic	0 (0%)	3 (16.7%)
African American	0 (0%)	1 (5.6%)
Acuity, *Mdn (IQR)*		
Right eye (OD)	1.0 (0.2)	1.0 (0.2)
Left eye (OS)	1.0 (0.2)	1.0 (0.2)
Optic neuritis, *n* (%)		
No	16 (76.5%)	13 (72.2%)
Yes	0 (0%)	4 (22.2%)
Possible	0 (0%)	1 (5.6%)
Duration of illness (years), *Mdn* (Range)		6.9 (1-17) (n = 14)
BDI-II total, *Mdn* (Range)	2 (0-13) (n = 16)	9 (0-35) (n = 17)
FSMC total, *Mdn* (Range)	28 (20-76) (n = 16)	43 (20-80) (n = 15)
Cog, *Mdn* (Range)	14 (10-36)	19 (10-39)
Motor, *Mdn* (Range)	13 (10-40)	24 (10-41)

Note. Number of participants per group is noted for missing data. BDI-II ≡ Beck Depression Inventory-II; FSMC ≡ Fatigue Severity and Motor Scale; Cog ≡ cognitive.

Table 2.

Group Means, Standard Deviations, and Mann-Whitney tests for the small check condition, right eye.

Variable	Controls ($n = 16$)	Patients ($n = 18$)	Mann-Whitney test U	p
P60-N75 Amplitude (μV)	8.39 ± 4.53	6.08 ± 5.92	93.0	0.08
N75-P100 Amplitude (μV)	15.90 ± 5.96	14.01 ± 2.19	109.0	0.24
P100-N135 Amplitude (μV)	13.57 ± 6.09	13.68 ± 1.90	141.0	0.93
Power Band 1 (μV^2/band)	3.00 ± 1.42	2.85 ± 1.47	138.0	0.85
Power Band 2 (μV^2/band)	2.20 ± 0.81	2.03 ± 1.22	120.0	0.42
Power Band 3 (μV^2/band)	0.80 ± 0.34	0.60 ± 0.35	93.0	0.08
Power Band 4 (μV^2/band)	0.30 ± 0.12	0.35 ± 0.16	123.0	0.48
MSC Band 1	0.37 ± 0.17	0.40 ± 0.18	130.0	0.65
MSC Band 2	0.35 ± 0.15	0.31 ± 0.20	113.0	0.30
MSC Band 3	0.26 ± 0.17	0.16 ± 0.10	77.0	**0.02***
MSC Band 4	0.14 ± 0.09	0.14 ± 0.08	144.0	1.00
P60 Peak Time (ms)	50.06 ± 7.02	59.83 ± 14.52	84.5	**0.04***
N75 Peak Time (ms)	70.69 ± 4.53	82.50 ± 12.61	46.5	***p* < .01****
P100 Peak Time (ms)	96.88 ± 8.34	111.67 ± 14.81	50.5	***p* < .01****
N135 Peak Time (ms)	150.69 ± 17.37	154.33 ± 18.57	125.0	0.53
H24 Phase Delay (ms)	95.31 ± 5.10	109.76 ± 15.92	43.0	***p* < .001*****
Band 2 Delay (ms)	82.36 ± 17.32	99.05 ± 25.54	49.0	**0.01***

Note. Values are $M \pm SD$. Patients ≡ patients with multiple sclerosis. Power ≡ square root of total power within a frequency band. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3.

Group Means, Standard Deviations, and Mann-Whitney tests for the small check condition, left eye.

Variable	Controls ($n = 16$)	Patients ($n = 18$)	Mann-Whitney test U	p
P60-N75 Amplitude (μV)	8.69 ± 5.25	6.31 ± 5.18	86.0	**0.05***
N75-P100 Amplitude (μV)	14.89 ± 6.27	13.60 ± 8.59	118.0	0.38
P100-N135 Amplitude (μV)	12.16 ± 5.47	13.34 ± 7.17	133.0	0.72
Power Band 1 (μV^2/band)	2.76 ± 1.31	2.84 ± 1.40	144.0	1.00
Power Band 2 (μV^2/band)	2.15 ± 0.75	1.97 ± 1.07	118.0	0.38
Power Band 3 (μV^2/band)	0.77 ± 0.36	0.60 ± 0.35	95.0	0.10
Power Band 4 (μV^2/band)	0.30 ± 0.13	0.34 ± 0.19	132.0	0.70
MSC Band 1	0.34 ± 0.17	0.39 ± 0.15	122.0	0.46
MSC Band 2	0.34 ± 0.15	0.31 ± 0.16	124.0	0.51
MSC Band 3	0.24 ± 0.13	0.16 ± 0.12	71.0	**0.01***
MSC Band 4	0.13 ± 0.10	0.13 ± 0.07	134.0	0.75
P60 Peak Time (ms)	51.81 ± 7.23	58.83 ± 11.76	99.5	0.13
N75 Peak Time (ms)	71.44 ± 4.70	80.83 ± 9.47	54.5	***p* < .01****
P100 Peak Time (ms)	98.38 ± 9.71	115.44 ± 17.15	48.5	***p* < .01****
N135 Peak Time (ms)	148.38 ± 16.13	162.22 ± 20.41	85.5	**0.04***
H24 Phase Delay (ms)	96.91 ± 7.74	109.56 ± 14.01	55.0	***p* < .01****
Band 2 Delay (ms)	79.72 ± 13.95	97.60 ± 22.03	56.5	**0.01***

Note. Values are $M \pm SD$. Patients ≡ patients with multiple sclerosis. Power ≡ square root of power. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 4.

Group Means, Standard Deviations, and Mann-Whitney tests for the large check condition, right eye.

Variable	Controls ($n = 16$)	Patients ($n = 18$)	Mann-Whitney test U	p
P60-N75 Amplitude (μV)	4.45 ± 3.25	3.96 ± 2.60	126.0	0.74
N75-P100 Amplitude (μV)	11.11 + 4.68	10.52 + 4.36	131.0	0.87
P100-N135 Amplitude (μV)	10.94 ± 4.28	10.30 ± 3.43	134.0	0.96
Power Band 1 (μV^2/band)	2.48 ± 1.00	2.14 ± .84	122.0	0.46
Power Band 2 (μV^2/band)	1.69 ± 0.54	1.48 ± 0.58	115.0	0.33
Power Band 3 (μV^2/band)	0.51 ± 0.17	0.64 ± 0.32	115.0	0.33
Power Band 4 (μV^2/band)	0.25 ± 0.07	0.42 ± 0.27	83.5	**0.04***
MSC Band 1	0.33 ± 0.18	0.31 ± 0.13	134.0	0.75
MSC Band 2	0.28 ± 0.16	0.24 ± 0.12	124.0	0.51
MSC Band 3	0.16 ± 0.09	0.14 ± 0.05	118.0	0.38
MSC Band 4	0.12 ± 0.06	0.13 ± 0.05	114.0	0.31
P60 Peak Time (ms)	45.38 ± 4.79	51.65 ± 14.53	108.5	0.33
N75 Peak Time (ms)	64.06 ± 4.40	74.88 ± 15.59	68.0	**0.01***
P100 Peak Time (ms)	91.44 ± 6.40	102.65 ± 14.43	69.0	**0.02***
N135 Peak Time (ms)	144.88 ± 25.02	154.24 ± 24.21	108.0	0.33
H24 Phase Delay (ms)	91.38 ± 10.65	102.14 ± 15.28	81.5	**0.03***
Band 2 Delay (ms)	75.93 ± 9.65	96.06 ± 18.47	28.0	***p* < .01****

Note. Values are $M \pm SD$. Patients ≡ patients with multiple sclerosis. Power ≡ square root of power. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 5.

Group means, standard deviations, and Mann-Whitney tests for the large check condition, left eye.

Variable	Controls (n = 16)	Patients (n = 18)	Mann-Whitney test	
			U	*p*
P60-N75 Amplitude (μV)	4.03 ± 2.46	3.48 ± 2.75	126.0	0.55
N75-P100 Amplitude (μV)	11.83 ± 3.90	9.41 ± 4.84	101.0	0.14
P100-N135 Amplitude (μV)	11.52 ± 3.61	9.45 ± 4.67	104.0	0.18
Power Band 1 (μV2/band)	2.54 ± 1.01	2.11 ± 0.75	108.0	0.22
Power Band 2 (μV2/band)	1.69 ± 0.42	1.46 ± 0.73	114.0	0.31
Power Band 3 (μV2/band)	0.50 ± 0.14	0.48 ± 0.28	101.0	0.14
Power Band 4 (μV2/band)	0.26 ± 0.10	0.29 ± 0.11	125.0	0.53
MSC Band 1	0.33 ± 0.16	0.32 ± 0.17	135.0	0.77
MSC Band 2	0.28 ± 0.13	0.26 ± 0.17	122.0	0.46
MSC Band 3	0.14 ± 0.07	0.11 ± 0.05	87.0	**0.05***
MSC Band 4	0.09 ± 0.04	0.11 ± 0.05	116.0	0.35
P60 Peak Time (ms)	42.81 ± 9.84	54.89 ± 15.87	69.0	**0.01***
N75 Peak Time (ms)	64.25 ± 5.39	76.28 ± 16.01	75.5	**0.02***
P100 Peak Time (ms)	92.81 ± 7.61	103.78 ± 16.84	81.5	**0.03***
N135 Peak Time (ms)	143.13 ± 22.41	148.61 ± 21.78	126.5	0.55
H24 Phase Delay (ms)	90.40 ± 5.55	103.37 ± 20.07	100.0	0.14
Band 2 Delay (ms)	80.32 ± 9.31	101.34 ± 34.40	42.0	0.09

Note. Values are $M \pm SD$. Patients ≡ patients with multiple sclerosis. Power ≡ square root of power. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 6.

Zero[th]-order correlations among amplitude and MSC/power frequency-band measures for controls (upper triangle, n = 16) and patients (lower triangle, n = 18) for the small check condition, right eye.

MEASURE	1	2	3	4	5	6	7	8	9	10	11
1) P60-N75 A	—	.661***	.032	-.321	.382	.523*	-.063	.-.048	.622*	**.789*** **	.021
2) N75-P100 A	.838***	—	.728***	.417	.573*	.157	-.348	**.582***	**.950*** **	**.699*** **	.414
3) P100-N135 A	**.674*** **	**.943*** **	—	.712**	.280	-.291	-.435	**.883*** **	.763***	.354	.563*
4) MSC BAND 1	.546*	.760***	.795***	—	.431	-.255	-.236	.672**	.318	-.177	.355
5) MSC BAND 2	**.853*** **	**.903*** **	.844***	.725**	—	.573*	.366	.132	.407	.259	.341
6) MSC BAND 3	.653**	.762***	.712**	.344	.620**	—	.566*	-.248	.058	.477	.039
7) MSC BAND 4	.664**	.786***	.790***	.592**	.698**	.650**	—	-.459	-.440	-.106	.168
8) POWER BAND 1	.686***	**.943*** **	**.952*** **	.820***	.779***	.756***	.736**	—	.586*	.370	.404
9) POWER BAND 2	.871***	**.961*** **	**.925*** **	.771***	**.945*** **	.665**	.756***	.872***	—	.740**	.477
10) POWER BAND 3	**.579***	.663***	.640***	.279	.499*	.723**	.381	.656**	.626**	—	.358
11) POWER BAND 4	.604***	.655***	.690***	.415	.574*	.452	.584*	.582*	.727**	.772***	—

Note. Amplitudes (A) for the three peak-to-trough amplitude measures, mean MSC values and square root of power estimates (based on vector-mean amplitudes for each individual observer) for the four frequency bands obtained in the current study: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz).

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed).

Table 7.

Zero[th]-order correlations among amplitude and MSC/power frequency-band measures for controls (upper triangle, n = 16) and patients (lower triangle, n = 18) for the small check condition, left eye.

MEASURE	1	2	3	4	5	6	7	8	9	10	11
1) P60-N75 A	—	.667***	.011	-.159	.680**	**.838***	.322	-.058	.723***	.814***	.354
2) N75-P100 A	.896***	—	.725***	.545*	.655**	.576*	.154	**.588***	**.882***	**.744***	**.522***
3) P100-N135 A	**.726***	**.921***	—	**.856***	.299	-.052	-.125	**.888***	.561*	.184	.311
4) MSC BAND 1	.416	.669**	.735**	—	.229	-.129	-.158	.830***	.247	.001	.173
5) MSC BAND 2	.795***	.805***	.668**	.505*	—	.544*	.459	.077	.624**	.448	.581*
6) MSC BAND 3	**.849***	.756***	.655**	.395	.696**	—	.612*	-.126	.501*	**.850***	.508*
7) MSC BAND 4	.041	-.032	-.069	-.187	.012	.060	—	-.223	.110	.276	.718**
8) POWER BAND 1	.697***	**.887***	**.964***	.696**	.532*	.628**	-.077	—	.463	.092	.151
9) POWER BAND 2	**.902***	**.973***	.893***	.613**	.848***	.735**	.012	.832***	—	.681**	.418
10) POWER BAND 3	**.814***	.807***	.804***	.405	.532*	.780***	-.152	.832***	.774***	—	.522*
11) POWER BAND 4	.148	.230	.355	-.014	-.036	-.013	.243	.410	.253	.485*	—

Note. Amplitudes (A) for the three peak-to-trough amplitude measures, mean MSC values and square root of power estimates (based on vector-mean amplitudes for each individual observer) for the four frequency bands obtained in the current study: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz).

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed).

Table 8.

Zero[th]-order correlations among amplitude and MSC/power frequency-band measures for controls (upper triangle, n = 16) and patients (lower triangle, n = 18) for the large check condition, right eye.

MEASURE	1	2	3	4	5	6	7	8	9	10	11
1) P60-N75 A	—	.547*	.026	.166	.711**	.663**	.384	.005	.474	.535*	.236
2) N75-P100 A	.737**	—	.636***	.782***	.757**	.358	.095	.685***	.479	.322	-.037
3) P100-N135 A	.393	.755***	—	**.845*** **	.471	-.038	.131	**.839*** **	.608*	.229	.319
4) MSC BAND 1	.095	.582*	.572*	—	.456	.040	-.070	**.861*** **	.354	.268	.129
5) MSC BAND 2	.545*	**.777*** **	.479	.574*	—	.549*	.214	.239	.648**	.358	.110
6) MSC BAND 3	.227	.415	.437	.310	.543*	—	.543*	-.213	.216	.698**	.213
7) MSC BAND 4	-.243	-.091	-.036	.036	.125	.572*	—	-.056	-.108	.122	.202
8) POWER BAND 1	.155	.547*	**.842*** **	**.678** **	.251	.277	-.040	—	.360	.068	.143
9) POWER BAND 2	**.750*** **	**.793*** **	.668***	.311	.705**	.431	.036	.350	—	.446	.477
10) POWER BAND 3	.023	.024	.447	-.044	-.193	.188	-.283	.265	.224	—	.668**
11) POWER BAND 4	-.162	-.070	.306	-.256	-.294	-.118	-.235	.123	.162	.734**	—

Note. Amplitudes (A) for the three peak-to-trough amplitude measures, mean MSC values and square root of power estimates (based on vector-mean amplitudes for each individual observer) for the four frequency bands obtained in the current study: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz).

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed).

Table 9.

Zero[th]-order correlations among amplitude and MSC/power frequency-band measures for controls (upper triangle, n = 16) and patients (lower triangle, n = 18) for the large check condition, left eye.

MEASURE	1	2	3	4	5	6	7	8	9	10	11
1) P60-N75 A	—	.507*	-.013	-.092	.492	.129	.264	.020	.551*	.137	.036
2) N75-P100 A	.722**	—	.763**	.527*	.423	-.257	.010	**.755*** **	.705***	-.158	.095
3) P100-N135 A	.314	.825***	—	**.573***	.279	-.108	.479	**.768*** **	.604*	-.104	.208
4) MSC BAND 1	.216	.601**	.703**	—	.158	-.223	.147	.746**	.040	-.074	.141
5) MSC BAND 2	**.779*** **	**.804*** **	.583*	.484*	—	.335	.033	.022	.578*	.061	-.285
6) MSC BAND 3	**.486***	.627**	.477*	.468	.741***	—	.432	-.433	.025	.471	-.111
7) MSC BAND 4	.146	.178	-.069	.207	.188	.019	—	.214	.006	.124	.398
8) POWER BAND 1	.267	.722***	**.840*** **	.687**	.331	.274	.088	—	.253	-.148	.302
9) POWER BAND 2	**.760*** **	**.876*** **	.741***	.337	**.816*** **	.577*	.051	.566*	—	.104	.070
10) POWER BAND 3	-.007	.318	.428	-.014	.041	.344	.292	.246	.366	—	.596*
11) POWER BAND 4	-.069	.253	.361	.245	-.048	.016	.314	.312	.168	.608**	—

Note. Amplitudes (A) for the three peak-to-trough amplitude measures, mean MSC values and square root of power estimates (based on vector-mean amplitudes for each individual observer) for the four frequency bands obtained in the current study: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz).

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed).

Table 10.

Zeroth-order correlations among time and frequency-domain latency measures for controls (upper triangle, *n = 16*) and patients (lower triangle, *n = 18*) for the small check condition, right eye.

MEASURE	1	2	3	4	5	6
1) P60 LATENCY	—	**.670****	.359	-.056	.077	-.139
2) N75 LATENCY	**.735****	—	**.767****	-.115	.444	.078
3) P100 LATENCY	**.626****	**.898*****	—	-.401	**.692****	**.519***
4) N135 LATENCY	**.623****	**.689****	**.792*****	—	-.158	-.414
5) H24 DELAY	.455	**.844*****	**.856*****	**.527***	—	**.613***
6) BAND 2 DELAY	.311	**.698****	**.831*****	**.759****	**.610***	—

Note. Latencies for the four peak time measures, 24th harmonic delay estimates (to predict P100 peak time), and Band 2 delay estimates.

* $p < .05$. ** $p < .01$. *** $p < .001$. (two-tailed).

Table 11.

Zero[th]-order correlations among time and frequency-domain latency measures for controls (upper triangle, *n = 16*) and patients (lower triangle, *n = 18*) for the small check condition, left eye.

MEASURE	1	2	3	4	5	6
P60 LATENCY	—	**.659****	**.566***	-.247	**.557***	.285
N75 LATENCY	**.623****	—	**.690****	.087	**.636****	.354
P100 LATENCY	.198	**.702****	—	.003	**.887*****	.268
N135 LATENCY	.030	**.547***	**.913*****	—	-.072	-.511
H24 DELAY	.071	**.539***	**.737*****	**.703****	—	.445
BAND 2 DELAY	.360	**.798*****	**.778*****	**.564***	**.547***	—

Note. Latencies for the four peak time measures, 24[th] harmonic delay estimates (to predict P100 peak time), and Band 2 delay estimates.

* $p < .05$. ** $p < .01$ (two-tailed).

Table 12.

Zero[th]-order correlations among time and frequency-domain latency measures for controls (upper triangle, *n = 16*) and patients (lower triangle, *n = 18*) for the large check condition, right eye.

MEASURE	1	2	3	4	5	6
1) P60 LATENCY	—	.435	-.069	-.279	.138	.063
2) N75 LATENCY	**.676****	—	.053	.092	.057	-.117
3) P100 LATENCY	.442	**.824*****	—	.319	**.815*****	.038
4) N135 LATENCY	**.510***	**.490***	**.540***	—	.009	**-.595***
5) H24 DELAY	**.487***	**.741****	**.928*****	.469	—	.530
6) BAND 2 DELAY	.279	.399	**.758****	.290	**.834*****	—

Note. Latencies for the four peak time measures, 24[th] harmonic delay estimates (to predict P100 peak time), and Band 2 delay estimates.

* $p < .05$. ** $p < .01$ (two-tailed).

Table 13.

Zeroth-order correlations among time and frequency-domain latency measures for controls (upper triangle, *n = 16*) and patients (lower triangle, *n = 18*) for the large check condition, left eye.

MEASURE	1	2	3	4	5	6
1) P60 LATENCY	—	.212	-.163	-.182	-.298	-.288
2) N75 LATENCY	**.694****	—	.260	-.306	.189	.214
3) P100 LATENCY	**.519***	**.852*****	—	.269	.091	-.351
4) N135 LATENCY	.396	**.664****	**.748*****	—	-.145	**-.565***
5) H24 DELAY	.353	**.748*****	**.935*****	**.688****	—	**.651***
6) BAND 2 DELAY	**.887*****	**.821****	**.828****	.476	**.841****	—

Note. Latencies for the four peak time measures, 24th harmonic delay estimates (to predict P100 peak time), and Band 2 delay estimates.

* $p < .05$. ** $p < .01$ (two-tailed).

Table 14.

Receiver operating characteristic (ROC) curve analysis – small checks

(MS: n = 18, Controls: n = 16)

Variable	AUC	*p*	95% CI	
			LL	*UL*
Power Band 3 (OD) (*n* = 18/16)	.68	.08	.50	.86
Power Band 3 (OS)	.67	.09	.49	.85
MSC Band 3 (OD)	**.73**	**.02***	.56	.90
MSC Band 3 (OS)	**.75**	**.01***	.59	.92
P60 Latency (OD)	**.71**	**.04***	.53	.88
P60 Latency (OS)	.66	.13	.46	.85
N75 Latency (OD)	**.84**	**< .01****	.71	.97
N75 Latency (OS)	**.81**	**< .01****	.66	.96
P100 Latency (OD)	**.83**	**< .01****	.69	.96
P100 Latency (OS)	**.83**	**< .01****	.70	.97
Band 2 Delay (OD) (*n* = 10/11)	**.84**	**< .01****	.66	1
Band 2 Delay (OS) (*n* = 10/11)	**.82**	**.01***	.62	1
H24 Delay (OD)	**.85**	**< .01****	.72	.98
H24 Delay (OS)	**.81**	**< .01****	.66	.96

Note. *n's* vary because of missing data, *n* ≡ number of patients/controls, *LL* ≡ lower limit, *UL* ≡ upper limit, A' *n*'s ≡ number of patients/controls

* $p < .05$. ** $p < .01$.

Table 15.

Receiver operating characteristic (ROC) curve analysis– large checks

(MS: n = 18, Controls: n = 16)

Variable	AUC	*p*	95% CI	
			LL	*UL*
P60 Latency (OD)	.60	.32	.40	.80
P60 Latency (OS)	**.75**	**.02***	.58	.92
N75 Latency (OD)	**.75**	**.01***	.57	.93
N75 Latency (OS)	**.72**	**.03***	.54	.90
P100 Latency (OD)	**.75**	**.02***	.58	.91
P100 Latency (OS)	**.72**	**.03***	.55	.89
N135 Latency (OD)	.60	.31	.40	.80
N135 Latency (OS)	.56	.58	.36	.76
Band 2 Delay (OD) (n = 10/11)	**.83**	**.01***	.65	1
Band 2 Delay (OS) (n = 10/11)	.66	.21	.42	.91
H24 Delay (OD)	**.72**	**.03***	.54	.89
H24 Delay (OS)	.65	.13	.46	.84

Note. *n's* vary because of missing data, $n \equiv$ number of patients/controls, *LL* ≡ lower limit, *UL* ≡ upper limit.

* $p < .05$. ** $p < .01$.

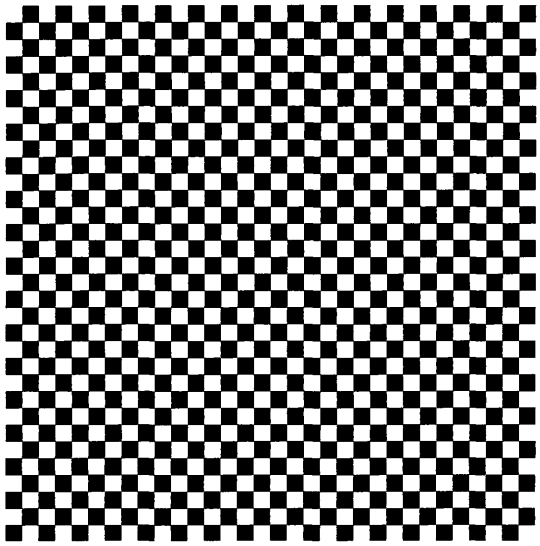


Figure 1. Contrast-reversing checkerboard stimulus pattern used to elicit tVEPs. Checkerboards were presented as 64 x 64 checks or 16 x 16 checks for small and large conditions, respectively.

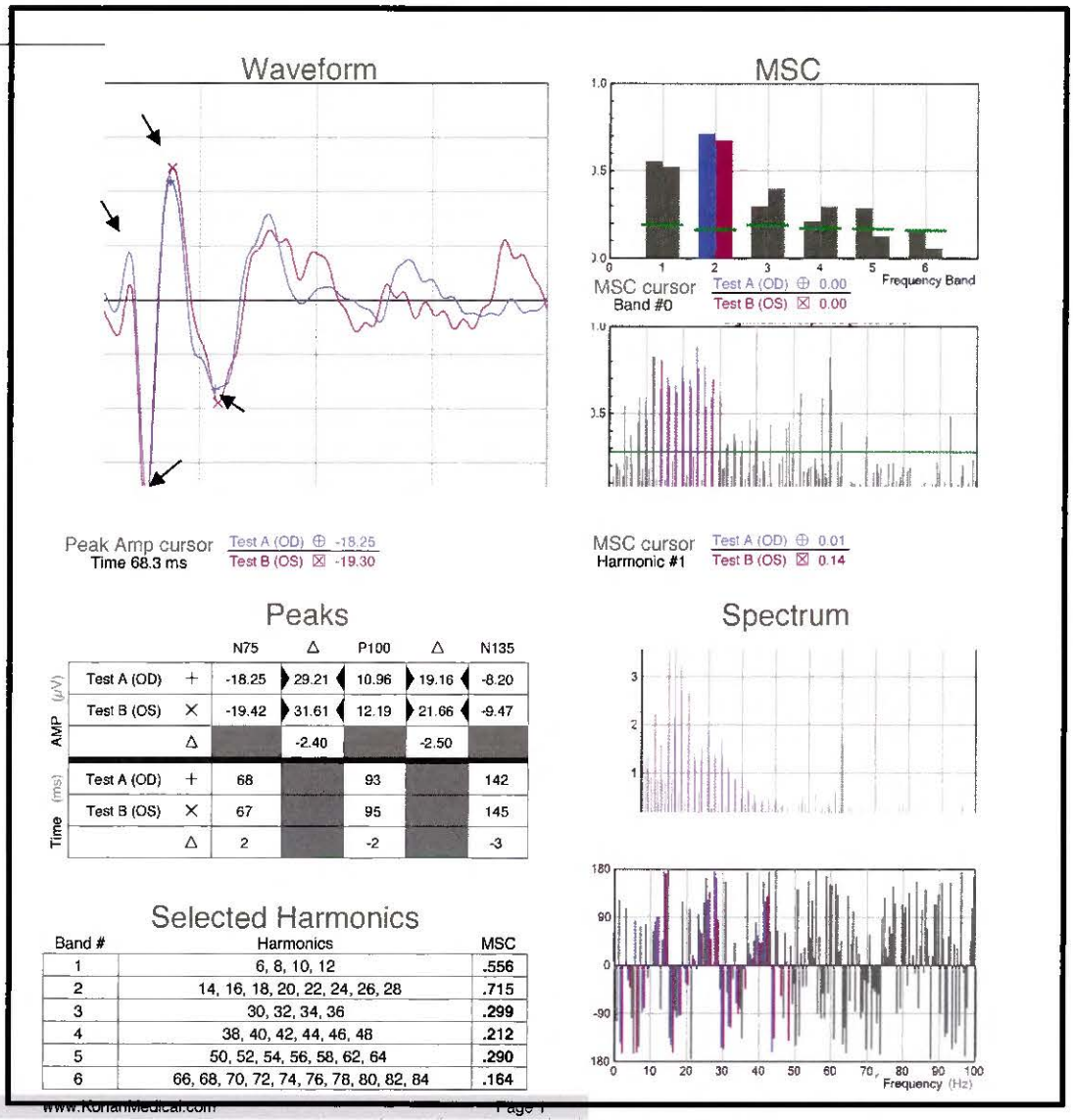


Figure 2A. Representative output from EvokeDx for a control participant for the small check condition.

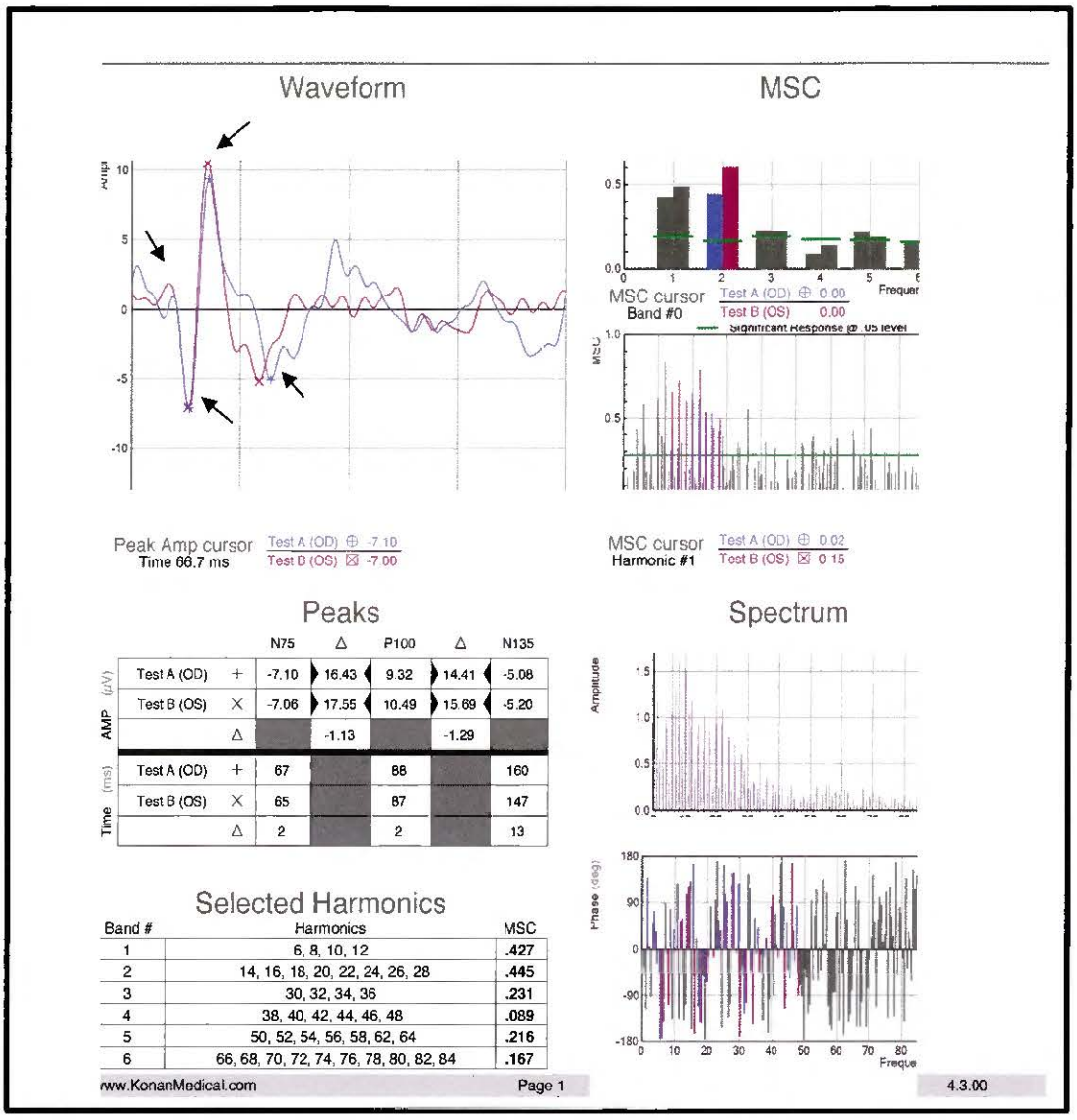


Figure 2B. Representative output from EvokeDx for a control participant for the large check condition.

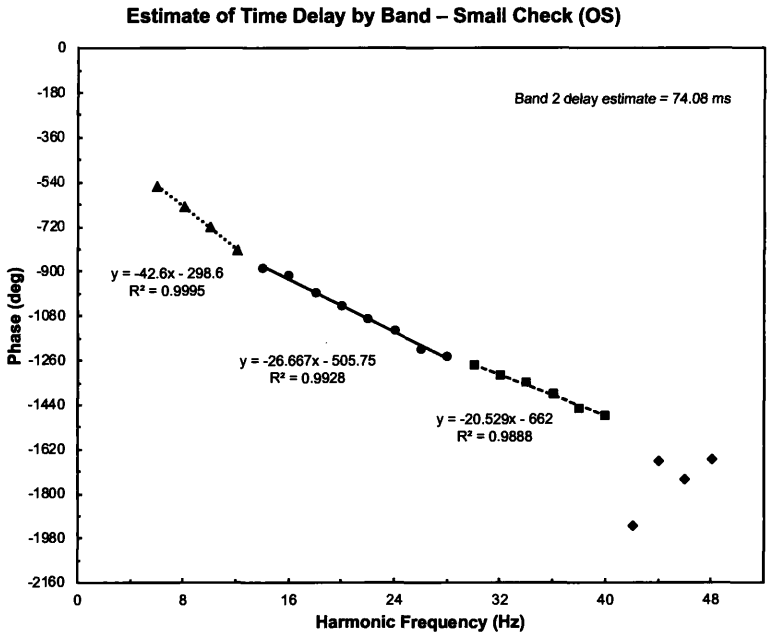
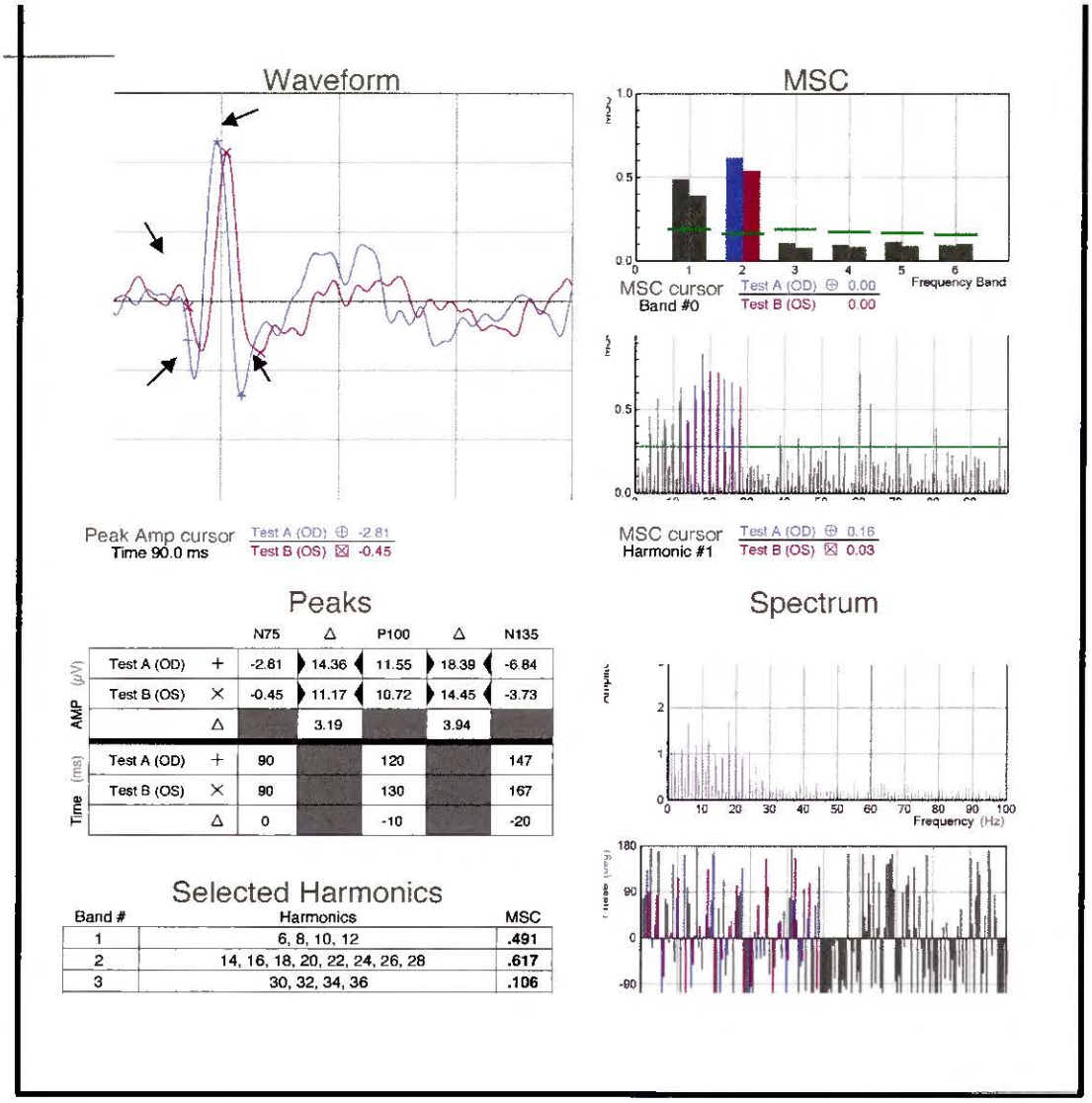


Figure 2C. Representative phase vs. harmonic frequency plot for a control participant for the small check condition – left eye. Slopes were plotted for Bands 1 through 3 and an estimate for Band 2 delay was calculated based on the derived slope.



		N75	Δ	P100	Δ	N135
AMP (μV)	Test A (OD) +	-2.81	14.36	11.55	18.39	-6.84
	Test B (OS) ×	-0.45	11.17	10.72	14.45	-3.73
	Δ		3.19		3.94	
Time (ms)	Test A (OD) +	90		120		147
	Test B (OS) ×	90		130		167
	Δ	0		-10		-20

Band #	Harmonics	MSC
1	6, 8, 10, 12	.491
2	14, 16, 18, 20, 22, 24, 26, 28	.617
3	30, 32, 34, 36	.106

Figure 3A. Representative output from EvokeDx for an MS participant for the small check condition. P100 peak time is prominent and delayed compared to the control participant waveform.

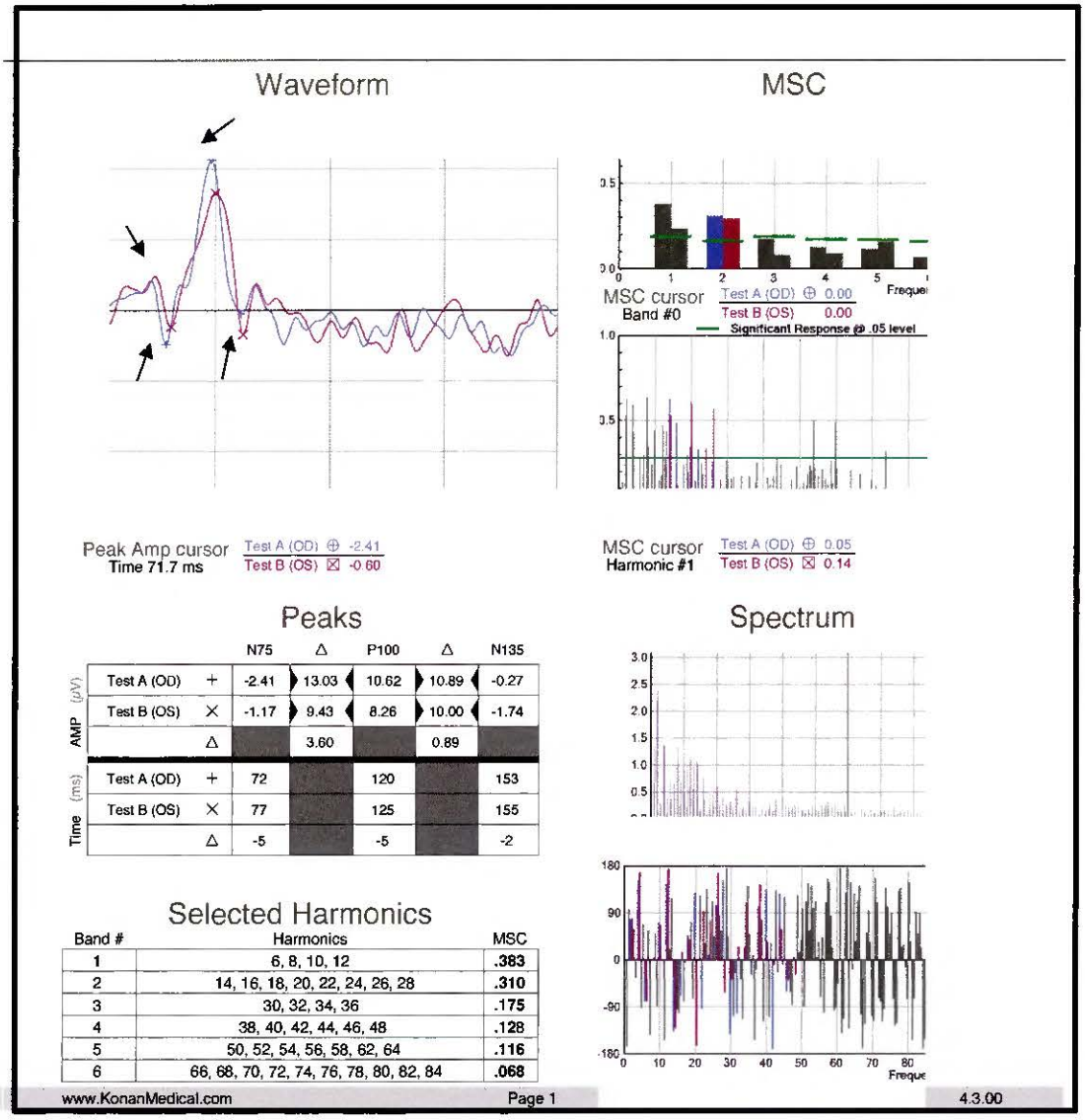


Figure 3B. Representative output from EvokeDx for an MS participant for the large check condition.

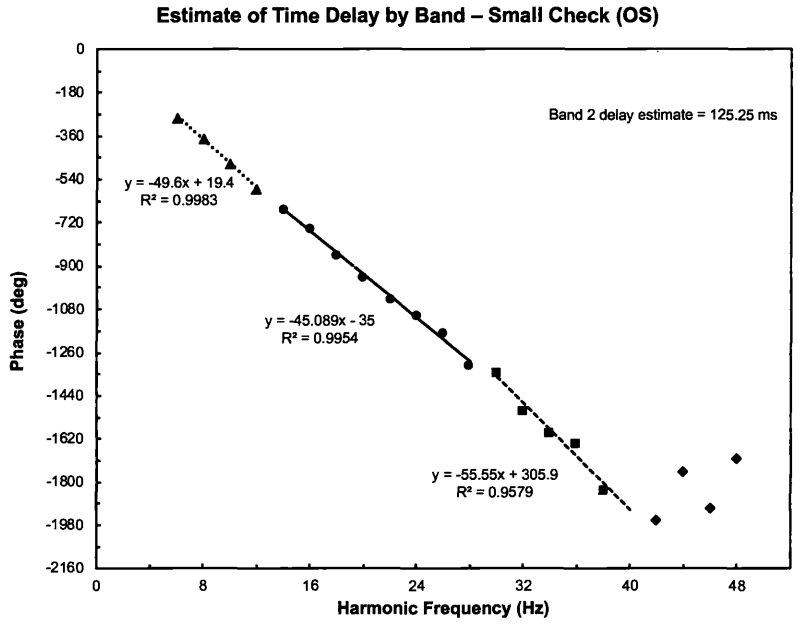


Figure 3C. Representative phase vs. harmonic frequency plot for an MS participant for the small check condition – left eye. Slopes were plotted for Bands 1 through 3 and an estimate for Band 2 delay was calculated based on the derived slope. The estimate of Band 2 delay for this participant is ~51 ms longer than that for the representative control participant.

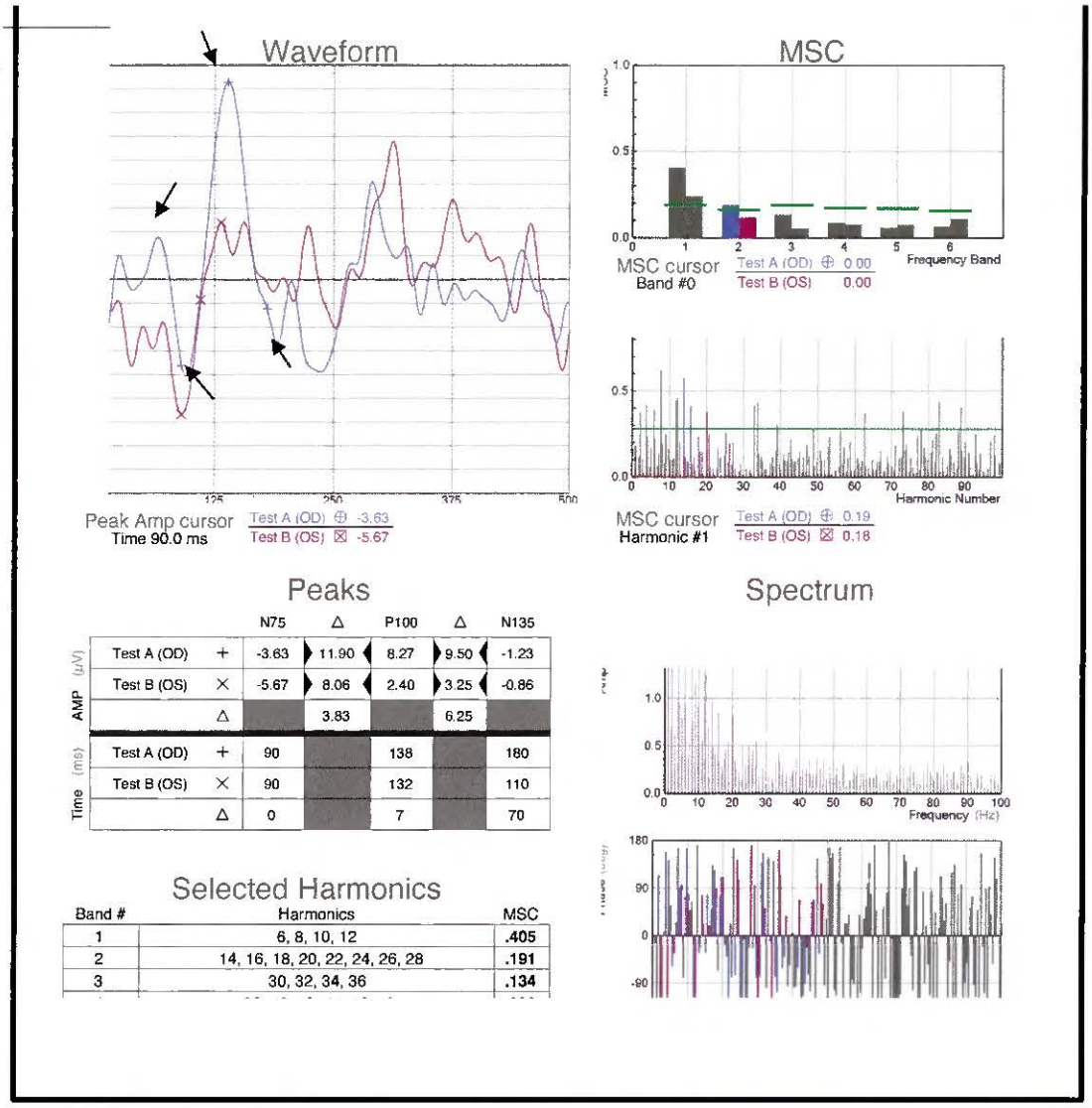


Figure 4. Representative output from EvokeDx for an MS participant with optic neuritis for the small check condition. In the waveform plot at the top left of the image, the blue waveform is the right (unaffected) eye and the purple waveform is left (affected) eye. There is a clear attenuation of the waveform in the left eye.

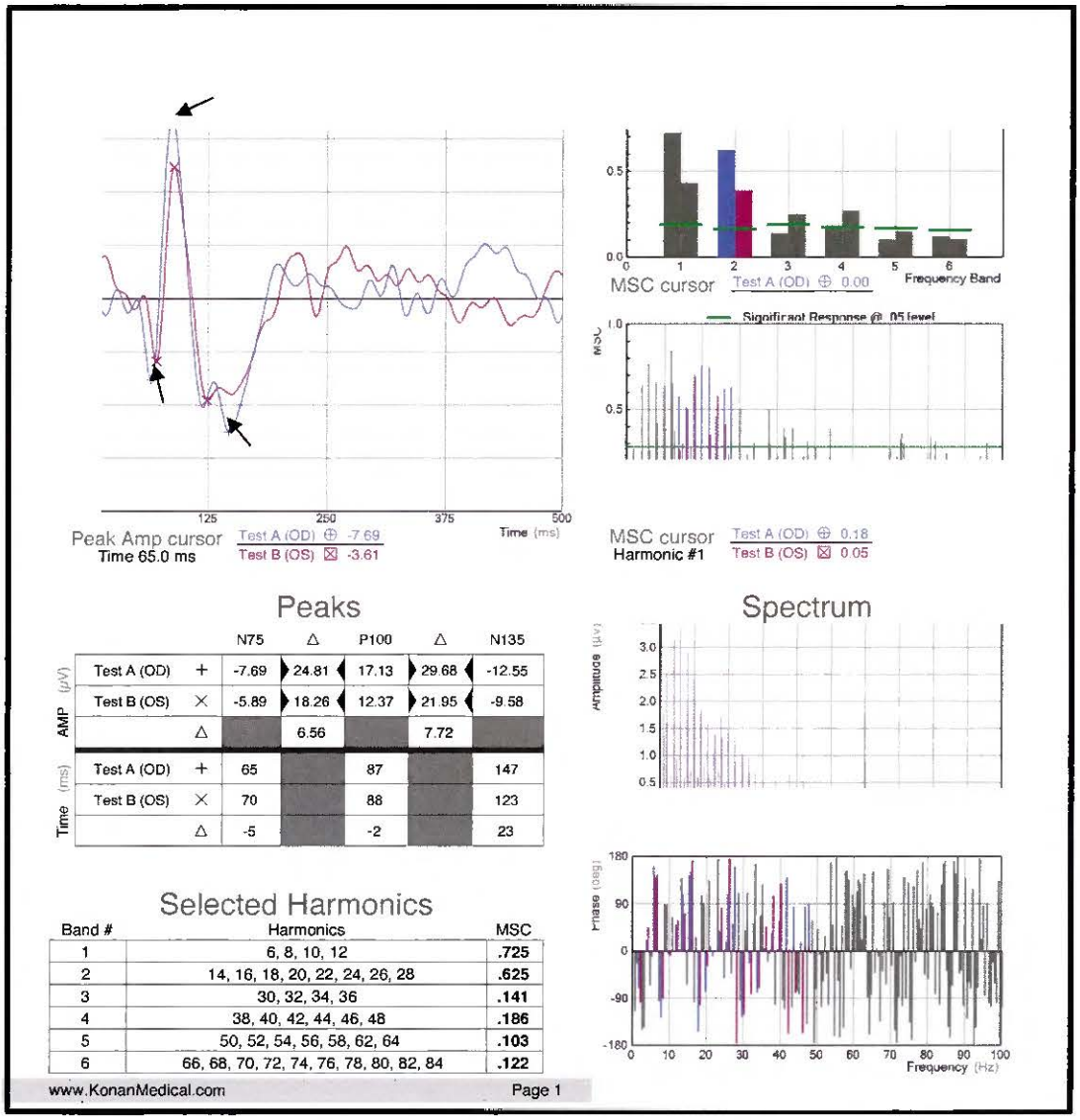


Figure 5. Representative output from EvokeDx for a control participant with medication stimulating GABAergic inhibition for the small check condition. P100 peak is quite prominent in the waveform, as expected given that this positive deflection depends on GABAergic activity, while other peaks appear attenuated and P60 peak is small or undetectable.

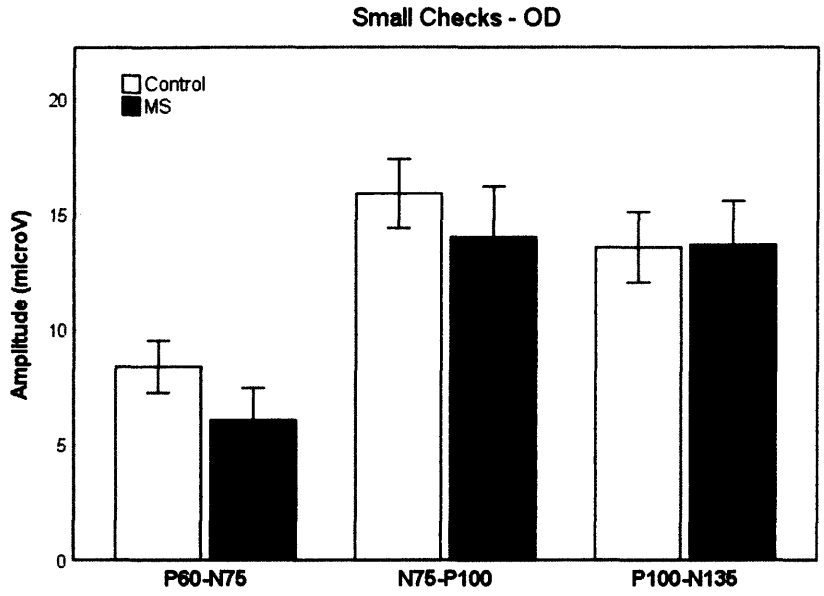


Figure 6A. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the small check condition, right eye. Error bars are ± 1 *SE*.

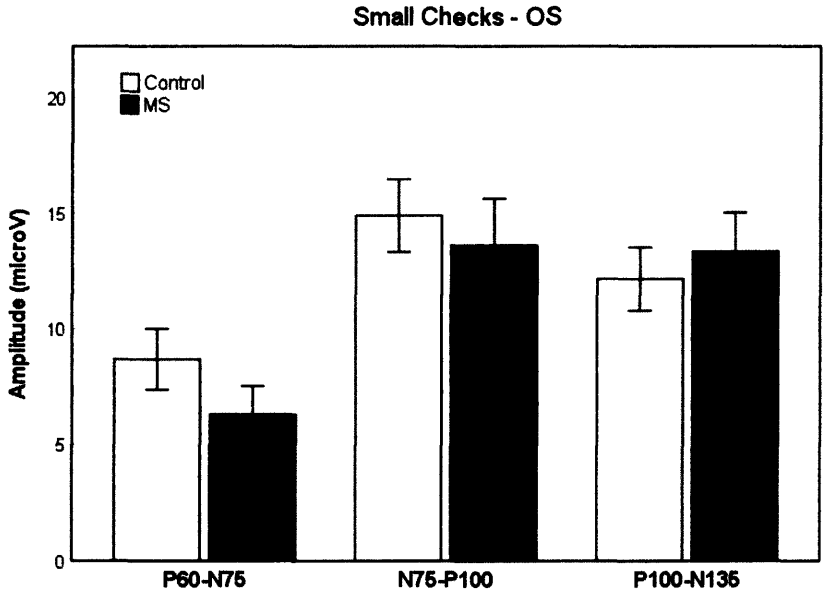


Figure 6B. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the small check condition, left eye. Error bars are ± 1 *SE*.

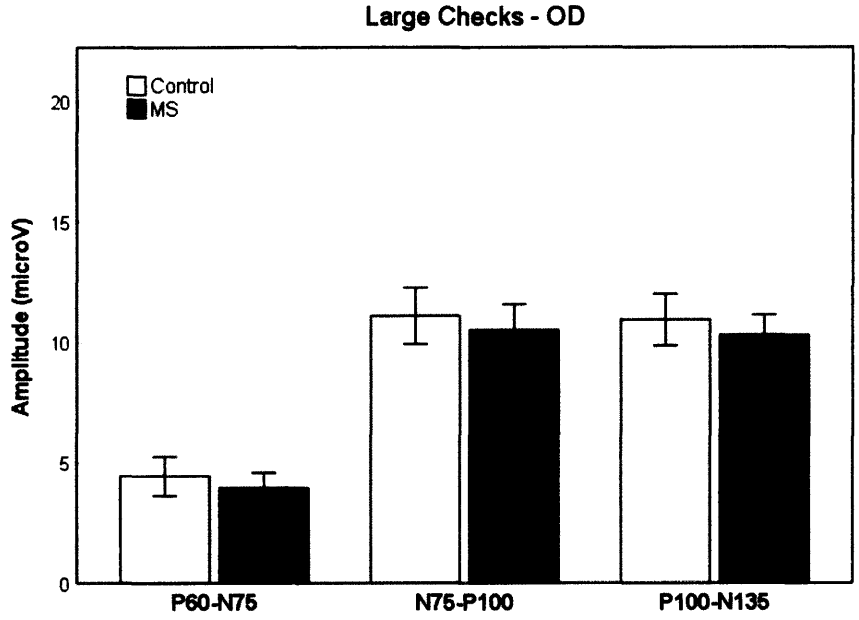


Figure 6C. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the large check condition, right eye. Error bars are ± 1 *SE*.

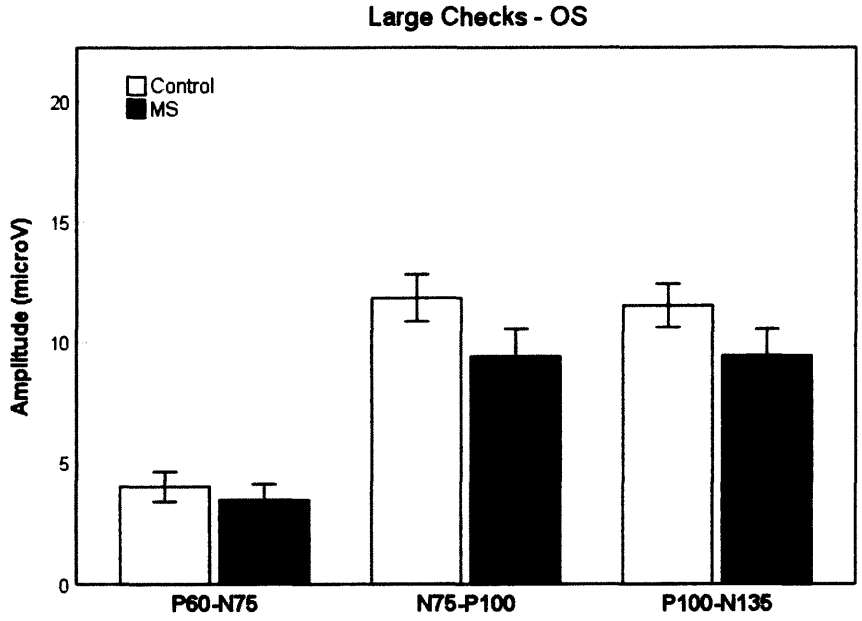


Figure 6D. Peak-to-trough amplitudes (mean amplitudes in microvolts) for the large check condition, left eye. Error bars are ± 1 *SE*.

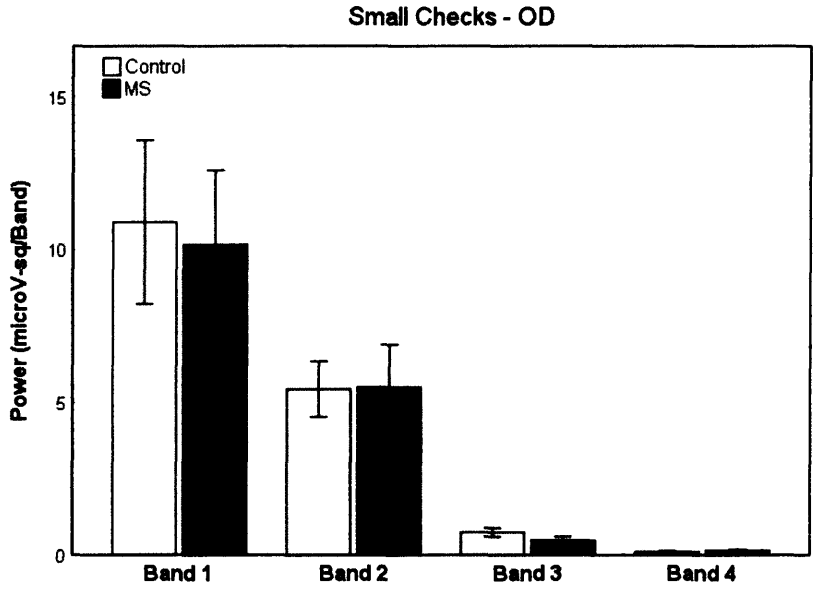


Figure 7A. Power (μV^2/band) plotted as a function of frequency band for the small check condition, right eye. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). Error bars are ± 1 *SE*.

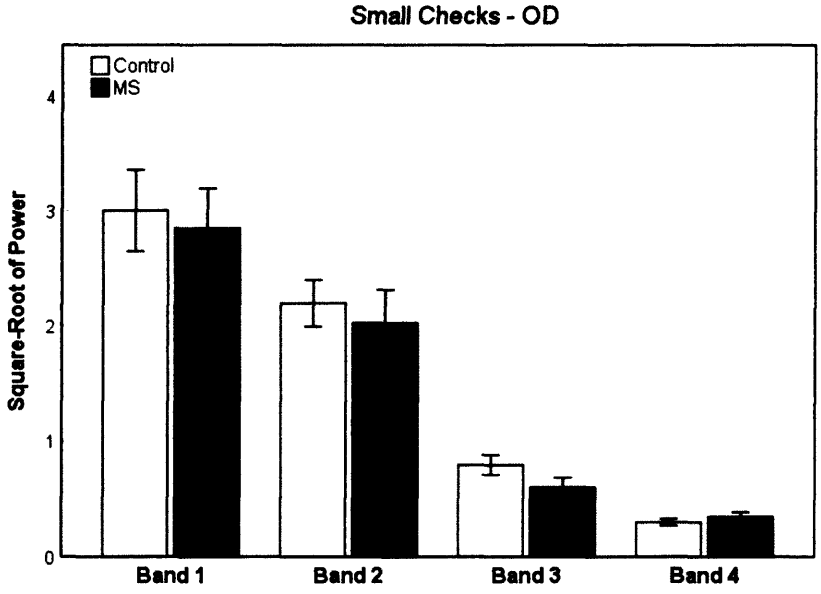


Figure 7B. Square root of power per band plotted as a function of frequency band for the small check condition, right eye. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). Error bars are ± 1 *SE*.

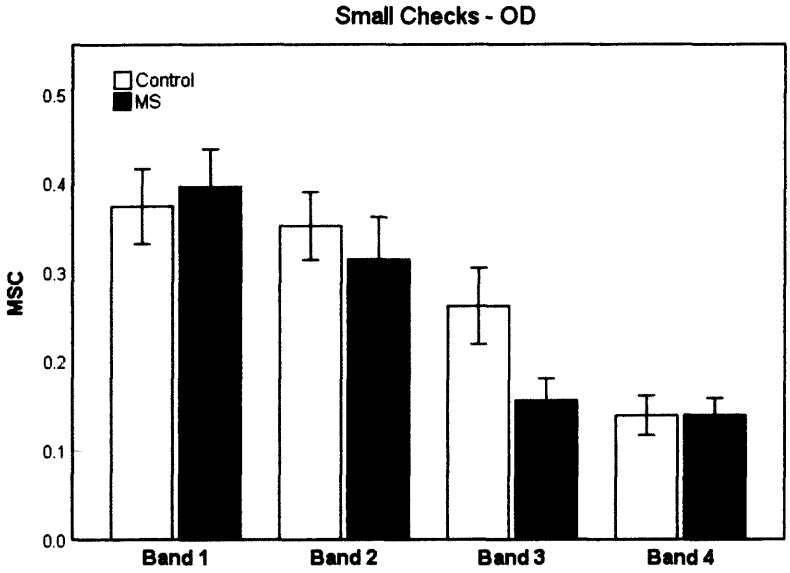


Figure 8A. Mean MSC per band plotted as a function of frequency band for the small check condition, right eye. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). Error bars are ± 1 *SE*.

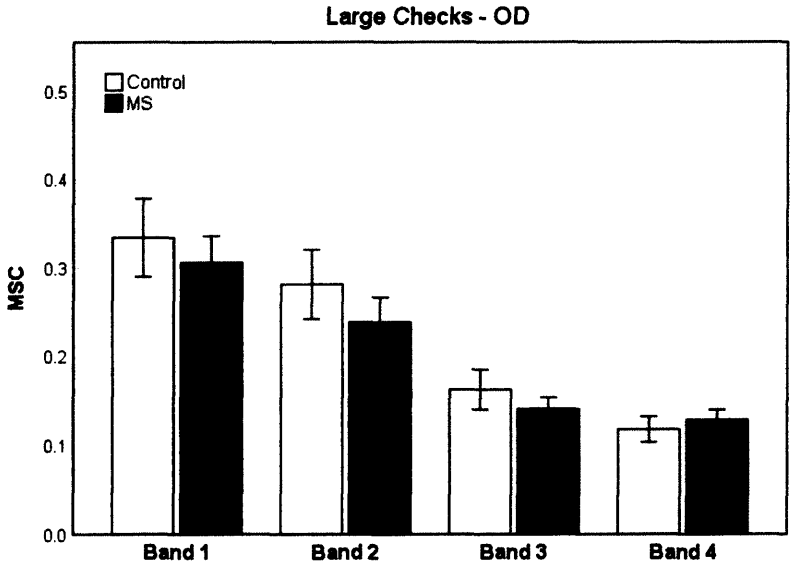


Figure 8B. Mean MSC per band plotted as a function of frequency band for the large check condition, right eye. Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). Error bars are ± 1 *SE*.

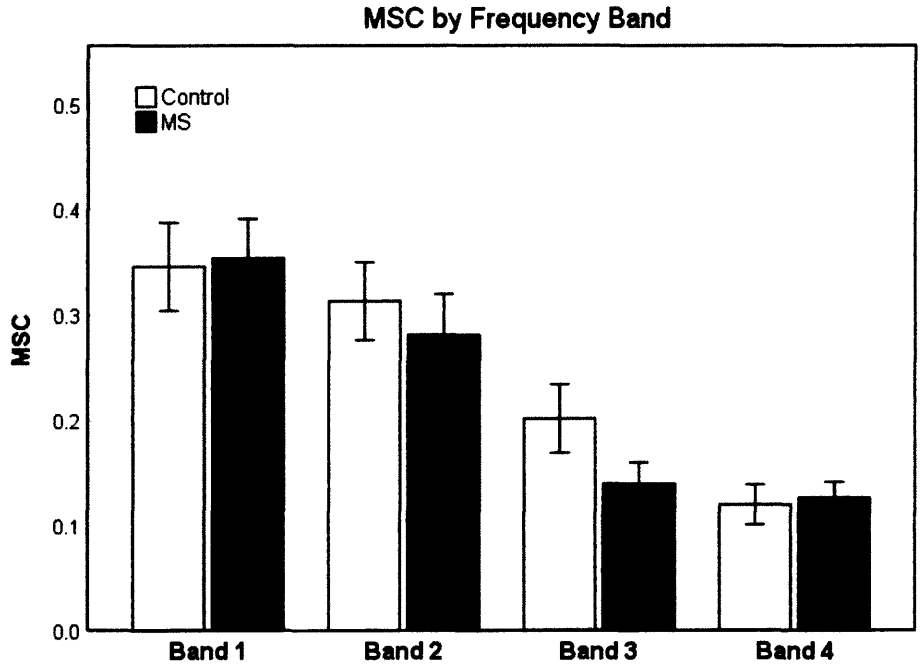


Figure 8C. Mean MSC per band plotted as a function of frequency band aggregated over stimulus conditions (small and large check) and eyes (OD and OS). Bands are defined as follows: Band 1 (6-12 Hz), Band 2 (14-28 Hz), Band 3 (30-40 Hz), Band 4 (42-48 Hz). Error bars are ± 1 *SE*.

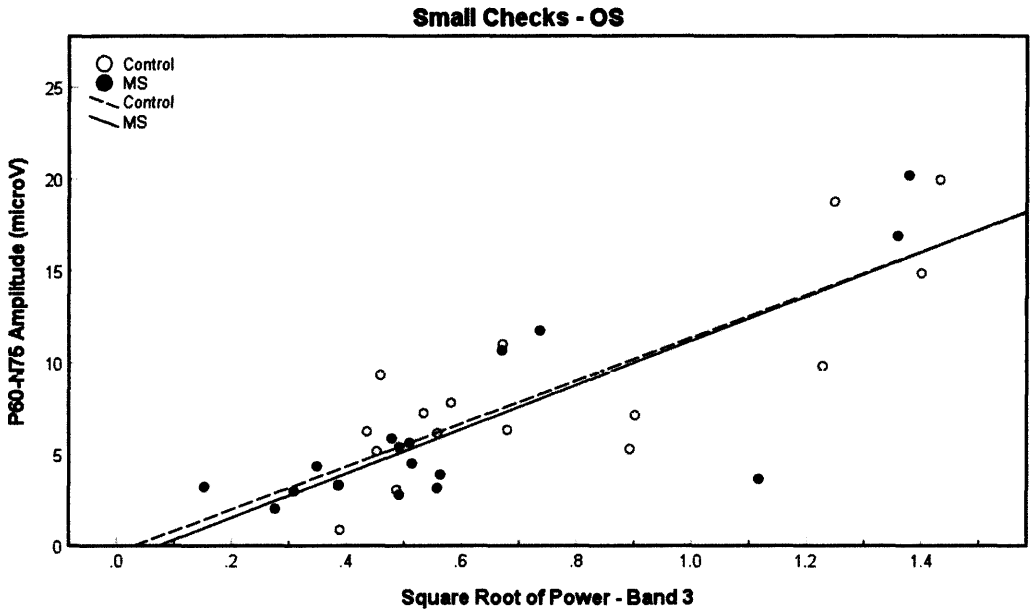


Figure 9. Scatterplot of P60-N75 amplitude vs. square root of power in Band 3 for the small check condition, left eye. Linear relationships between these two measures are moderate and equal between the two groups. Control participants had a regression equation of Y = 0.24 + 9.71*X (R^2 = .66) while MS participants had a regression equation of Y = 0.14 + 10.4*X (R^2 = .66).

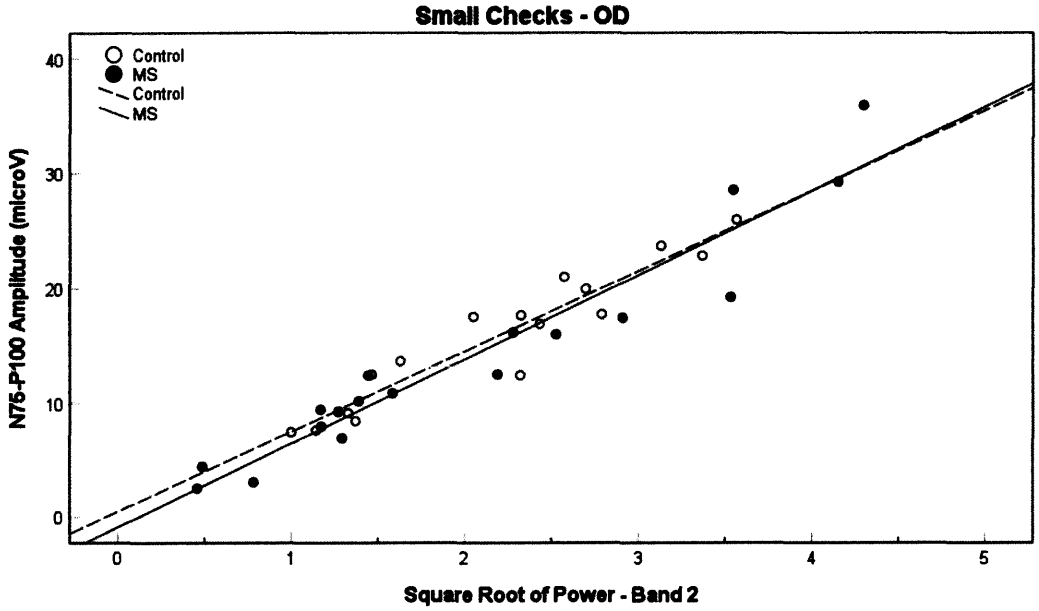


Figure 10. Scatterplot of N75-P100 amplitude by the square root of power in Band 2 for the small check condition, right eye. Both groups exhibit strong linear relations between these measures. The MS participants have a stronger linear relationship in this condition and eye (R^2 =.92, equation: Y = -0.85 + 7.33*X) compared to controls (R^2 =.90, equation: Y = 0.53 + 6.99*X).

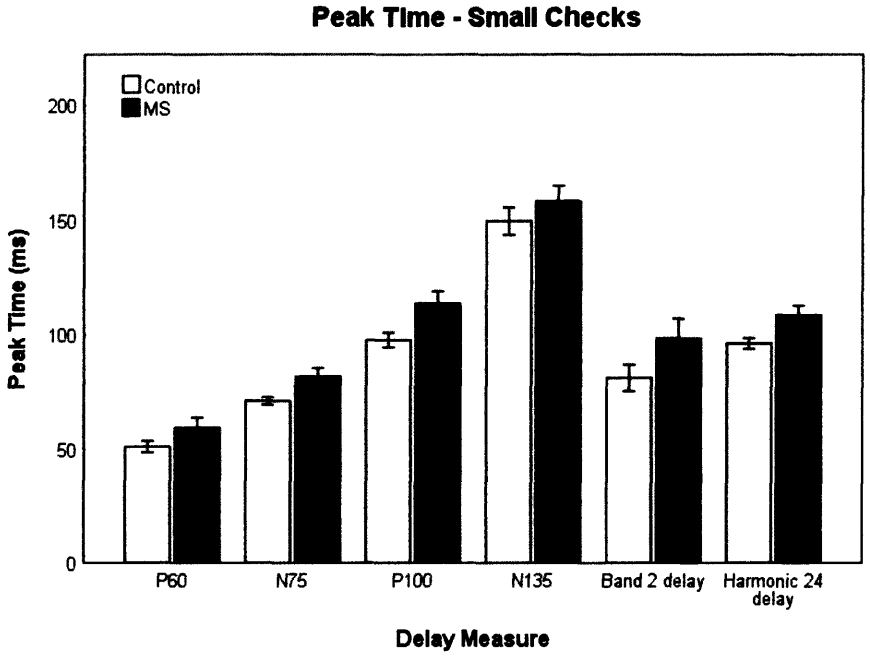


Figure 11A. Mean peak time estimates (ms) for time- and frequency-domain delay measures collapsed over eyes for the small check condition. Error bars are 95% CI.

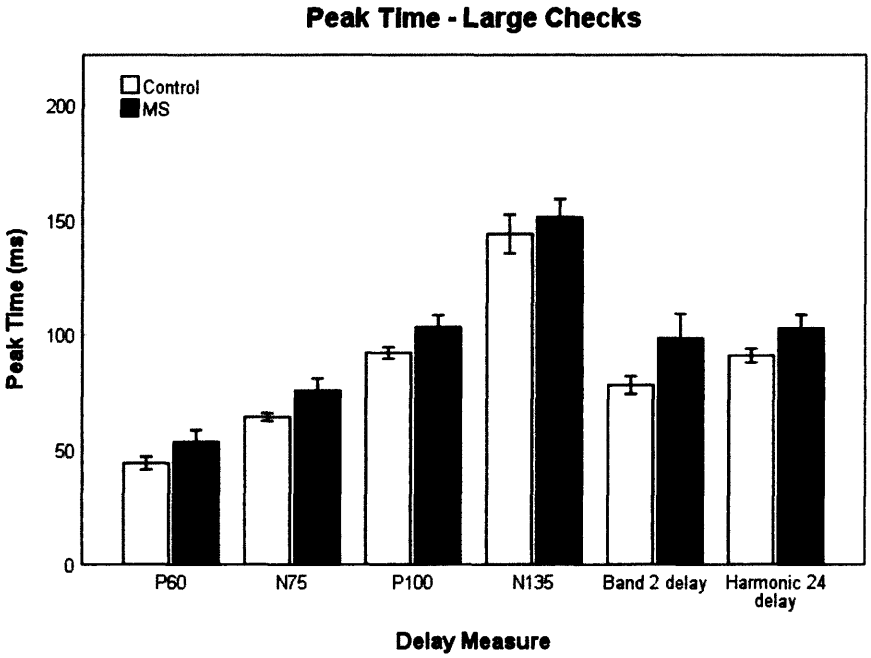


Figure 11B. Mean peak time estimates (ms) for time- and frequency-domain delay measures collapsed over eyes for the large check condition. Error bars are 95% *CI.*

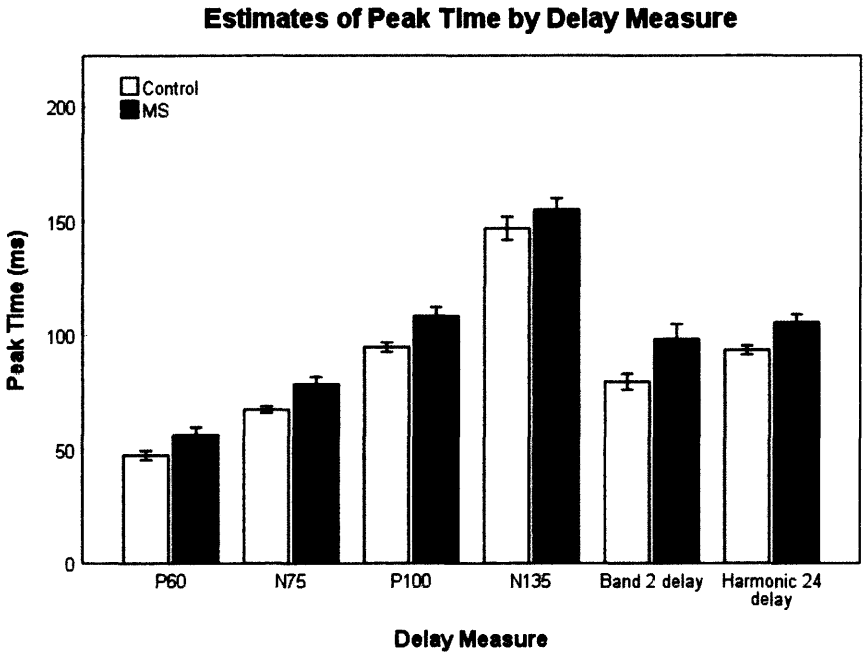


Figure 11C. Mean peak time estimates (ms) for time- and frequency-domain delay measures collapsed across conditions and eyes. Error bars are 95% CI.

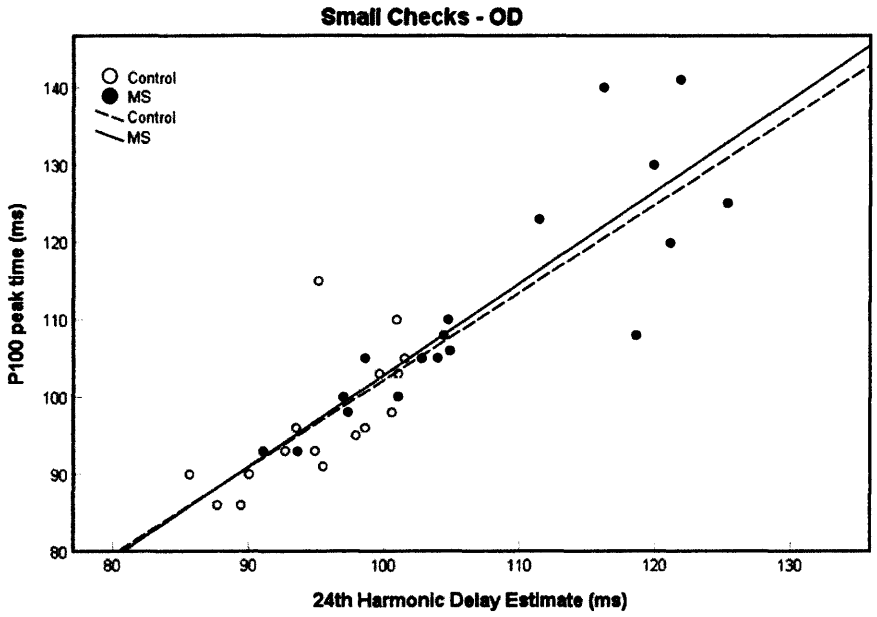


Figure 12A. Scatterplot of P100 peak time vs. 24th harmonic delay estimate for the small check condition, right eye. Both groups exhibit linear relations. The MS participants have a much stronger linear relationship in this condition and eye (R^2 =.73, equation: Y = -11.05 + 1.13*X) compared to controls (R^2 =.48, equation: Y = -15.87 + 1.19*X).

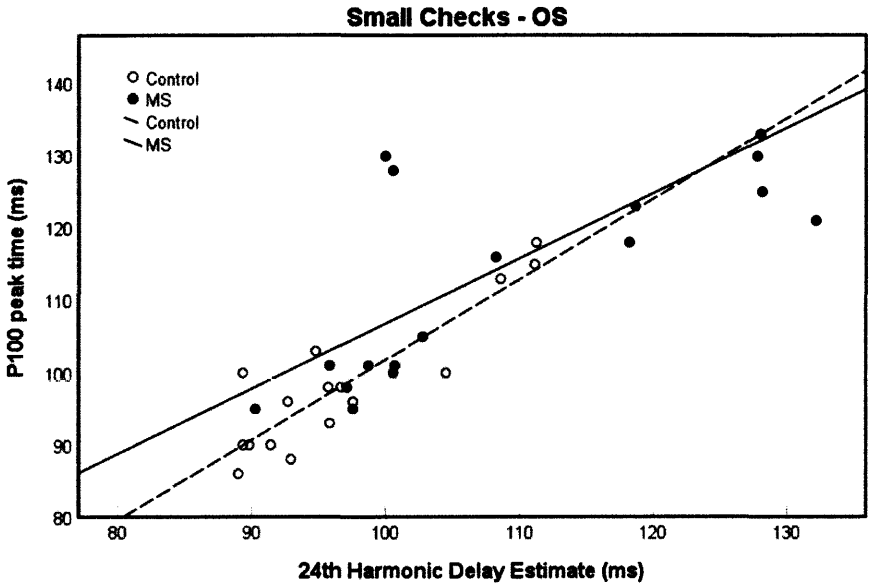


Figure 12B. Scatterplot of P100 peak time vs. 24th harmonic delay estimate for the small check condition, left eye. Both groups exhibit linear relations. The control participants have a stronger linear relationship in this condition and eye (R^2 =.79, equation: Y = 9.52 + 1.11*X) compared to controls (R^2 =.54, equation: Y = 16.59 + 0.9*X).

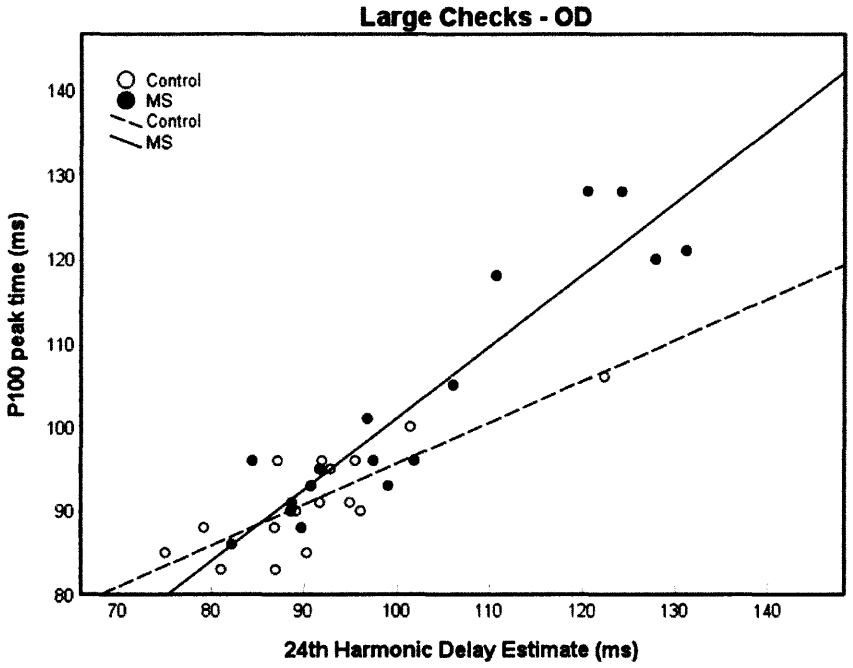


Figure 12C. Scatterplot of P100 peak time vs. 24th harmonic delay estimate for the large check condition, right eye. Both groups exhibit linear relations. The MS participants have a stronger linear relationship in this condition and eye (R^2 =.86, equation: Y = 15.74 + 0.85*X) compared to controls (R^2 =.67, equation: Y = 46.69 + 0.49*X).

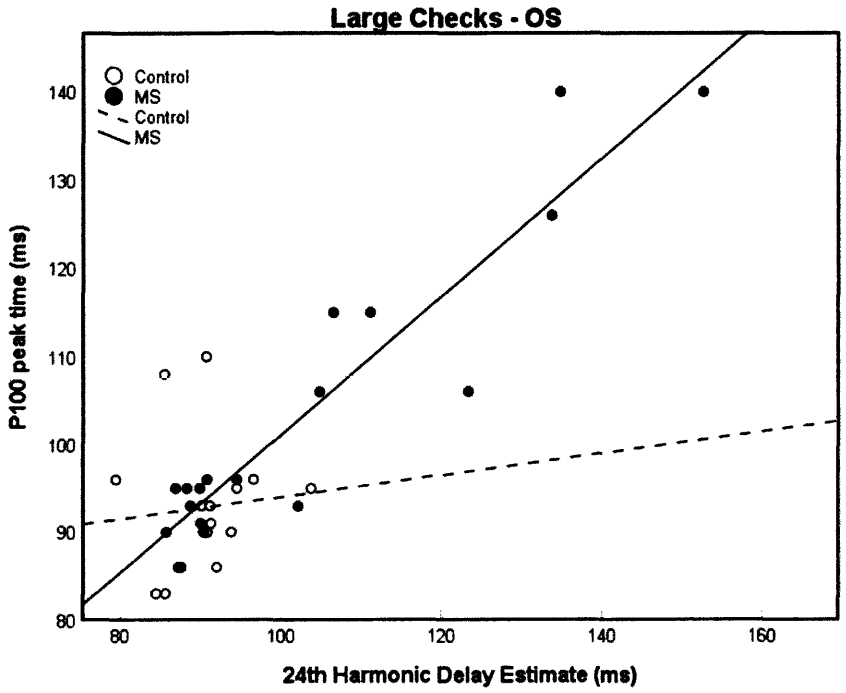


Figure 12D. Scatterplot of P100 peak time vs. 24th harmonic delay estimate for the large check condition, left eye. Both groups exhibit linear relations. The MS participants have a much stronger linear relationship in this condition and eye (R^2 =.88, equation: Y = 922.66 + 0.78*X) compared to controls (R^2 =.01, equation: Y = 81.57 + 0.12*X).

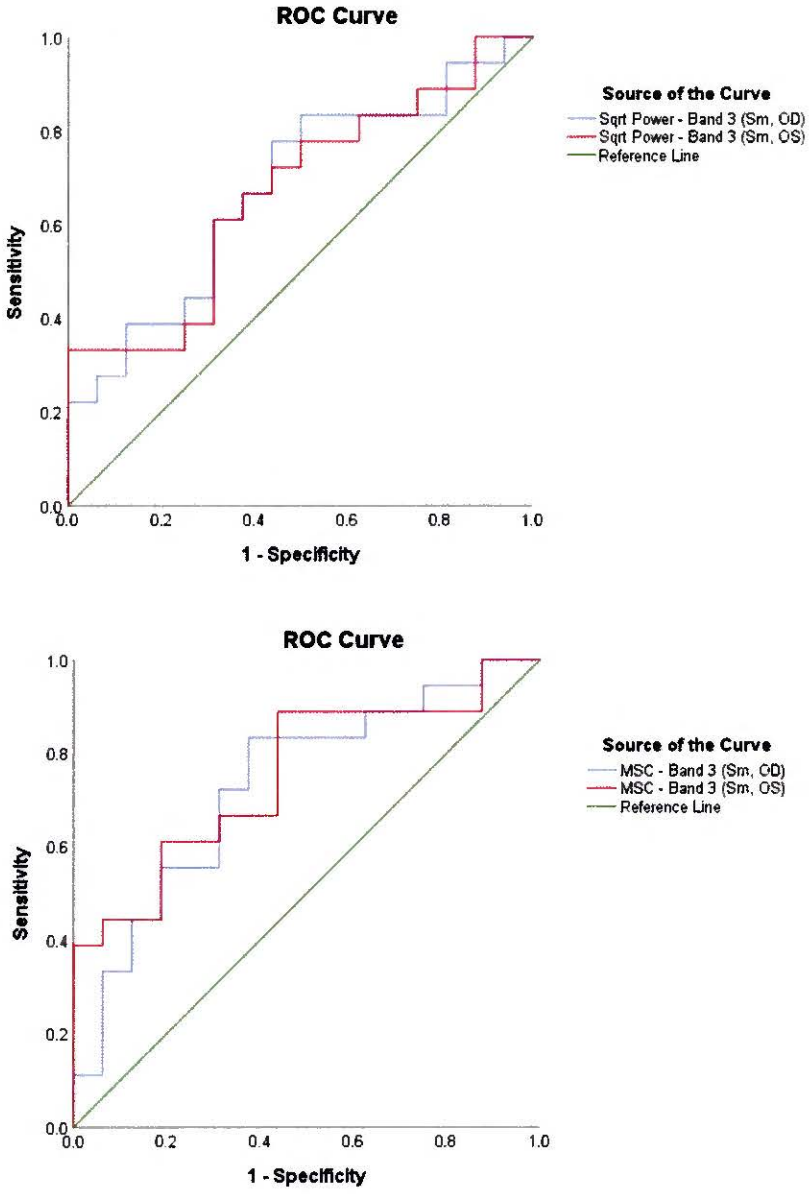


Figure 13. Receiver operating characteristic (ROC) curves of square root of power in Band 3 (top) and MSC in Band 3 (bottom) for small checks for classification of group membership.

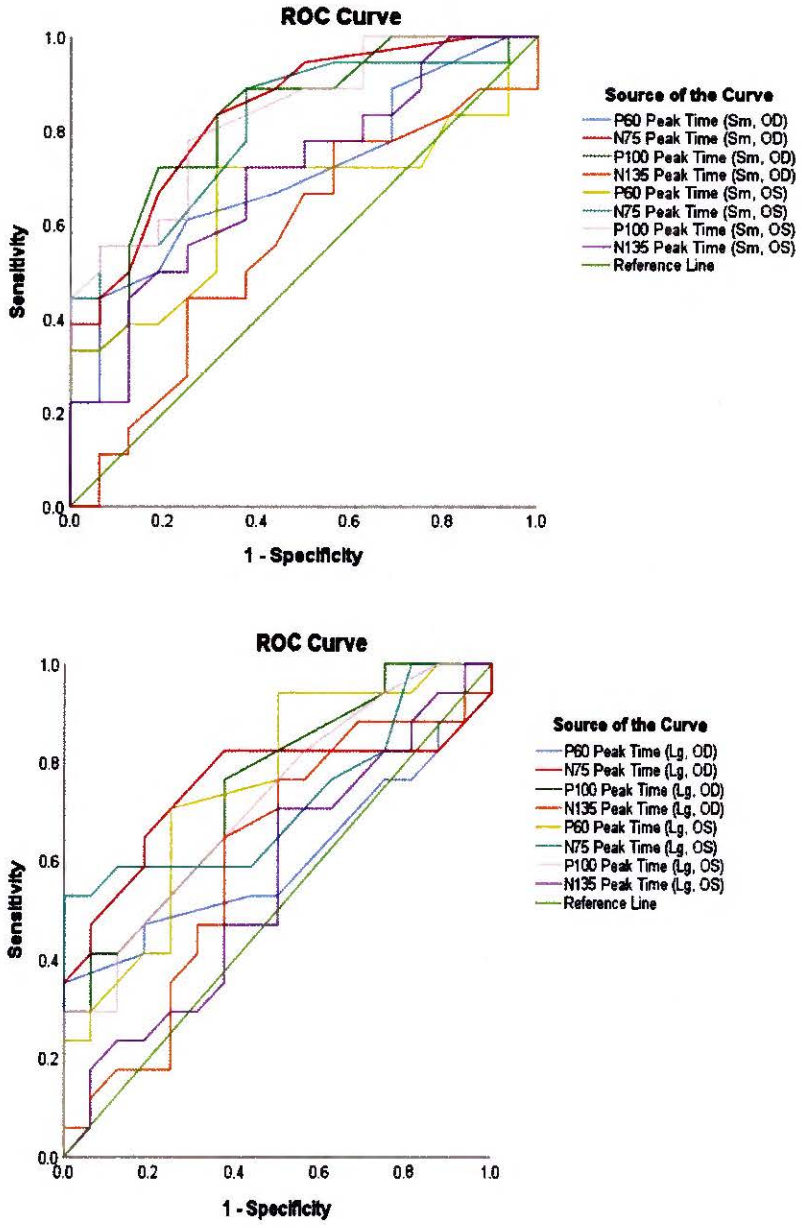


Figure 14. ROC curves for peak time measures for small (top) and large (bottom) checks for classification of group membership.

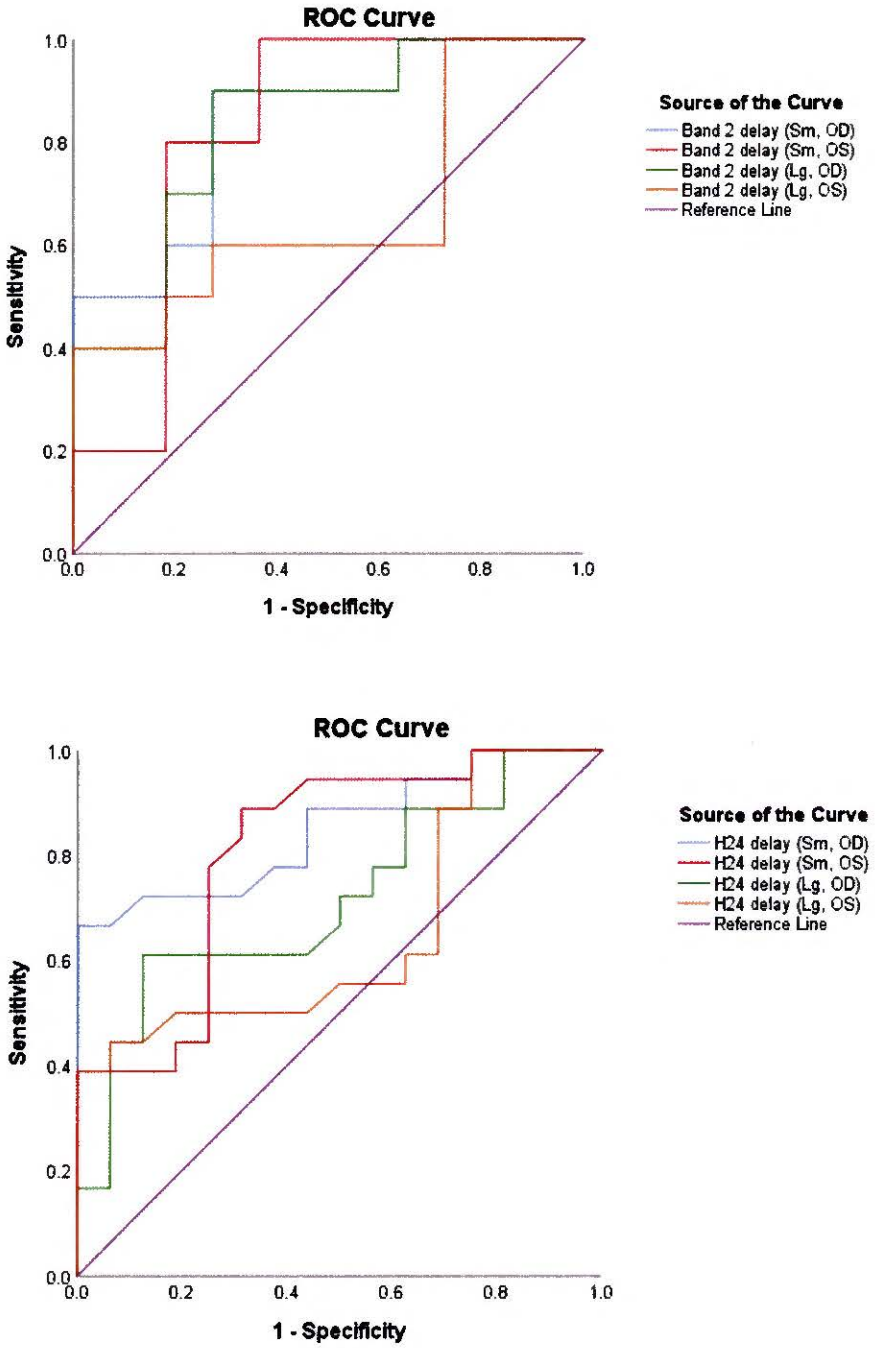


Figure 15. ROC curves for frequency-domain delay measures for small and large checks. ROC curves are for Band 2 delay (top) and H24 delay (bottom) for classification of healthy controls and MS participants. These frequency-domain measures provided greater classification accuracy than did time-domain measures.

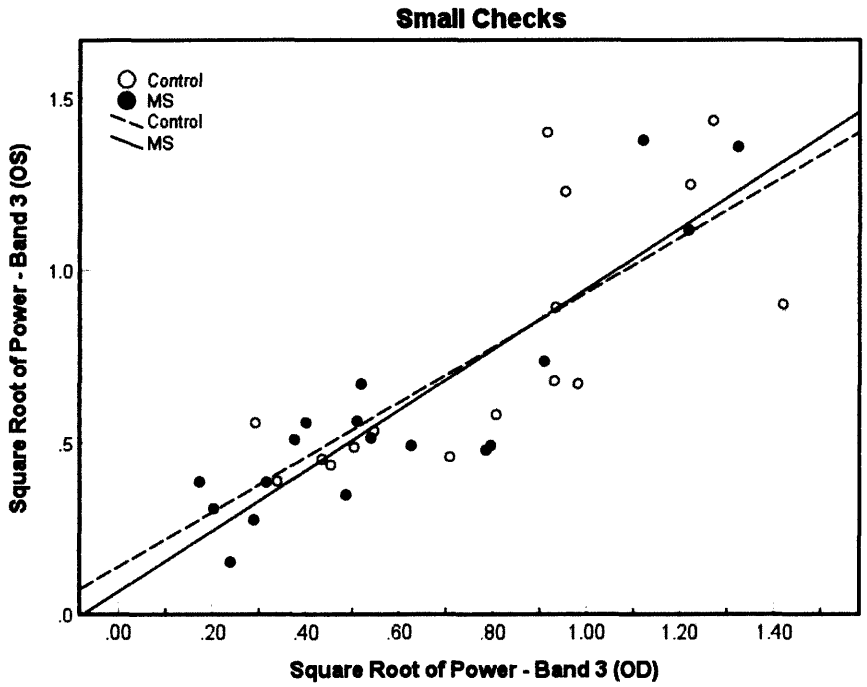


Figure 16A. Interocular scatterplot of square root of power in Band 3 for the small check condition. Both groups exhibit moderate linear relations between eyes. The MS participants have a regression equation of Y = 0.07 + 0.88*X (R^2 = .79) and control participants have a regression equation of Y = 0.14 + 0.8*X (R^2 = .56).

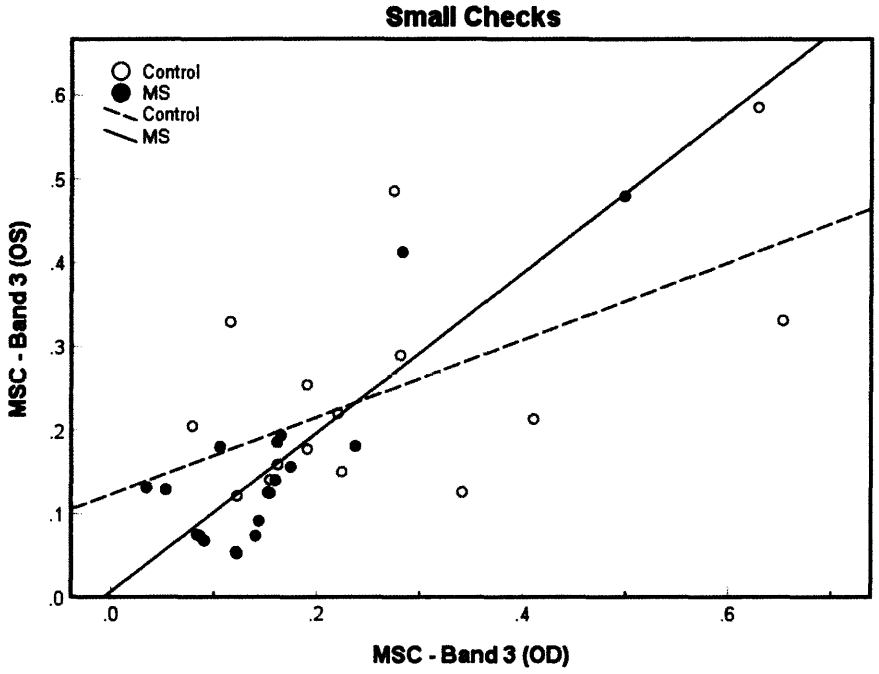


Figure 16B. Interocular scatterplot of MSC in Band 3 for the small check condition. Both groups exhibit moderate linear relations between eyes. The MS participants have a regression equation of Y = 6.49E-3 + 0.95*X (R^2 = .73) and control participants have a regression equation of Y = 0.12 + 0.46*X (R^2 = .73).

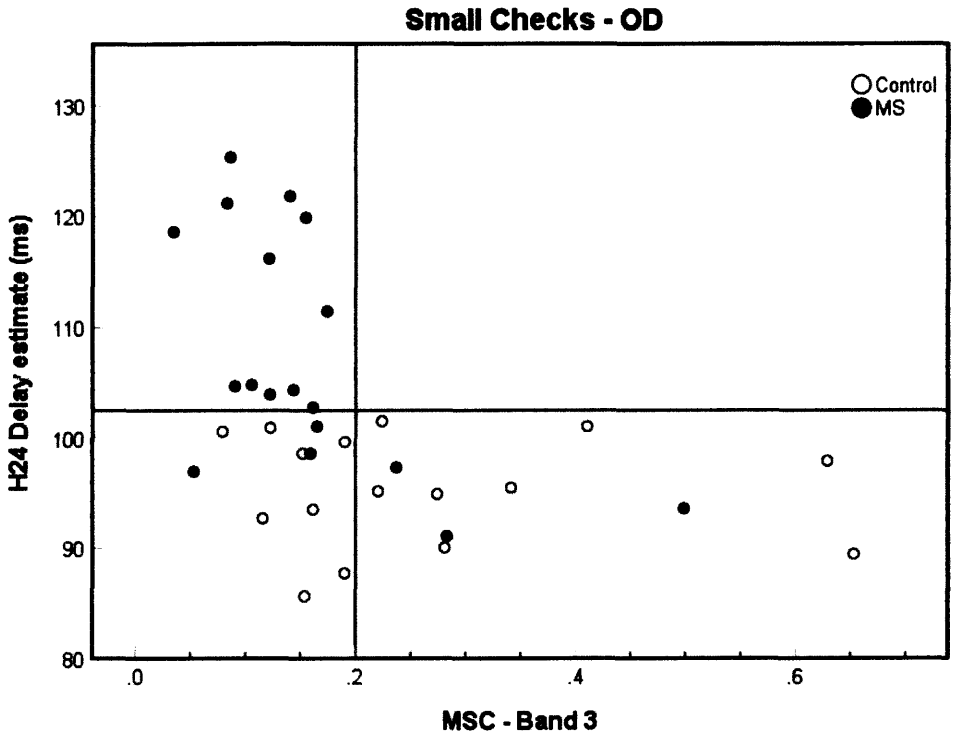


Figure 17A. Classification accuracy using a composite measure of H24 delay and MSC in Band 3 for the small check condition, right eye shows that many in the MS group have long delays and low MSC values. Inclusion criteria for the prediction of MS are indicated by horizontal (H24 delay $\geq$ 101 ms) and vertical (MSC Band 3 $\leq$ .2) reference lines. Note that 12 of 18 patients are classified correctly and that no controls are misclassified.

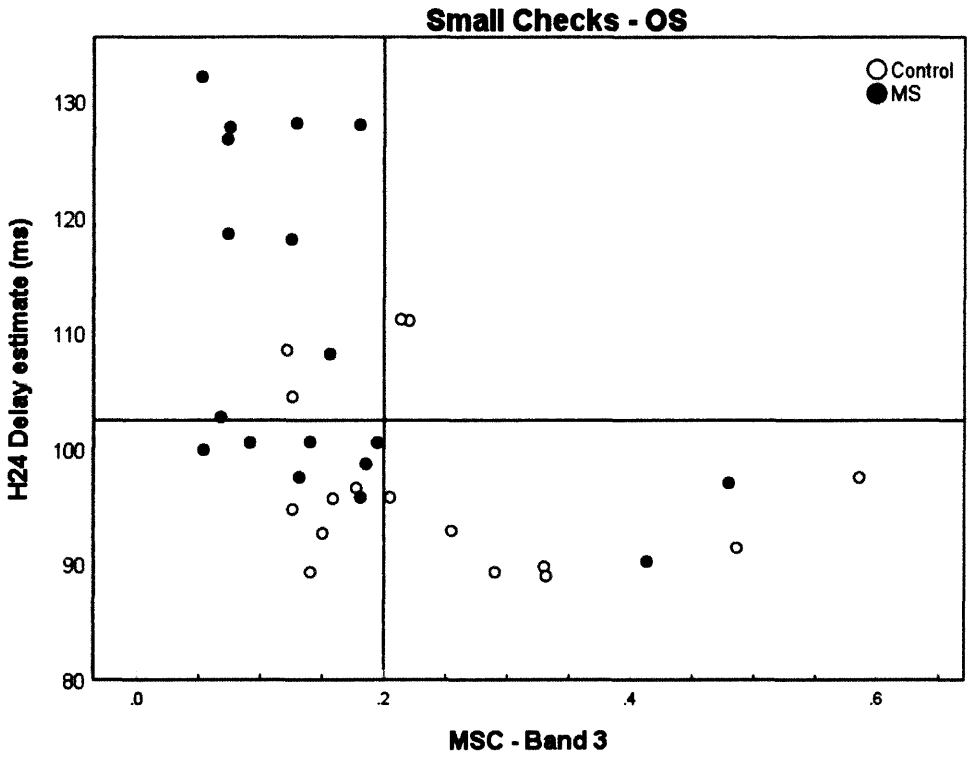


Figure 17B. Classification accuracy using a composite measure of H24 delay and MSC in Band 3 for the small check condition, left eye shows that many in the MS group have long delays and low MSC values. Inclusion criteria for the prediction of MS are indicated by horizontal (H24 delay $\geq$ 101 ms) and vertical (MSC Band 3 $\leq$ .2) reference lines. Note that 9 of 18 patients are classified correctly and that 2 controls are misclassified.

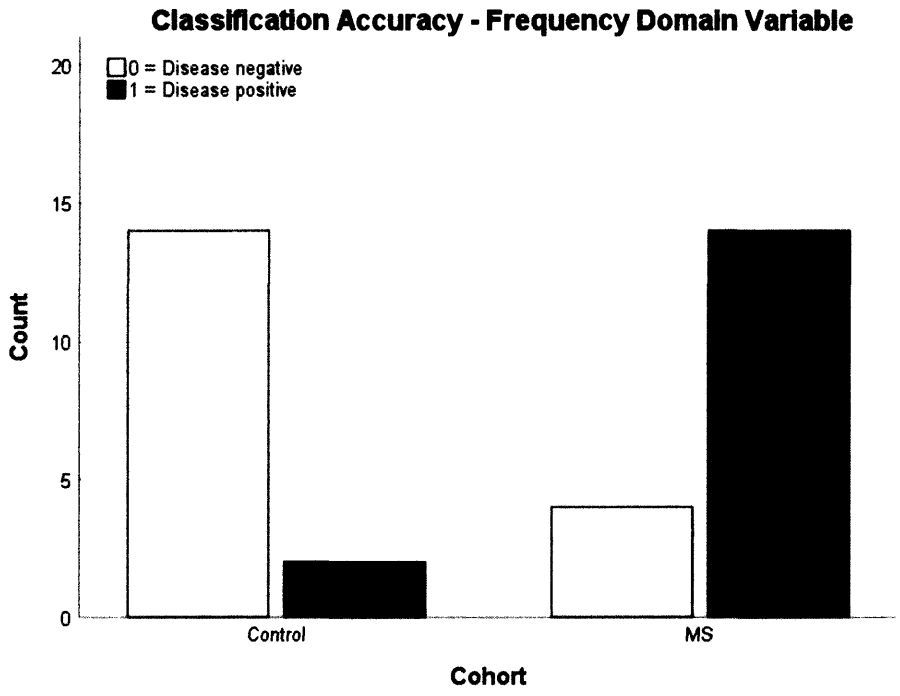


Figure 18A. Classification accuracy using the composite frequency-domain measure. **Frequency-Domain Variable** = H24 delay with MSC-Band 3. Prediction: **Disease negative** = participants with H24 delay ≤ 101 ms and MSC-Band 3 ≥ .2; **Disease positive** = participants with H24 delay ≥ 101 ms and MSC-Band 3 ≤ .2. Using this measure, specificity = 87.5% and sensitivity = 77.8%.

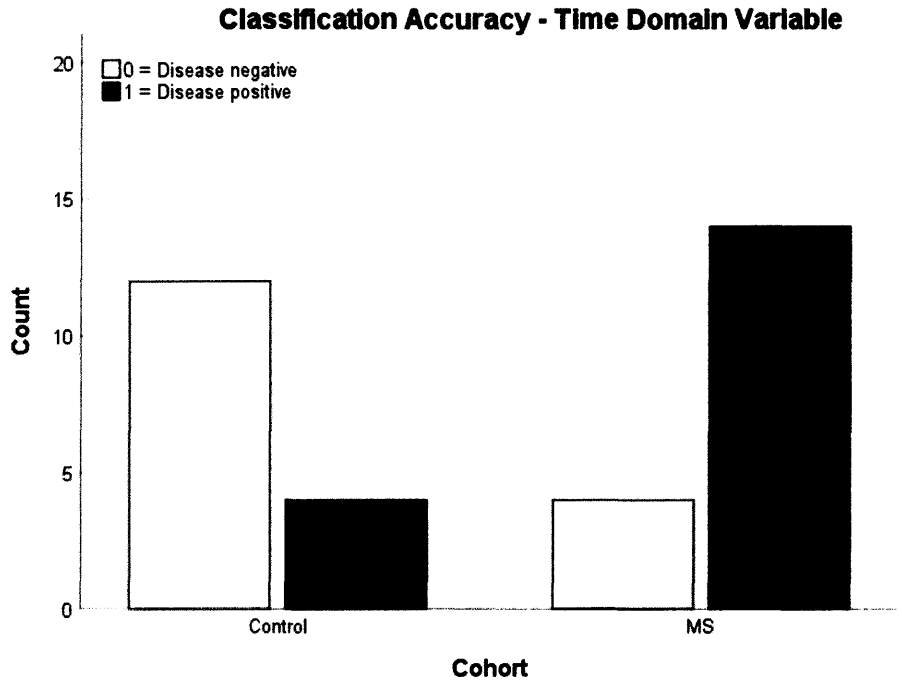


Figure 18B. Classification accuracy using the composite time-domain measure. **Time Domain Variable** = P100 long time-domain variable. **P100 long** = variable created for the small check condition and participants were included if P100 latency was ≥ 104 ms (based on ROC coordinates). Prediction: **Disease negative** = participants with P100 peak time < 104 ms; **Disease positive** = participants with P100 peak time ≥ 104 ms. Using this measure, specificity = 75% and sensitivity = 77.8%.